Search and Matching in the Housing Market[1]

David Genesove[2]

Lu Han[3]

January 5, 2012

Abstract

Housing markets clear partly through the time buyers and sellers spend on the market, and the readiness with which they transact with each other. Applying a random matching model to unique multi-year, multi-market survey data on both buyers and sellers, we examine how demand affects housing market liquidity. We find that buyer time on the market, the number of homes buyers visit, and especially seller time on the market all decrease with demand, with a much greater sensitivity to demand growth than its level. This is consistent with a straightforward matching model with a lag in seller response. Our findings imply that the elasticity of the hazard that any given seller will be contacted by a buyer, with respect to the buyer-seller ratio, is 0.84, assuming a constant returns to scale matching function.

Keywords: real estate, search, matching, liquidity

---

[1]Thanks to Ed Baryla, Paul Bishop, Glenn Crellin, Harold Elder and Len Zumpano for generally providing us with data, codebooks and questionnaires; Travis Chow for excellent research assistance; the editor Stuart Rosenthal; Paul Anglin, Carlos Garriga, Shouyong Shi, and other seminar participants at Toronto, Guelph, Ben-Gurion, Hebrew University, Madison-Wisconsin, UBC and Duke, for their helpful comments.
[2] Hebrew University of Jerusalem and C.E.P.R. Email: genesove@mscc.huji.ac.il.
[3] University of Toronto. Email: lu.han@rotman.utoronto.ca.

1. Introduction

Housing markets are highly illiquid, as it takes time and effort for buyers to find suitable homes and vice versa. Past studies on time on the market have focused entirely on one side of the market – generally sellers, but sometimes buyers. Such information suffices for assessing the determinants of search outcomes among agents on one side of a market. However, an equilibrium analysis requires data from both sides of the market, and across different markets. Perhaps for this reason, no previous paper has empirically investigated the determinants of housing market liquidity.

This paper assembles a unique dataset based on information from both the buyers' and sellers' sides, across geographical markets and over time. With this dataset, we empirically investigate demand shocks as determinants of housing market liquidity, interpreting the patterns we uncover by a straightforward matching model that integrates both buyers' and sellers' search behaviour. We show that a simple dynamic extension to the standard model fits the data well. We find that sellers' search behaviour is affected most, and that of both buyers and sellers is affected much more by a change in demand than by the level itself.

Specifically, we consider the response of three measures of liquidity - seller time on the market (*STOM*), buyer time on the market (*BTOM*), and the number of homes visited by the buyer during his search (*NHVIS*) – to market level demand. Looking at both sides of the market, and interpreting the number of home visits as an inverse proxy for the probability of a transaction given that a buyer and seller meet, allows us to construct the *buyer-seller ratio* and buyers' and sellers' *contact hazards* (the probability one side meets the other), and so better understand why *STOM*, *BTOM* and *NHVIS* respond the way they do. It also allows us to estimate the elasticity of the contact hazards with respect to the buyer-seller ratio, making this paper the first to estimate the matching function for housing.

We aggregate micro data from the National Association of Realtors' (NAR) buyers and sellers surveys to the Metropolitan Statistical Area (MSA) level, for available years from 1987-2008, to form a panel dataset with market-level measures of *BTOM*, *STOM*, and *NHVIS*. These variables are formed from the responses to the questions: 'How long did you actively search before you located the home you recently purchased?' (*BTOM*), 'How long was this home on the market?' (*STOM;* both in weeks), and 'Including the home you purchased, how many homes did you walk through and examine before choosing your home?' (*NHVIS*). We aggregate individuals' responses to the market level (MSA/year), since our focus is on how markets respond to overall changes in demand. Aggregation prevents us from investigating the extent to which individuals react differently to market level variables, but we leave that to future research. The NAR surveys are the only multi-year, multi-market surveys of both buyers' and sellers' search experience, which we need for our market level analysis. However, their response rates are quite low; we explore whether that is likely to induce substantial bias in the estimates, and detail the conditions under which it will not.

We preface the empirical work with a theoretic analysis of the standard constant returns to scale random matching model that is meant to capture the mechanism through which a demand shock affects the housing market liquidity.[4] The analysis turns on which of buyer or seller inflow to the market is more sensitive to the value of search. Assuming that buyers are, the model predicts that an increase in market-wide housing demand decreases *STOM* and *NHVIS*; its effect on *BTOM* is ambiguous, since while the frequency of a buyer's contact with a seller declines, each meeting is more likely to end in a transaction. However, the model's predictions for the buyer (seller) hazard, the log difference between *NHVIS* and *BTOM* (*STOM*), are unambiguous: the buyer (seller) contact hazard falls (increases). Also, the buyer-seller ratio increases. We provide a number of extensions to the model to demonstrate the robustness of these predictions.

[4] Novy-Marx (2007) presents a similar theoretical model. Its main difference from ours is its focus on shifts in the outside option, rather than the match quality distribution. Other theoretical papers that have looked at housing markets through the lens of matching models include Wheaton (1990), Krainer(1997)and Albrecht et. al. (2007). The prevalence of multiple listing services in U.S. real estate markets might be thought to make Coles and Smith's (1996) marketplace model a better fit. However, with market conditions summarized solely by the ratio of the stock of buyers and sellers, random matching requires but a single statistic in an empirical analysis of the *STOM* and *BTOM* distributions. For the marketplace model, buyer and seller inflows matter as well, so at least two statistics are required.

We then turn to the empirical analysis, which is our focus. We first regress *BTOM*, *STOM* and *NHVIS* on our demand proxies, average income and population, controlling for year and market fixed effects. In principle, we would like to use consumption and production amenities as demand indicators but difficulties in measuring yearly variations in them force us to rely on income and population instead (see Section 3). Consistent with the basic matching model, we find that larger population is strongly associated with shorter *BTOM* and *STOM* and fewer *NHVIS*, with the effect on *STOM* greatest, and that on *BTOM* least. Income follows the same pattern, although the relationship is significant only for *STOM*.

Having found that demand affects market liquidity, we then consider how it affects the buyer-seller ratio, and the buyer and seller contact hazards. These variables are not directly observed. We construct them as linear combinations of *STOM*, *BTOM* and *NHVIS* that are derived from the model. Our estimates show that income is positively associated with both the buyer-seller ratio and the seller contact hazard, and negatively with the buyer contact hazard, all of which is consistent with the model.

We also estimate a dynamic specification that distinguishes between a demand shock's short and long run effects. We find that the short run effects on *STOM* and *NHVIS* of an income or population increase are substantive and negative; in contrast, the long run effects, which are also negative, are an order of magnitude smaller and sometimes insignificant. *BTOM* displays the same, but much more muted, pattern, with relatively small and insignificant short run effects.

These findings are consistent with sellers reacting to demand with a lag and an effectively negligible interest rate. We find supporting evidence in, respectively, the behaviour of the sales to list price ratio and a comparison of mean *STOM* to the discount rate implied by Genesove and Mayer (1997). Our results provide an instrumental variables estimate of the elasticity of the contact hazards with respect to the buyer-seller ratio, specifically a seller contact hazard elasticity of 0.84 (equivalently, a buyer contact hazard of -0.16). Thus a doubling of the buyer-seller ratio would increase the likelihood of any given seller being contacted by a buyer by 79 percent. We argue that such a large seller contact hazard elasticity is consistent with the seller listing institutions in U.S. real estate markets.

To conclude, we find the effects of demand shock on market liquidity – in terms of home buyers' and sellers' search behaviour -- well captured by a simple matching model. We assess a number of alternative explanations - changes in matching technology and search effort, the dynamic response from housing supply, and measurement error in *BTOM* – but none alter our conclusion.

This paper is related to a large body of work on time on the market in the housing markets. On the sellers' side, studies have explored determinants of *STOM*, such as idiosyncrasy of the property (Haurin, 1988), seller motivations (Glower et al., 1998), initial offer price (Anglin, Rutherford and Springer, 2003), owner equity (Genesove and Mayer, 1997), previous purchase price (Genesove and Mayer, 2001), initial list price (Anglin, Rutherford and Springer, 2003) and brokers (Levitt and Syverson,2005; Hendel, Nevo and Ortalo-Magne, 2009; and Bernheim and Meer, 2008). On the buyers' side, the literature is much more limited, comprising perhaps only Baryla and Zumpano (1995), Anglin (1997), Elder, Zumpano and Baryla (1999, 2000), D'Urso (2002), and Anglin (1994). These studies typically focus on individual level determinants of either *BTOM* or *STOM* and make no cross-market comparisons.[5] In contrast, our work provides an equilibrium analysis of *market* determinants of *both* buyer and seller behaviour, using data *across* geographical markets and time.

2. The Model

2.1 Baseline Model

We begin with a matching model to guide the subsequent empirical analysis. It is designed to be simple and to generate transparent predictions. Section 2.2 develops a dynamic extension. Section 2.3 assesses the baseline model's robustness by relaxing a number of its assumptions.

The matching function $m(B,S)$ maps the number of buyers (B) and sellers (S) to the number of contacts between them. We make the usual constant returns to scale assumption that the probability that

[5] Merlo and Ortalo-Magne (2004) consider four markets, defined by two areas and two periods. D'Urso (2002) and Baryla, Zumpano and Elder (1999, 2000) examine housing units located in different geographical markets. There is also work on the determinants of bargaining power between buyer and seller (Harding, Rosenthal, and Sirmans, 2003); it, too, is based on individual characteristics, and not market level factors.

an agent makes a contact with the other side depends only on the buyer-seller ratio, denoted by $\theta \equiv B/S$. Thus a given seller makes contact with a buyer with probability $q(\theta) \equiv m(B,S)/S = m(\theta,1)$, while a given buyer contacts a seller with probability $h(\theta) \equiv m(B,S)/B = m(1,1/\theta)$. Thus $q \equiv h\theta$. Standard and intuitive assumptions on m imply that $q'(\theta) > 0$ and $h'(\theta) < 0$.

The net present value of a given buyer owning a given home is X, a random variable with distribution $1 - G(X - v)$. The distribution G is assumed known to both buyers and sellers.[6] Shifts in the parameter v represent movements in demand. X is idiosyncratic to the buyer-home match; its value says nothing about the value to the same buyer of any other home, or the value to any other buyer of owning that home. X is unknown to both buyer and seller before contact, but observed by both upon contact. If there is no transaction, both buyer and seller continue searching.

V^B (V^S) denotes the value of continued searching for the buyer (seller). We assume efficient bargaining, so there is a transaction if and only if the value of owning the home exceeds the sum of the value of searching for buyer and seller, i.e., $X \geq V^B + V^S \equiv y$. The probability of a transaction given a meeting is therefore $G(y - v)$. The expected surplus of a consummated transaction is $E[X|X \geq y] - y$, where E is the expectation operator for $1 - G(X - v)$.

We assume Nash bargaining, with the seller obtaining β of the surplus. Thus the price is $P = V^S + \beta(X - y)$. Sellers face a cost c^S of search; buyers, c^B. For interest rate r, the asset equations for the value of the seller's search and the buyer's search are

(1) $rV^S = -c^S + q(\theta)\beta G(y - v)(E[X|X \geq y] - y)$

(2) $rV^B = -c^B + h(\theta)(1 - \beta)G(y - v)(E[X|X \geq y] - y)$

[6] Thus sellers know the value of their home, so we are abstracting from that sort of learning, which is discussed in Lazear (1986) and Sass (1988), although our motivation for the dynamic analysis of Section 2.2. has sellers' learning about the market more slowly than buyers.

With two equations in three unknowns (V^S, V^B and θ), we need an additional restriction to complete the model. This baseline model assumes an infinite supply of buyers at $\bar{V}^B$. Implicitly, it assumes that buyers have a large number of markets to choose among, while sellers are tied to a specific one.

Section 2.3 presents a variant model that assumes that the inflow of buyers is more sensitive to the value of search than is the seller inflow. It yields the same qualitative results. We assume a greater buyer inflow sensitivity because we think this assumption reasonable for housing markets. Nonetheless, it is easy to see what the model predicts if seller inflow sensitivity is greater. As it turns out, the first assumption generates predictions consistent with the data, whereas the second assumption does not.

With V^B a constant, one can rewrite the above equations as

$$(3)\quad ry = -(c^S - r\bar{V}^B) + q(\theta)\beta G(y - v)(E[X|X \geq y] - y)$$

$$(4)\quad 0 = -(c^B + r\bar{V}^B) + h(\theta)(1 - \beta)G(y - v)(E[X|X \geq y] - y)$$

Each equation can be interpreted as the asset equation of a hypothetical searcher with offer distribution $1 - G$ and optimal reservation value y. In equation (3), the search cost is $c^S - r\bar{V}^B$, the interest rate is r, and the offer arrival rate is $q(\theta)\beta$. In equation (4), the search cost is $c^B + r\bar{V}^B$, the interest rate zero, and the offer arrival rate $h(\theta)(1 - \beta)$. Since $q'(\theta) > 0$, the solution in y of (3) is upward sloping in θ (Mortensen, 1986); call that curve $\mathcal{S}$. Since $h'(\theta) < 0$, the solution to (4) in y (the curve $\mathcal{B}$) is downward sloping in θ. Where $\mathcal{S}$ and $\mathcal{B}$ cross defines the unique equilibrium y and θ (see Figure 1).[7]

It is easy to see that an increase in v shifts $\mathcal{B}$ up one for one but shifts $\mathcal{S}$ up less than one for one, unless $r = 0$ (Mortensen, 1986, p. 864). This implies that a demand shock increases equilibrium θ and y, but decreases $y - v$, as in Figure 2.

[7] For existence it suffices that (a) $q(\theta) \longrightarrow 0$ as $\theta \to \infty$ and (b) a bounded support of $1 - G$. (b) can be replaced by $1 - G$ a Generalized Pareto distribution, with shape parameter $c > -1$, which permits it to be unbounded.

Intuitively, $\mathcal{B}$ shifts up one for one since, for given θ, the expected surplus value, i.e. $G(y - \nu)(E[X|X \geq y] - y)$, must remain the same to ensure a constant V^B. This can only be ensured by the threshold value y adjusting fully to the new demand, leaving $y - \nu$, and so the acceptance rate $G(y - \nu)$ unchanged. But full adjustment of y at the initial θ would, by (1), leave V^S unchanged as well. That would mean, in turn, that those positive surplus transactions with match quality values of X exceeding the unchanged sum of buyer and seller search values but falling below the new threshold y would remain unconsummated. This contradicts efficient bargaining. Thus, $\mathcal{S}$ must shift up less.

The consequences of a demand shock for market liquidity are thus as follows:

- Since $y - \nu$ falls, the acceptance rate rises. This is the main prediction of an endogenous acceptance rate model, or 'stochastic matching' (Pissarides 2000, ch. 5). It implies that *NHVIS* falls.
- On the seller side, an increased θ implies that the contact hazard $q(\theta)$ increases. Since the acceptance rate also increases, the probability of selling $q(\theta)G(y - \nu)$ increases. So *STOM* decreases.
- For buyers, the outcome is ambiguous: the contact hazard $h(\theta)$ falls, but the purchase probability $h(\theta)G(y - v)$ may go either up or down. Which effect dominates depends upon G's shape.[8]

The size of these responses depends on the interest rate. For $r = 0$, a positive demand shock shifts up $\mathcal{S}$ one for one as well, so that y increases by the same amount as ν and θ remains unchanged - as will therefore all the other variables. Obviously, the changes will be small for a small positive r.

2.2 Lagged Seller Response

As will been seen, some dynamics are necessary to rationalize our empirical results. As a simple dynamic extension, say that sellers only react to shifts in $1 - G$ with a lag. This is reasonable: when buyers arrive with large offers, sellers will think at first that these are simply large offers from an

[8]The change in $h(\theta)G(y - v)$ is $h'G\frac{d\theta}{dv} + hG'\left\{\frac{dy}{dv} - 1\right\} = h'G\frac{r/\beta}{\Delta}hG + hG'\left(\frac{h^2AG}{\Delta} - 1\right) = \frac{r/\beta}{\Delta}h'hG^2\left\{1 + \right.$ $G'G(EXX \geq y - y)$, where $\Delta = A(-h'.(r\beta) + h2G)$ and $A \equiv Gy - \nu EX|X \geq y - y$. For the Generalized Pareto Distribution (GPD) $G(X; c, k, v) \equiv (1 - c(X - \nu)/k)^{1/c}$ $(X \geq \nu, c > -1)$, we have $G'/G = -1/(k - c(y - v))$ and $E[X|X \geq y] - y = (k - c(y - v))/(1 + c)$. So hG is unchanged for $c = 0$ (exponential distribution), increases if $1 - G$'s hazard is decreasing ($c < 0$), and decreases if it is increasing ($c > 0$).

unchanged distribution, and not fully adjust their reservation price, as in Lucas (1972). More frequent buyer visits may also be mistakenly attributed at first to chance. Moreover, sellers list and buyers do not, and so while buyers see the change in sellers' numbers (and, un-modelled here, their list prices), sellers do not see the change in the number of buyers or their willingness to pay, and so will lag in apprehending the changed environment, at least relative to buyers. Stein (1995) and Genesove and Mayer (2001) offer additional reasons for slow seller adjustment when prices are falling, but as we shall see, the vast majority of our data cover periods in which prices were rising.

Consider then a two period model in which sellers maintain their initial reservation price in the first period and then fully adjust to the new stationary state in the second. Assuming as before that buyers flow in and out of the market so as to keep V^Bconstant, the enhanced demand must be offset by a first period increase in θ. In Figure 2, θ increases to θ', the value that ensures that equation (3) holds at the new match quality distribution, with $\bar{V}^B$ and V^S (and so y as well) held constant. Consequently, q increases and h decreases. The acceptance rate increases by an amount much greater than in the steady state (for which the increase is due solely to y not quite fully adjusting to the higher offer distribution when $r > 0$). So *NHVIS* and *STOM* fall substantially. With h and $G(y - v)$ moving in opposite directions, the effect on *BTOM* is ambiguous.

To recall, we assume full adjustment to the new steady-state equilibrium (E^*) in the second period. Thus sellers increase their reservation price, so that y now increases, while θ falls to a level above its initial one. Relative to the first period, q and $G(y - v)$ fall, although to levels above their initial values. h increases, although to below its initial value. There is thus substantial overshooting, with long and short run effects in the same direction, but with the latter much greater. Both *NHVIS* and *STOM* fall substantially at first, then only partially recover, with full recovery for $r = 0$. *BTOM* 's behaviour is ambiguous.

We are also interested in the expected price, which equals $V^S + \beta(X - y) = \beta E[X|X \geq y] + (1-\beta)y - \bar{V}^B$. In the short run, y is constant, so price increases by $\beta \partial E[X|X \geq y]/\partial v$.[9] In the long run, y increases as well, so price increases by a further $\{\beta\, \partial E[X|X \geq y]/\partial y + (1-\beta)\}\, dy/dv$. The total long run price increase is positive.

2.3 Static Theoretical extensions.

This section relaxes several of the model's assumptions in order to assess its robustness.

The first is the infinite elasticity of the buyer inflow. Assume instead that the inflow is a linear function of the value of buyer search, $a_B + d_B V^B$, and that the inflow of new sellers is $a_S + d_S V^S$. The latter may represent new home construction and entry of existing homes, whose owners go elsewhere. Ideally, the existing homes inflow would be proportional to the stock of homes not offered for sale; our specification is an adequate approximation when the share of homes on the market is sufficiently small.

In the stationary state, these flows equal each other, which makes both V^S and V^B linear functions of y,[10] allowing us to write the buyer and seller asset equations as

(5) $$r\frac{d_B}{d_B+d_S}y = -\left(c^S - \frac{a_S-a_B}{d_B+d_S}r\right) + q(\theta)\beta G(y-v)(E[X|X \geq y] - y)$$

(6) $$r\frac{d_S}{d_B+d_S}y = -(c^B + \frac{a_S-a_B}{d_B+d_S}r) + h(\theta)(1-\beta)G(y-v)(E[X|X \geq y] - y)$$

From Mortensen (1986), we have that at the equilibrium θ, $\mathcal{S}$ and $\mathcal{B}$ shift up respectively by

(7) $$dy/dv = [q(\theta)\beta G(y-v)]/[r d_B (d_B + d_S)^{-1} + q(\theta)\beta G(y-v)]$$

(8) $$dy/dv = [h(\theta)(1-\beta)G(y-v)]/[r d_S (d_B + d_S)^{-1} + h(\theta)(1-\beta)G(y-v)] \;,$$

so that $\mathcal{B}$ shifts up more *iff* $d_S/d_B < (1-\beta)h(\theta)/\beta q(\theta) = (1-\beta)/[\beta\theta]$. This generalizes our baseline model, for which $d_B = \infty$. It seems reasonable to assume that even if not infinite, buyer inflow

[9] Price can fall in the short run. This will occur, e.g., for a GPD with $c < 0$. In the exponential case ($c = 0$) price stays constant in the short run: $E[X|X \geq y] = k + y$, so $\partial E[X|X \geq y]/\partial v = 0$.
[10] $V^S = (d_B + d_S)^{-1}[-(a_S - a_B) + d_B y]$, while $V^B = (d_B + d_S)^{-1}[(a_S - a_B) + d_S y]$.

is more sensitive to the search value than is seller inflow, both because buyers have a number of markets to choose among and because building new units takes time. Unless sellers have very little bargaining power, or the buyer-seller ratio is very large, the baseline model's conclusions still hold.

Note that we have not modelled owners who offer their homes for sale and search for a new home to buy in the same market. Since the inflow of such 'dual' agents has no effect on the net inflow of buyers less sellers, ignoring them is inconsequential to the steady state analysis and does not affect the linear relationships between y and V^S and V^B, so long as their subsequent actions as buyers or sellers are independent of each other. This approach is in line with papers such as Williams (1995) and Krainer (2001), although admittedly not with others, notably Wheaton (1991).

Note also that variation in v generates a positive price-volume correlation. The inflow (and so the steady state outflow, i.e., the number of transactions), is increasing in y, which is increased by v; while price is also increased by v. Such a correlation is widely thought to hold,[11] and has generated a number of possible explanations (such as Stein, 1995 and Genesove and Mayer, 2001). This analysis shows that a simple matching model will generate the correlation as well.

We also relax a number of other assumptions. Appendix A allows buyers to anticipate becoming sellers at some point in the future, by having owners have a constant hazard of becoming mismatched with the home. In Appendix B sellers act as monopolists, offering a take-it-or-leave-it price to the buyer regardless of the buyer's quality match to the home (which of course leads to an inefficient number of transactions). Neither modification makes any qualitative difference to the theoretical results. Nor does allowing bargaining power to vary with the buyer-seller ratio by having β increase in θ.[12]

[11] Stein (1995) and Berkovec and Goodman (1996) document a positive time series correlation for the US, Andrew and Meen (2003) and Ortalo-Magné and Rady (2004) for Britain;, and, using panel data, Leung, Lau, and Leong (2002) for Hong Kong. However, Follain and Velz (1995) find zero correlation in a fixed effects panel study of 22 US MSAs, while Hort (1999) finds a negative correlation for 19 Swedish regions.

[12] All that matters is that $\beta(\theta)q(\theta)$ is increasing in θ, and $\left(1-\beta(\theta)\right)h(\theta)$ decreasing. Both sequential bargaining under the threat of the arrival of an alternative seller or buyer (Rubinstein and Wolinsky, 1985) and auction models in which the arrival rate of bidders is increasing in θ (Julien, Kennes and King (2000)) have this property.

Appendix C models search effort. Following Pissarides (2000, chapter 5), it assumes that an agent can increase its contact hazard by expending more effort; agents' effort on the other side of the market increases this hazard, while that of other agents on its own side decreases it. Assuming an efficiency unit specification, as in Pissarides (2000), but adding stochastic matching, it shows that the comparative static results for y and θ still hold so long as the marginal cost of effort rises sufficiently fast, seller effort is sufficiently dissipative, or the buyer hazard is sufficiently decreasing in θ . The importance of induced effort relative to changes in the buyer-seller ratio and the acceptance rate is discussed in Section 5.6 in the context of the empirical results.

3. From the Model to the Data

Our empirical analysis focuses on three directly observed measures of market liquidity: *BTOM*, *STOM* and *NHVIS*. The model's stationarity implies that *BTOM* and *STOM* are distributed exponentially, and *NHVIS* geometrically. Thus, the median log *STOM* is $-\ln{(qG)}$ and the median log *BTOM* is $-\ln{(hG)}$, ignoring a common constant; we approximate the median log *NHVIS* by (a constant plus) $-lnG$.[13] We use the median because of top coding.

Linear combinations of these three variables yield the expected log contact hazards (which we label *lnq* and *lnh*) and the log buyer to seller ratio (*lnθ*). The above implies that $lnq = lnNHVIS - lnSTOM$. Similarly, $lnh = lnNHVIS - lnBTOM$. Lastly, and recalling that $q = h\theta$, $ln\theta = lnBTOM - lnSTOM$. Intuitively, if buyers spend half the time on the market that sellers do, they must be twice as likely to transact as sellers. However, as buyers and sellers leave the market in pairs, this implies that there are twice as many sellers as buyers. There is thus no need of measures of the stock of sellers and buyers. These are typically used in empirical labour matching models; however as buyers are not listed in North American housing markets, their stock is impossible to construct for empirical work.

[13] To check this, we simulated the median $lnNHVIS$ on the grid $G\epsilon\{.01, .02, \dots, .4\}$. The least squares regression of the simulated value on lnG is $median(lnNHVIS) = -.36 - .988 \times lnG$, with an R^2 of .993.

Although in principle, we would like to use consumption and production amenities as indicators of demand, the difficulty in measuring their yearly variation leads us to rely on income and population instead. We justify income as a proxy for demand in two ways. First, higher income people have a higher willingness to pay for consumption amenities, so that improvements in these amenities will be accompanied by higher income people moving in. Second, complementarity between individual skill and productive opportunities implies that improving the latter will attract higher wage people (Moretti, 2004). Population should also be correlated with consumption amenities: in an open city model, improving an amenity draws in more people until the increased commuter cost at the new urban edge just offsets the greater amenity value. This shifts up the bid-offer curve throughout the city. Note that sellers' willingness to pay also increase under these scenarios, but by less that of incoming buyers. For simplicity, our model considers the effect of shocks to buyers' willingness to pay only.

In theory, other types of shocks can generate the opposite correlations between housing demand and income and population proxies. Population increases can follow from zoning laws relaxations, although to the extent that they endogenously respond to demand pressure, as Wallace (1988) and McMillen and McDonald (1991) strongly suggest, no bias is introduced. With homogenous labour, pure consumption amenities should increase home prices but decrease wages (Roback, 1982) and so decrease the correlation between house prices and income. But the empirical evidence is consistent with our claims: in Gyourko and Tracy (1991), as in Gabriel et al. (2003), any significant effects of amenities on wages and prices are almost always of the same sign, Rauch (1993) shows that cities with higher average education exhibit both higher housing prices and higher labour income, while Capozza, Hendershott and Mack (2004) show that prices depend positively on median real income and population in both pooled and fixed effects regression. Gallin (2006) and Mikhed and Zemcik (2009) cast doubt on the consistency of Capozza *et al*'s (2004) estimates, given the inability to reject the null of no co-integrating relationship. However, these tests have little power, as Gallin emphasizes. Indeed, Glaeser and Gyourko (2005), who use long differences of thirty years, show the relationships quite robustly: population growth and price growth are positively correlated; they react similarly to the technological shock of weather adaptation; and growing cities have an increasing share of college graduates. These results further justify our use of income and population as demand proxies.

Data constraints force us to identify the market geographically with the MSA. In reality, however, an MSA is comprised of heterogeneous spatial and characteristic-based sub-markets which differ in their liquidity. We assume that heterogeneity can be captured by the Vaupel et al. (1979) frailty model with either a non-negative stable distribution (Hougaard, 1986) or a compound Poisson distribution (Aalen, 1988); the frailty distribution may differ by MSA but must evolve over time independently of the demand proxies. That ensures that the demand proxies' coefficient estimates will be unaffected by within-MSA heterogeneity. Under our assumptions, cross-MSA differences in the degree of within market heterogeneity are absorbed by the MSA effects. Because we will be using the median of the micro level response, and not the mean, we must however assume that the liquidity of each sub-market is equally responsive to the demand proxies.[14]

4. Data

We constructed a panel dataset from three sources. *STOM*, *BTOM* and *NHVIS* come from micro data of twelve separate NAR surveys of homebuyers conducted biannually between 1987 to 2003, and annually between 2003 and 2008. We lack only the 1997 and 1999 surveys. We aggregated the 53,505 survey responses up to the MSA level, by year. The combined sample covers 334 unique MSAs and primary metropolitan statistical areas, which we refer to collectively as MSAs. We include only MSAs that appear more than one year. The resulting panel is heavily unbalanced: the mean number of years an

[14] Thus we assume that for outcome measure y, and unit (i.e., buyer or seller) j in MSA m in year t, $lny_{jmt} = a_j + X_{mt}b$, where X is the demand proxies, and the distribution of exp (a_j) is one of the aforementioned. This implies that, the median lny equals $X_{mt}b + \mathrm{H}_{mt}$, where H_{mt} is a function solely of the frailty distribution. For b to be consistently estimated, we require H_{mt} to be uncorrelated with X_{mt}, conditional on the MSA and year effects For example, in the case of a mean one gamma distribution with variance σ_{mt}^2, $\mathrm{H}_{mt} = \ln\left((2^{\sigma_{mt}^2} - 1)/\sigma_{mt}^2\right)$. Were the model $lny_{jmt} = a_m + X_{mt}b_j$ instead, there would be a bias: assuming a gamma distribution for $\exp(X_{mt}(b_j - b))$, $\sigma_{mt}^2 \approx X_{mt}^2 Var((b_j)$, imparting an approximate bias of $2\bar{X}\zeta Var((b_j)$, $\zeta \in (.386, .692)$, to the estimate of b, where $\bar{X}$ is the mean X in the sample. This bias would not arise were we able to work with mean log outcome.

MSA is observed with *STOM* is 6.3 (the standard deviation is 3.9; the maximum, 17). The unbalanced panel, together with our lacking two surveys and the biannual frequency of the early surveys, forces our dynamic analysis to be rudimentary. Yet the results are so stark that it is difficult to believe they would not be present in a more sophisticated dynamic specification, were the data permitting.

The questionnaires were sent to recent home buyers, who reported on their home buying experience, and when they had sold a previous home, on their selling experience as well. We use selling information only for those who sold two or three years before the survey date, as only then was the sale year reported. As prior to 2007, the city of previous residence was not asked, we only use sales for which the respondent reported moved fifty miles or less. Looking at sellers who buy within two or three years in the same MSA market should not be too restrictive, since the post-2007 data show that such sellers remain on the market a mere 3 percent longer than those who buy elsewhere within two years, while Sinai (1997) shows that three quarters of sellers buy again within two years of selling. To avoid the post-2007 foreclosures, we use only pre-2008 transactions; consequently, our use of the 2008 survey is restricted to the pre-2008 sales that are reported there.

A more serious concern is the extremely low response rates of the NAR surveys, which never exceed 19 percent and fall as low as 6 percent in one year. Whether or not this leads to a serious sample selection bias depends on the pattern of survey response. One possibility is that the decision to respond to the survey depends on how the subject views his position in the distribution of outcomes in his market. For example, those on the market longer than that market's norm may respond in order to vent their annoyance; alternatively, those who transact quickly may respond in order to celebrate or gloat. Appendix D shows that there will be no bias other than on the constant if the subject correctly perceives the outcome distribution. Even if the perceived distribution lags that of the true one, as the 'dynamic extension' assumes, the bias is shown to be quantitatively unimportant. Another possibility is that the propensity to respond is correlated with the subject's economic and demographic characteristics. For example, high income people have higher opportunity costs of time, making them less likely to respond to the surveys and less willing to search longer. Conditional on fixed effects, this introduces a sample selection bias, which might contaminate the estimated effect of market level income shocks. To address this bias, we include respondents' mean economic and demographic characteristics, as a linear approximation of a Heckman selection term. As shown in Section 5.2, doing so will have little effect on the coefficients of the market level demand proxies.

With the individual NAR level data, we construct the following time-varying MSA level medians: *BTOM, STOM,* and *NHVIS*. We use the median and not the average, since the variables are top coded in certain years. As responses are requested in open form in some survey years and in an unchanging set of brackets in others, we assign the midpoint of the corresponding bracket as the value of the variable for the open form years. Our results are nearly identical when we alternatively use the reported value for the open form years in which brackets are not used and assign the midpoint value otherwise.

MSA-level population and income are obtained from the Bureau of Economic Analysis. Yearly repeat sales housing price indices are derived from the Office of Federal Housing Enterprise Oversight (OFHEO), which tracks average single-family house price changes in repeat sales or refinancing. On average, about three thousand repeat transactions underlie a given year and MSA's index value.

Table 1 provides summary statistics. An observation is an MSA X year combination. All statistics are weighted by the number of underlying micro level responses. We use three different samples: the Seller sample, for which *STOM* is available; the Buyer sample, for which *BTOM* and *NHVIS* are available; and the Joint sample (the intersection of the previous two). Note that the year of survey does not necessarily correspond to the year of transaction. The Seller sample includes observations for 1986-1996 and 2000-2007, the Buyer and Joint samples for years 1986-1993, 1995-1996, 2000-2001, and 2003-2007. The Seller sample is substantially smaller than the Buyer sample because the survey is sent to a sample of purchasers; information on selling comes only from those purchasers who also sold. However the former is not a subset of the latter, as one may sell in a different year than one purchases.

The weighted average (across MSA X year observations) of the median time on the market for sellers in the Seller sample is 7.3 weeks. That for buyers in the Buyer sample is 8.2 weeks. Multiplying these numbers by 1.44 (the ratio of the mean to median for the exponential distribution) yields an estimate of the mean under the model's stationarity assumption. The weighted average median *NHVIS* is 9.9. These numbers vary little across the samples. Table A.1 shows that figures reported for other survey-based samples (with much better response rates) are very similar to these. Although our reported statistics for the mean *BTOM* and *STOM* are quite close to one another, the mean $ln\theta$ is 0.25. All three liquidity measures vary substantially over time and across space, a necessary condition for their playing a role in how the market adjusts to changes in demand.

The means for log average income and log population are reported. Average population is two percent greater in the Seller than Buyer sample, and three percent greater again in the Joint sample. Average income is two percent smaller in the Buyer than in the Seller and Joint samples. Finally, average yearly price appreciation is six percent. In only 7 to 9 percent of the observations is price falling, suggesting that the equity and loss aversion effects demonstrated in Genesove and Mayer (1997, 2001) are not relevant for the vast majority of markets in our sample.

5. Reduced-Form Analysis of Demand Effects on Matching Behavior

5.1 Income and Population as Demand Proxies

Table 2 examines how buyer and seller behavior changes in response to variations in the demand proxies. Standard errors are robust and clustered at the MSA level to account for any kind of auto correlation in the errors[15] as well as cross-MSA heteroskedacticity. This appreciably increases the standard errors by at least a quarter and sometimes much more, but except in one case, which we will point out, has no impact on the significance status at conventional levels. All specifications include

[15] This is a concern only for the *NHVIS* regressions, which occasionally fail Wooldridge's (2002, p. 275) test for no serial correlation, with the predicted value under zero correlation $M^{-1}\sum_i 1/(T_i+1)$, where M is the number of MSAs and T_i the number of observation for MSA i. The *STOM* and *BTOM* regressions always pass the test.

dummies for both the year of the survey and the year of transaction. We also include MSA fixed effects to control, e.g., for differences in the efficacy of the MLS system. All variables are in logs. To adjust for heteroskedacticity, we weight each observation (MSA/year) by the underlying number of transactions.

Columns (1), (3), and (5) regress *STOM*, *BTOM*, and *NHVIS* on income and population. As predicted, greater income is associated with a shorter *STOM* and fewer homes visited, although the latter is insignificant. The income coefficient of -1.43 in column (1), for example, implies that an extra ten percent in average income is associated with about a thirteen percent shorter *STOM*, while Column (5)'s coefficient of -0.44 predicts three percent fewer home visits. At *mean* values, this is slightly more than nine days fewer on the market, and one third of a home less. *BTOM*, for which the model provides no unambiguous predicted response to demand, has a negative but very small and insignificant income coefficient. Population has significant negative effects on all the variables. A ten percent greater population is associated with an eleven percent shorter *STOM* and a four and a half percent shorter *BTOM*, and five percent fewer *NHVIS*. These results are in line with the basic model.

Additional predictions of the model can be assessed by comparing across the columns. Theoretically, log time on the market should equal the log *NHVIS* minus the log hazard rate. Thus, the sensitivity to the demand proxies should be greater for *STOM* than for *NHVIS* (since the seller hazard should increase with demand), but greater for *NHVIS* than for *BTOM* (since the buyer hazard should decrease with demand). Also, *BTOM* should decrease less than *STOM* as their log difference is the log buyer-seller ratio. All those predictions hold, both for income and population.

Those comparisons, however, may be contaminated by the non-coincidence of the Buyer and Seller samples. Columns (1), (3) and (5) of Table 3 thus regress lnq, lnh, and $ln\theta$ on the demand proxies, using the Joint sample. All the coefficient estimates for lnq and $ln\theta$ are positive, although only income is significant. As the sample is smaller, the estimates are noisier than before. The effects for lnh are negative, although never significant. The sign patterns are all consistent with the model.

The remaining columns of Tables 2 and 3 distinguish between short run and long run effects by adding income growth and population growth. This is a barebones dynamic specification, but as noted above, the realities of our panel force that upon it. Furthermore, the data's frequency forces us to identify the short run with a year, which is longer than we would prefer, given the rarity of either a buyer or seller being on the market that long, and the likely speed of the information flow. Nonetheless, too long a period for the short run should bias the estimated short run effect downwards in magnitude. Also note that the clustered standard errors account for any pattern of autocorrelation in the errors.

Given the MSA and year fixed effects, the short and long run effects are identified by variations in the level and growth rate around an MSA's average (relative to the calendar year). To see that, write the regression as $y_{it} = x_{it}\beta + \Delta x_{it}\gamma + u_t + v_i + e_{it}$, where i indexes the MSA, t the transaction year, u_t (v_i) the time (MSA) effect and $\Delta x_{it} \equiv x_{it} - x_{it-1}$. Then our regression amounts to

(9) $$y_{it} - \bar{y_i} - \bar{y_t} + \bar{y} = (x_{it} - \bar{x_i} - \bar{x_t} + \bar{x})b + (\Delta x_{it} - \overline{\Delta x_i} - \overline{\Delta x_t} + \overline{\Delta x})g + \{e_{it} - \bar{e_i} - \bar{e_t} + \bar{e}\}$$

(Balestra, 1992). Consider, e.g., x constant over periods 1 through s, and constant at a different level from s+1 until T. Intuitively, the difference between mean *STOM* over 1 through s and the mean from s+2 through T (all zero growth periods) identifies the level effect, while the difference between *STOM* in s+1 (a growth period) and the mean in all other periods identifies the growth effect.

Column (2) of Table 2 exhibits the striking result that both income growth and population growth strongly decrease *STOM*. The semi-elasticity of *STOM* with respect to income growth is about seven, while that with respect to population growth is about 14.5. In contrast, the effects of the income and population levels are much smaller, with the income coefficient insignificant.

These responses are consistent with our dynamic extension of the model, in which sellers and buyers react to changes in the economic environment at different speeds: in the short run (given by the sum of the coefficients on the level and change, and shown in the table's second panel) there is a large fall in *STOM* upon a demand increase, as buyers offer more, while sellers do not yet adjust to the new demand reality. In the long run (seen in the level coefficient alone), sellers increase their reservation value, and so

STOM falls much less dramatically. The estimates indicate that a one percent increase in income decreases *STOM* by three and a half days in the short run, but by only an (insignificant) quarter of a day in the long run; an equivalent increase in population is associated with a full week's decline in *STOM* in the short run, but only a (significant) half day in the long run.

The dynamic extension also implies that *NHVIS* decreases substantially in the short run with demand, and indeed Column (6) shows significant short run population and income elasticities of -4.5 and -1.2. Here, too, the long run effects are much smaller, although under the robust standard errors one cannot reject the null of equal short and long run effects.[16]

Turning to *BTOM*, we see in column (4) that although the four regressors all have negative signs, only the population level coefficient is significant. To recall, the model makes no unambiguous prediction here, as the buyer contact hazard is predicted to fall while the acceptance rate is expected to increase. Since, in contrast, the seller contact hazard rate is predicted to increase, the model does predict that time on the market for buyers can not fall more than that of sellers, and indeed the coefficients are all much smaller in magnitude than those for *STOM*, as expected.

Columns (2), (4) and (6) of Table 3 add income and population growth to the lnq,lnh , and $ln\theta$ regressions, run on the Joint sample. Not surprisingly given the previous estimates, and in line with the dynamic extension to the model, short run effects are much larger than long run effects, with the latter always insignificant. The short run effects are significant in all cases except for that of income on lnh, and have the predicted signs. The effects for the buyer contact hazard are much smaller in magnitude than those for the seller. Section 5.4 offers an explanation based on real estate listing institutions.

The assumption that income and population affect the dependent variables only through their correlation with a common effect (the model's v) is easily tested. Within a regression, this is the test of the null hypothesis that the ratio of the coefficient on population to that on income equals the corresponding ratio for the change in the variables, i.e., $\frac{b_{pop}}{b_{inc}} = \frac{g_{pop}}{g_{inc}}$, where b and g are the coefficient

[16] One can reject under the usual standard errors. Only here does significance hinge on robust standard errors.

vectors in equation (9). This test, "F-stat" in Tables 2 and 3, is never rejected. A more comprehensive test is that for the null that the ratio is the same not only within, but also across, all three regressions. Distributed asymptotically as $\chi^2(5)$, it equals 0.53 (p-value of 0.99) for Columns (2), (4) and (6) of Table 2 and 2.99 (p-value of 0.70) for the same columns of Table 3. Clearly the null cannot be rejected.

Our model also makes predictions about prices. Given the many studies on the relationship between housing prices and income and population (see Section 3) none of which required our unique panel, we conduct only a cursory investigation. Column (1) of Table 4 shows the regression of the log OFHEO price index on the demand proxies for the Buyer sample, with MSA and year fixed effects and robust standard errors, clustered at the MSA level. With no breaks in the data series for the dependent variable, we are able to include its lagged value. The estimates imply a long-run elasticity of price with respect to average income of 0.063/(1-0.938)=1.02 (s.e. of 0.41) and a short-run elasticity of 0.063 +0.437 =0.50 (s.e. of 0.07). The long and short run elasticities with respect to population are 2.48 (s.e. of 0.55) and 1.55 (s.e. of 0.38). The long run effect exceeding the short run is consistent with the model's dynamic extension. Consistent with the hypothesis that income and population proxy a single demand shock, one cannot reject that the ratio of the long to short run effect is equal for both proxies.

5.2 Additional Controls

Not all population changes are equally indicative of demand changes. Certain age groups are more mobile than others, and so should move more readily with changes in consumption and productive amenities. Table 5 thus adds the share of population between ages 25 to 40 and its first difference to the set of regressors. We expect their coefficients to have the same sign as those on population and its change. That turns out to be the case when the population share coefficients are significant – in the static *BTOM* regression and on the change in the population share in the dynamic *NHVIS* regression. In the *STOM* regressions, the coefficient is of the wrong sign, but always insignificant. Including population share has little effect on the income and population coefficient estimates.

As noted earlier, if response patterns are correlated with demographic characteristics, MSA income and population coefficient estimates might reflect sample selection. To address that bias, we include mean respondents' economic and demographic characteristics: log respondent income, respondent age, number of children in a household, and couple status, all measured at the MSA/year level. These variables have all been viewed as individual level determinants of duration of home search before, through their effect on search costs (e.g., Anglin, 1997; Baryla and Zumpano, 1994), and as good predictors of the propensity to respond. Their presence in the regression should be seen as a linearized Heckman selection term that is estimated simultaneously with the equation of interest. This is a feasible strategy for estimating the coefficients on the MSA level demand proxies, as there is no reason to think that sample selection should be a function of these demand proxies directly. Our lack of data on non-respondents, which precludes us from estimating a sample selection equation, is thus no impediment. Since there may be other respondent characteristics that are residually correlated with the MSA demand proxies, this approach will not eliminate all the selection bias, but comparing its estimates to Table 2 should give a good indication of the seriousness of sample selection bias here.

Table 6 repeats the estimation in Table 2 on the sample for which respondent demographics are available. Controlling for the demographics has little impact on the estimates of demand effects. In the *STOM* regressions of Columns (1)-(4) the population and income level terms have the same negative signs as before, whether we control for average respondent characteristics (column (2)) or not (column (1)). They are no longer significant, but given the much reduced (by nearly 40 percent) sample size, this is not surprising. When we add the growth terms, the pattern is similar to that of Table 2: *STOM* decreases substantially with demand in the short run but almost fully recovers in the long run (column (3)). Adding the controls changes neither the estimates nor the demand coefficients' significance much. No demographic control is statistically significant, although they are jointly so at the 10% level.

Columns (5)-(8) show *BTOM* regressions. Here, too, the controls, none of which are significant, barely affect the coefficients of interest, whose pattern is as before, with long run effects greater than short run, and population effects greater than income. The *NHVIS* regressions are also similar to Table 2, except that the income growth coefficient is now smaller in magnitude than the level coefficient, though both remain insignificant. The controls are highly jointly significant, with income and couple status positively predicting *NHVIS*.

Other *BTOM* and *NHVIS* regressions (not shown) include the percentages of first-time home buyer (reflecting home buying experience, Baryla and Zumpano, 1994), and new-home buyer (for differential effort required in home inspections, if new homes involve less hidden information). The only significant estimate for these two variables is that of the percentage of new homes in positively predicting *NHVIS*, but, again, including them is inconsequential for the variables of interest.

Of course, even conditional on these various controls (and the fixed effects), the mix of respondents may vary across observations. However, these controls are likely to capture the major dimension of heterogeneity among respondents. As including them has little effect, we conclude that it is unlikely that unobservable types introduce any significant bias into the main empirical analysis.

5.3 Estimating the Contact Hazard Elasticities

Our estimates of the effect of the demand proxies on the buyer-seller ratio and the contact hazards provide estimates of the matching function, or more precisely, the elasticity of the contact hazard with respect to θ. From Table 3, we see that, e.g., a one percent increase in income growth increases $ln\theta$ by 7.64 percent, while decreasing lnh by 1.19 percent, implying that the elasticity of the buyer hazard with respect to the buyer-seller ratio is -0.16. By construction of our dependent variables, the elasticity of the seller hazard with respect to θ equals one plus the buyer hazard elasticity, or 0.84. Using population growth yields a buyer hazard elasticity of -0.20.

These calculations are essentially instrumental variables estimates. Two-stage least squares regression of lnh on $ln\theta$, with the usual MSA and year fixed effects, and income growth and population growth as instruments, yields an estimate of -0.17, with a robust, MSA clustered standard error of 0.10. As the theoretical range of the elasticity is $[-1,0]$, this should be seen as a small estimate. Of course, these results are valid only to the extent that income and population growth are uncorrelated with within city changes in the matching technology – of which more later.

5.4 Supporting Evidence

(1) Lagged seller response: If sellers' response lags, so should the list prices they set, and the price premium – the sales to list price log ratio– should increase with demand in the short run. Table 4's Column (2) shows this occurring for both demand proxies. The long run effect is much smaller and insignificant, consistent with sellers' eventually adjusting to the new demand reality.[17]

(2) Small buyer contact hazard: The small buyer contact hazard is consistent with the matching institution of most U.S. real estate markets. This is the multiple listing service, in which sellers list their property but buyers typically do not advertise themselves. Sellers list, not buyers, because it is easier to describe a home and its list price than to describe a buyer's preferences over attributes and price. That buyers seek out sellers but not vice versa suggests that q will be more sensitive than h to θ. Consider a limiting case: a buyer's opportunity to search arises independently of other buyers' at a constant rate ρ, and if he (or his agent) chooses homes to visit randomly, then the buyer contact hazard is $h = \rho$,[18] and $q = \rho\theta$. Reality is not so extreme: buyers' search opportunities are clustered over the week and offers last over some time interval, during which another buyer can show up. So a buyer's contact hazard will be decreasing in the number of buyers, and increasing in the number of sellers, which increases the chance that other buyers will end up elsewhere. Nevertheless, the asymmetry in the parties' roles in the listing institution makes it reasonable to presume that h will be less responsive than q to θ, and that is indeed what the estimated elasticity implies.

---

[17] Our model has no role for list prices, but in Julien, Kennes and King (2000),which assumes that the buyer pays the list price (unless multiple buyers show up and an auction ensues), shows that in large markets, the list price is set equal to the seller's reservation price. One might object that as list prices are set before sales prices, the premium will be increasing in price growth, even in the absence of lagged seller adjustment, if list prices are set according to contemporaneous sales prices. But Table 1 shows the average median *STOM* is 7.3 weeks, implying a mean of 10.5 weeks, i.e., 1/5 of a year. Clearly, the growth in prices from Column (1) is too small to explain Column (2)'s result.

[18] See Mortensen and Pissaridies (1996). Wheaton (1991) uses this specification.

(3) A negligible interest rate: Our results show small long run demand effects on market liquidity. This is consistent with a small interest rate, for which under our baseline model neither the buyer-seller ratio nor the acceptance rate, and thus neither $STOM$ nor $BTOM$, changes much with demand. We see that more formally in equation (7), which shows that, when $r\frac{d_S}{d_B+d_S}$ is much smaller than $q\beta G$, increased demand shifts up S almost one by one, so that neither y nor θ changes much. Does this condition hold empirically? Our discussion of Table 1 shows that the mean *STOM* is 7.32 X ln2 X 7 days, which is equivalent to 1/qG in the steady state. This implies a daily hazard rate of sale (qG) of 1.36 percent. In contrast, the 20% interest rate inferred by Genesove and Mayer (1997) and subsequently used by Levitt and Syverson (2005) implies a daily rate of only one-twentieth of a percent. Since $d_B > d_S > 0$, unless sellers receive only a very small share of the surplus, $r\frac{d_S}{d_B+d_S}$ is clearly much smaller than $q\beta G$.[19] Thus, small long run effects are precisely what we should expect.

5.5 Alternative Dynamic Supply Explanations

Can inelastic short run supply coupled with elastic long run supply explain our dynamic results? Intuitively, a demand shock will first increase the buyer-seller ratio, as new buyers flow in. In the long run, an increased price will induce developers to build new homes. This inflow of newly constructed homes will then lower the buyer-seller ratio. Thus even without recourse to a formal matching model, a slowly adjusting housing supply seems capable of explaining why *STOM* decreases substantially in the short run but largely recovers in the long run.

However, this alternative explanation fails to explain the remaining dependent variables. Most crucially, inelastic short run and elastic long run supply implies that price will overshoot following a demand shock, while Table 4's column (1) shows price rising more in the long than in the short run, consistent with our model of a lagged seller response.

[19] Sufficiently high seller bargaining power is an appropriate assumption, given Chen and Rosenthal 's (1996) theoretical result that setting a 'ceiling' list price accords seller full bargaining power in certain circumstances.

The alternative explanation is also at odds with NHVIS's behaviour. In this alternative model, prices come down in the long run due to the inflow of sellers willing to transact at a lower price that covers only land and construction costs, and not the additional idiosyncratic ties that owners have to their homes. These sellers are willing to accept lower offers than the existing home sellers – that is the meaning of the long run supply curve being more elastic. So in the transition from the short run to the long run, sellers' reservation prices fall, and with demand remaining at its new higher level, the probability of sale must increase. Thus *NHVIS* should increase more in the short than in the long run, according to the alternative. This is not what we see, however.

Nor can out-of-steady state transition dynamics explain our results. Pissarides' (2000) analysis has the system immediately jumping to the long run θ and y, which implies that all three measures of liquidity also adjust immediately, in contrast to the empirical results.

5.6 Variations in Technology and Search Costs

Our interpretation of the empirical results requires that the matching function not vary with demand. However, the sample period was one of dramatic advances in communication and digitalization.[20] Fixed effects will pick up national level changes in technology over time, or fixed differences across MSAs, but not non-uniform technological adoption across markets. New technology that increases the matching rate for a given buyer-seller ratio is more likely to be adopted in high income and, due to fixed costs of adoption, large markets. Thus technological advances might seem to explain why income and population are negatively correlated with *BTOM* and *STOM*.

Our empirical results are not thus easily explained, however. Since a contact hazard is simply the matching function divided by the number of agents of that type, technological adoptions that increase the matching rate will have the same direct effect on the buyer as on the seller contact hazard (holding the buyer seller ratio constant). Yet that is not at all the case. As Column (4) of Table 3 shows, *lnh* does not

[20] Photographic developing and processing improvements – the spread of one hour photo kiosks - put pictures in agents' hands quickly; camera digitalization decreased costs and further sped up processing. Computerization sped the dissemination of MLS updates, with brokers receiving weekly update diskettes, rather than paper inserts. The Internet further sped up MLS updating, and permitted buyers to visually assess properties from home.

increase with the demand proxies, as one would expect from a correlation with technological improvements; instead, it falls with income in the short run.

Also, technological changes that increase the matching rate will not produce the observed correlation between homes visited and the demand proxies. An increase in the hazard rates will increase each side's search reservation value for a given θ, thus increasing y and so decreasing the acceptance rate and increasing homes visited - the opposite of what we observe.

Technological advance has difficulty explaining the dynamic regression results as well. No obvious reason explains why income and population growth should be associated with more efficient matching technology. While real estate firms might *adopt* newer technologies in a high growth period, the *level* of technology, which is the relevant factor for the *level* of time on the market or homes visited, should not necessarily be greater in times of greater growth, and certainly would not explain such large effects. (Recall that the MSA fixed effects will control for differences in MSAs' average growth rates in the dynamic regressions.) So our large estimates of short run effects are likely to be robust to the presence of economically exogenous technological change.

What about demand induced technological change? Although a full analysis is beyond the scope of this paper, Appendix C's analysis of endogenous search effort can guide us. It shows seller effort to be increasing in y (and buyer effort constant). This provides an additional mechanism by which the seller hazard increases with demand, while mitigating the fall in the buyer hazard, or even causing it to rise on net. We are unable to empirically isolate these additional effects, but the negative coefficients in the buyer hazard regressions do show that any induced effort or technological change there may be is dominated by the primary mechanism of the changing buyer-seller ratio.

Our main argument for the robustness to technological change is empirical, however. Although we lack measures of adoption of the earlier technologies, we do have a good one for Internet use: the fraction of buyers who report finding their home via the Internet. This variable reflects the conflation of buyer access to the Internet and sellers' (or their agents') postings on websites. Figure 3 shows the evolution of the variable mean over time. Obviously, Internet use is not a possible response in the earlier surveys, but it does appear as far back as 1995, when commercial Internet use was in its infancy (as that year's mean shows). We thus set the variable to zero for all earlier years.

Table 7 shows regressions with and without Internet use. The Internet coefficient is always large, and is significant for *BTOM* and *NHVIS*. The estimates predict that if all buyers find homes via the Internet, *BTOM* will be twenty four percent[21] and *NHVIS* thirty percent greater than if none do. The effect on *STOM* is of a similar size as that on *BTOM*, but is insignificant. Crucially, including the Internet variable has no noticeable effect on the demand coefficients. As the Internet is surely the dominant technological change over the last decade of our sample period, and that our measure of its use is evidently good enough to show a strong correlation with *BTOM* and homes visited, it is difficult to imagine that other, unmeasured, technological changes are responsible for our earlier results.

Finally, search costs may be higher for both buyers and sellers in high income cities, if higher income people value their time more. At a given buyer-seller ratio, this would lead to lower reservation values and so a higher acceptance rate, and so impart a negative bias to the long run effect of income on *NHVIS*. However, that should not impart any direct bias to the hazard rate regressions, and there being no dynamic element in this explanation, to differences between the short and long run effects.

5.7 Reporting Errors in *BTOM*

The relative insensitivity of *BTOM*, although predicted by the model, is also consistent with it being less well defined than *STOM*. As a rule, sellers begin searching by contracting with an agent; whereas a buyer might visit homes intermittently and haphazardly. So *BTOM* might be reported more noisily than *STOM*. Since *BTOM* and *STOM* are dependent variables, only precision is an issue. In fact, the *BTOM* regressions' standard errors are typically only half those in the corresponding *STOM*

[21] This is very close to that of D'Urso (2002), who uses a single cross-section. The Internet use variable clearly suffers from classical measurement error, and so attenuation bias. The samples are a tad smaller than in Table 2, as the variable's construction requires buyer responses in the year and MSA. Using two other measures, the number of broadband operators and a Current Population Survey based measure, instead, leads to similar results, although much smaller samples and much noisier estimates, as these measures do not go as far back in time.

regressions; it is the size of the estimated coefficients that accounts for the insignificant results. The relative indefiniteness of searcher status for buyers might alternatively lead people to report focal values that do not vary with the actual experience; belying that explanation is the fact that *BTOM* does vary significantly with income level and Internet use, and (although not shown) across MSAs.

6. Conclusion

This paper is the first to empirically test a housing market matching model. Using unique data, it shows that both buyers' and sellers' search experiences are consistent with a straightforward random matching model amended by a lag in the seller response to demand shocks. In the short run, both *STOM* and *BTOM* decline in response to an increase in demand --the effect for sellers being quite large, while that for buyers much more moderate. These effects are much smaller in the long run.

To understand why market liquidity responds the way it does, we also examine how demand affects the underlying components of time on the market. In the short run, a positive demand shock increases the buyer-seller ratio dramatically. There is a corresponding large increase in the seller contact hazard, and a small decrease in the buyer contact hazard. Conditional on meeting, the probability that an offer is accepted also falls. This pattern explains the large decrease in STOM and moderate decrease in BTOM in the short run and is consistent with our assumption that sellers keep their reservation prices unchanged in the short run. In the long run, the buyer-seller ratio falls but to a level above its initial value, while the acceptance rate increases to below its initial value, consistent with sellers eventually adjusting their reservation prices to the higher demand. Thus the long run effects are all in the same direction as short run effects, but of a much smaller magnitude.

The timing and coverage of the NAR surveys prevent us from fully describing the dynamics of our dependent variables.[22] But the basic pattern should be clear: a large, short run response, and a muted,

[22] Running the regressions with the lagged (by one year or two, as available) dependent variable included, we have obtained similar results to those reported (although much noisier, as we have had to both the sample's first year and 2001, given the missing 1997 and 1999 surveys, as well as additional years giving the incomplete coverage).

if any, long run response for *STOM, NHVIS* and the buyer-seller ratio. This pattern is consistent with the seller listing institution used in US real estate markets, the relative size of the interest rate and the sales to list price ratio.

We also estimate the elasticity of the contact hazard rates with respect to the buyer-seller ratio. In line with the seller listing institution of U.S. real estate markets, we find a relatively small contact hazard elasticity of -0.16 for buyers (equivalently, a large elasticity of 0.84 for sellers). This implies that sellers must have substantial bargaining power for the market to be near efficient (Hosios, 1990). This, and our other quantitative results, should prove helpful in search-based calibration models of the housing market, a number of which have already appeared,[23] and more of which we expect to see in the future in light of the recent boom and bust in US markets. Comparing our estimates to those from markets with different matching institutions, like England which lacks multiple listing services, or from the most recent turbulent time period, should also prove fruitful.

[23] Examples include Ngai and Tenreyro (2009),Caplin and Leahy (2008) and Diaz and Jerez (2009).

Appendix A: Buyers Who Anticipate Becoming Sellers

Assuming buyers do not anticipate being sellers at some point in the future is not a critical element of the analysis. The return to being an owner when there is a constant hazard λ of becoming mismatched with the house is $rV(u) = u + \lambda(V^S - V(u))$, where u is the flow value of the owner-house match. Thus there is a transaction when $u \geq (r+\lambda)y - \lambda V^S \equiv z$. The expected surplus, conditional on it being positive, is $E[(r+\lambda)^{-1}(u+\lambda V^S)|u \geq z] - y = (r+\lambda)^{-1}\{E[u|u \geq z] - z\}$. The asset equations are as below, so that all our results follow as before:

$$0 = -(c^B + r\bar{V}^B)(r+\lambda) + h(\theta)(1-\beta)G(z-v)(E[u|u \geq z] - z)$$

$$rz = -r(c^S - (r+\lambda)\bar{V}^B) + q(\theta)\beta G(z-v)\frac{r}{r+\lambda}(E[u|u \geq z] - z).$$

Appendix B: Seller Monopoly Pricing

Assume the seller does not observe X, but knows G, and makes a take it or leave it offer P^M to the buyer. He accepts it if and only if $X - P^M \geq V_B$. Thus P^M maximizes $\mathrm{G}(P^M + V_B)P^M + \big(1 -$ $GPM+VBVS$, and so $PM=VS+j(PM+VB)$, where $j\equiv-G/G'$. Let y□$=PM+VB$.

The buyer's expected surplus, conditional on a transaction, is $[X|X \geq y^*] - (P^M + V_B) = E[X|X \geq y^*] - y^*$. The seller's is $P^M - V_S = y^* - y$. The asset equations are thus $ry = -(c^S - r\bar{V}_B) + q(\theta)G(y^* - v)(y^* - y)$ and $0 = -(c^B + r\bar{V}_B) + h(\theta)G(y^* - v)\{E[X|X \geq y^*] - y^*\}$. Label these as curves $\mathcal{S}^*$ and $\mathcal{B}^*$ in (θ, y^*) space. Clearly $\mathcal{B}^*$ slopes down. Totally differentiating the seller equation yields $dy/d\theta = q'(\theta)G(y^* - v)(y^* - y)/[r + qG] > 0$. But we need to know $dy^*/d\theta$. For the GPD, $j(y^*) = k - c(y^* - v)$, so that $y^* = (k + cv + y)/(1+c)$. For $c > -1$ (only then does the mean exist), $\mathcal{S}^*$ slopes up in (θ, y^*) space.

$\mathcal{B}^*$ shifts up one for one with v, as does $\mathcal{S}^*$ if $r = 0$. More generally, totally differentiating the seller equation yields $dy/dv = qG/(r+qG)$. Under the GPD, $d\,y^*/dv = (1+c)^{-1}\{c + dy/dv\}$. Thus for $c > -qG/(r+qG)$, $\mathcal{S}^*$ shifts up less than one for one with v. For $-1 < c < -qG/(r+qG)$, a rather narrow range, given the sale hazard (see our discussion in the text),$\mathcal{S}^*$ actually shifts down; this also leads to an increased θ, and changes none of our qualitative predictions.

Appendix C: Induced Search Effort

Let i_S (i_B) indicate the individual seller's (buyer's) effort, and $\bar{\imath}_S$ ($\bar{\imath}_B$) that of all other sellers (buyers). Search cost is an increasing convex function of own effort: $c_S = c_S(i_S)$ and $c_B = c_B(i_B)$. The contact hazards are now written as $q(\theta, i_S, \bar{\imath}_S, \bar{\imath}_B)$ and $h(\theta, i_B, \bar{\imath}_B, \bar{\imath}_S)$; subscripts will indicate partial derviatives. Let $A \equiv G(y - v)(E[X|X \geq y] - y)$. Note that $dA/dy = -G\, , dA/dv = G$.

Sellers choose effort so that $0 = -c_S'(i_S) + q_2\beta A$. Consider the symmetric equilibrium in which $i_S = \bar{\imath}_S$ and $i_B = \bar{\imath}_B$. Assuming an efficiency unit specification, $q_2 = q/\bar{\imath}_S$ (Pissarides, 2002, p. 128, eq. 5.12). The seller asset equation (with $\bar{V}^B = 0$ for simplicity) implies $ry = -c_S(\bar{\imath}_S) + q\beta A = -c_S(\bar{\imath}_S) + \bar{\imath}_S c_S{}'(\bar{\imath}_S)$. This implies $\bar{\imath}_S{}'(y) = r/ic_S{}'' > 0$. Similarly, $0 = \bar{\imath}_B c_B{}'(\bar{\imath}_B) - c_B(\bar{\imath}_B)$, so that $\bar{\imath}_B$ is constant.

Differentiating the seller equation yields $\mathcal{S}$'s slope: $q_1\beta A/\{r + q\beta G - q_3\beta A\bar{\imath}_S'(y)\} > 0$, since $q_3 < 0$ (more effort by other sellers reduces the chance that a buyer will contact the given seller). An increase in v shifts $\mathcal{S}$ as before: $0 < \ dy/dv = q\beta G/\{r + q\beta G - q_3\beta A\bar{\imath}_S'(y)\} < 1$.

$\mathcal{B}$'s slope is $h_1 A/\{hG - h_4 A\bar{\imath}_S'(y)\}$, which is negative as long as $h_4 A\bar{\imath}_S'(y) < hG$. Since $q = h\theta$ and $h_4 = \theta^{-1}[q_2 + q_3]$, this condition can be written as $\big(1 + (q_3/q_2)\big)(r/\beta)/qG < \epsilon \equiv ic_S''/c_S'$ (the elasticity of the marginal cost of effort). Our baseline model has ϵ infinite. The condition also holds if $q_3 = -q_2$, i.e., fully dissipative seller effort: increases in it only steal buyers away from other sellers and leave the overall matching rate unchanged. Also, we argue in the text, $(r/\beta)/qG$ is likely to be small. We assume the condition holds.

Increases in v shift up $\mathcal{B}$ by $dy/dv = hG/\{hG - h_4 A\bar{\imath}_S'(y)\} > 1$, and so increase both θ and y, and thus seller effort as well. Thus q increases, and the change in h is ambiguous. To know the effect on the acceptance rate requires solving the system. Doing so, we get $dy/dv = h^2G/\Delta$, where $\Delta = h^2G -$

$h_1(r/\beta) + (h_1 q_3 - q_1 h_4)Ai'$. This implies that $dy/dv \leq (\geq) 1$ as

$1 + \epsilon \geq (\leq) \big(1 + (q_3/q_2)\big)/|dlnh/dln\theta|$. Thus if effort is completely dissipative, $dy/dv < 1$ and the acceptance rate increases, as in the baseline model.

Appendix D: Rank Determined Response Propensity

Since $BTOM$ is distributed exponentially, median log *BTOM* at the market level is $m = ln2 - ln\lambda \equiv ln2 - X\beta$, where X represents the demand proxies and β their coefficients. However, if it is the top 100α'th percent who respond, then the median *reported BTOM*, $\widetilde{m}$, in a given market will be such that $\exp(-\lambda\widetilde{m}) = \alpha/2$, i.e., $\widetilde{m} = ln2 - ln\alpha - ln\lambda = X\beta - ln\alpha$. If only those with the bottom 100α'th percent outcomes respond, $\widetilde{m} = X\beta - \ln{(2 - \alpha)}$. More generally, so long as the propensity to respond depends on one's relative rank in the outcome distribution, $\widetilde{m}$ will differ m by a constant only. The result does not require the exponential distribution but does require homoskedasticity.

However, there is survey response bias if respondents perceive their rank as other than what it truly is, as in our dynamic model, where sellers misperceive the offer distribution in the short run after a demand shift. If only those who see themselves as in the top 100α'th percent according to the previous period's distribution respond, than sellers will respond if and only if $\lambda_{t-1} BTOM \geq -ln\alpha$, where $ln\lambda_{t-1} = -X_{t-1}\beta$, implying $\exp(-\lambda_t m) = \exp\big((\lambda_t/\lambda_{t-1})ln\alpha\big)/2$, i.e., $m = \ln(ln2 - (\lambda_t/\lambda_{t-1})ln\alpha) - ln\lambda_t \ = \ln(ln2 - \exp(-\Delta X_t \beta)\, ln\alpha) + lnX_t\beta$. It is straightforward to show $-1 < \partial m/\partial(\Delta X_t \beta) < 0$.

If it is those seemingly in the bottom 100α'th percent who respond, $m = \ln(ln2 - \ln{(1 +}$ $\exp(exp{-}\Delta Xt\beta \ln(1{-}\alpha))+Xt\beta$, and $\partial m\partial\Delta Xt\beta{=}xexln2{-}\ln 1{+}ex1{+}ex$, where $x \equiv \exp(-\Delta X_t\, \beta) ln(1 - \alpha)$. Since $x < 0$, this derivative is negative; visual inspection of the extended function on $x \leq 0$ shows its absolute value monotonically increasing and reaching 1 at $x = 0$.

Thus the empirical model when there are no true dynamic effects can be approximated by

$$m = X\beta - \zeta\Delta X\beta + u$$

where u is a regression error, i.e., with mean zero and uncorrelated with X and ΔX, and $0 < \zeta < 1$ is a constant. Under this model, the coefficients on ΔX have the opposite sign of and are no greater in size than those on X. Since, in contrast, the estimated coefficients are of the same sign with those for ΔX being of a much greater size, it is clear that survey response bias of this type can not explain the results.

The empirical model where are true dynamic effects (which we denote γ) can be written as

$$m = X\beta + \Delta X\gamma - \zeta\Delta X\beta + u = X\beta + \Delta X(\gamma - \zeta\beta) + u$$

Again, given the large difference between the coefficient estimates on X and ΔX, accounting for survey response bias would make little difference to the interpretation of our results.

References

Aalen, Odd O., "Heterogeneity in Survival analysis", *Statistics in Medicine*, 1988, Vol. 7, p. 1121-1137.

Abraham, Jesse and Patric Hendershott, "Bubbles in Metropolitan Housing Markets", *Journal of Housing Research*, 1996, Vol. 7, p. 191-207.

Albrecht, James, Anderson Axel, Eric Smith, and Susan Vroman, "Opportunistic Matching in the Housing Market," *International Economic Review,* 2007, Vol. 42, p. 641-664.

Andrew, Mark and Geoff Meen, "House Price Appreciation, Transactions and Structural Change in the British Housing Market: A Macroeconomic Perspective", *Real Estate Economics*, Vol. 31 (1):99-116, 2003.

Anglin, Paul M., "A Summary of Some Data on Buying and Selling Houses", Working paper, University of Windsor, 1994.

Anglin, Paul M., "Determinants of Buyer Search in a Housing Market", *Real Estate Economics*, 1997, Vol. 25 (4):567-589.

Anglin, Paul, Ronald Rutherford and Thomas Springer, "The Trade-off Between the Selling Price of Residential Properties and Time-on-the-Market: The Impact of Price Setting", *Journal of Real Estate Finance and Economics* , 2003, Vol. 26 (1), p.95-111.

Balestra, "Fixed Effect Models and Fixed Coefficient Models", chapter 3 in László Mátyás and Patrick Sevestre, eds., *The Econometrics of Panel Data: Handbook of Theory and Applications*, Kluwer Academic Publishers, 1992.

Baryla, Jr., Edward A., and Leonard V. Zumpano, "Buyer Search Duration in the Residential Real Estate Market: The Role of the Real Estate Agent", *Journal of Real Estate Research,* 1995, Vol. 10(1), p. 1-14.

Baryla, Jr., Edward , Leonard Zumpano and Harold W. Elder, "An Investigation of Buyer Search in the Residential Real estate Market Under Different Market conditions", *Journal of Real Estate Research*, 2000, Vol. 20(1), p.75-91.

Bernheim, Douglas and Jonathan Meer, “How Much Value Do Real Estate Brokers Add? A Case Study”, NBER Working Paper No.13796, 2008.

Caplin, Andrew and John Leahy, “Trading Frictions and House Price Dynamics,” *NBER Working Paper, 2008* No. w14605

Capozza, Dennis, Patric Hendershott and Charlotte Mack, “An Anatomy of Price Dynamics in Illiquid Markets: Analysis and Evidence from Local Housing Markets”, *Real Estate Economics*, 2004, Vol. 32(1), p. 1-32.

Case, Karl E. and Robert J. Shiller, “The Efficiency of the Market for Single-Family Homes”, *The American Economic Review*, 1989, Vol. 79(1), p. 125-137.

Chen, Yongmin and Robert Rosenthal, “On the use of ceiling-price commitments by monopolists”, *RAND Journal of Economics*, Vol. 27(2) Summer 1996, p. 207-220.

Coles, Melvin and Eric Smith, “Cross-Section Estimation of the Matching Function, Evidence from England and Wales”, *Economica*, 1996, Vol. 63(252), p.589-597.

Díaz, Antonia and Belén Jerez, “House Prices, Sales and Time on the Market: A Search-Theoretic Framework”, mimeo, March 2009.

D’Urso, Victoria, “Home Buyer Search Duration and the Internet”, MIT Working Paper, 2002.

Elder, Harold W., Zumpano, Leonard V. and Edward A. Baryla, “Buyer Search Intensity and the Role of the Residential Real Estate Broker”, *Journal of Real Estate Finance and Economics*, 1999, Vol. 18(3), p.351-368.

Federal Trade Commission, *The Residential Real Estate Brokerage Industry*, U.S. Government Printing Office, Washington, 1983.

Gallin, Joseph, “The Long-Run Relationship between House Prices and Income: Evidence from Local Housing Markets”, *Real Estate Economics*, 2006, Vol. 34(3), p. 417-438.

Gabriel, Stuart A., Joe P. Mattey and William L. Wascher, "Compensating differentials and evolution in the quality-of-life among U.S. states", *Regional Science and Urban Economics*, 2003, Vol. 33, p. 619-49.

Genesove, David and Christopher Mayer, "Equity and Time to Sale in the Real Estate Market", *The American Economic Review*, 1997, Vol. 87(3), p.255-269.

Genesove, David and Christopher Mayer, "Loss Aversion and Seller Behavior: Evidence from the Housing Market", *Quarterly Journal of Economics*, 2001, Vol. 116(4), p.1233-1260.

Glaeser, Edward L., Joseph Gyourko, "Urban Decline and Durable Housing", *Journal of Political Economy*, 2005, Vol. 113 (2), p.345-375.

Glower, M., D. R. Haurin, and P. H. Hendershott, "Selling Time and Selling Price: The Influence of Seller Motivation", *Real Estate Economics,* 1998, Vol. 26(4), p.719-40.

Gyourko, Joseph and Joseph Tracy, "The Structure of Local Public Finance and the Quality of Life", Journal of Political Economy, 1991, vol. 99(4),p. 774-806.

Harding, John, Stuart Rosenthal and C. F. Sirmans, "Estimating Bargaining Power in the Market for Existing Homes", *Review of Economics and Statistics*, 2003, Vol. 83(1), p. 178-188.

Haurin, Donald, "The Duration of Marketing Time of Residential Housing", *American Real Estate and Urban Economics Association Journal*, 1988, Vol. 16(4), p.396-410.

Hendel, Igal , Aviv Nevo and Francois Ortalo-Magne , "The Relative Performance of Real Estate Marketing Platforms: MLS versus FSBOMadison.com," *American Economic Review*, forthcoming.

Hort, Katinka, "Prices and turnover in the market for owner-occupied homes", *Regional Science and Urban Economics*, 2000, Vol. 30, p.99–119.

Hosios, Arthur J., "On the Efficiency of Matching and Related Models of Search and Unemployment", *Review of Economic Studies*, 1990, Vol. 57, p. 279-98.

Hougaard, P., “Survival Models for Heterogeneous Populations Derived from Stable Distributions”, Biometrika, 1986, Vol. 73(2), p. 387-396.

Krainer, John, “A Theory of Liquidity in Residential Real Estate Markets”, *Journal of Urban Economics*, 2001, Vol. 49, p.32-53.

Johnson, Norman, Samuel Kotz and N. Balakrishnan, *Continuous Univariate Distributions*, Volume 1, second edition, 1994.

Julien, Benoit, John Kennes and Ian Paul King, “Bidding for Labor”, *Review of Economic Dynamics*, 2000, Volume 3, p. 619-649.

Lazear, Edward, “Retail Pricing and Clearance Sales”, *American Economic Review*, 1986, Vol. 76, p. 14-32.

Leung, C. K. Y., G. C. K. Lau, and Y. C. F. Leong, “Testing Alternative Theories of the Property Price-Trading Volume Correlation”, *Journal of Real Estate Research*, 2002, Vol. 23, p. 253-264.

Levitt, Steven and Chad Syverson, “Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate”, *Review of Economics and Statistics*, 2005, Vol. 90(4), p.599-611.

Lucas, Robert E., Jr., “Expectations and Neutrality of Money”, *Journal of Economic Theory*, 1972, Vol. 4, p.101-24.

McMillen, D., and J. McDonald, “Urban land value functions with endogenous zoning”, *Journal of Urban Economics*, 1991, Vol. 9, p. 14-27.

Mortensen, Dale, “Labor Search and Market Analysis”, chapter 15 in O. C. Ashenfelter, R. C. Layard, *Handbook of Labor Economics*, North Holland, 1986.

Margaret Stevens, "New Microfoundations for the Aggregate Matching Function", *International Economic Review, 2007,* Vol. 48, p. 847-868.

Merlo, Antonio and Francois Ortalo-Magne, "Bargaining over Residential Properties: Evidence from England", *Journal of Urban Economics*, 2004, Vol. 56, p. 192-216.

Mikhed, Vyacheslav and Petr Zemčik, “Do House Prices Reflect Fundamentals? Aggregate and Panel Data Evidence”, *Journal of Housing Economics*, 2009, Vol. 18(2), 140-149.

Moretti, Enrico. “Human Capital Externalities in Cities” in *Handbook of regional and urban economics. Volume 4. Cities and geography*, 2004, pp. 2243-91, Elsevier.

Mortensen, Dale and Christopher Pissaridies, “Job Reallocation, Employment Fluctuations, and Unemployment”, *Handbook of Macroeconomics*, 1996.

Ngai, Rachael and Silvana Tenreyro (2009), “Hot and Cold Seasons in the Housing Market”, LSE wp, 2009.

Novy-Marx, Robert, “Hot and Cold Markets,” *Real Estate Economics*, 2009, Vol. 37, p. 1-22.

Otalo-Magné, F. and Rady, S., “Housing Transactions and Macroeconomic Fluctuations: A Case Study of England and Wales”, *Journal of Housing Economics*, 2004, Vol. 13, p. 288–304.

Pissardies, Christopher, *Equilibrium Unemployment Theory*, Second edition, 2002.

Quigley, John M. and Steven Raphael, “Is Housing Unaffordable? Why Isn't It More Affordable? “ *Journal of Economic Perspectives,* 2004, Vol. 18(1), p. 191-214.

Roback, Jennifer, “Wages, Rents and the Quality of Life”, *Journal of Political Economy*, 1982, Vol. 90 (December), p. 1257-1278.

Rubinstein, Ariel and Asher Wolinsky, “Equilibrium in a Market with Sequential Bargaining”, *Econometrica*, 1985, Vol. 53(5), p. 1133-1150.

Sass, Tim R., “A Note on Optimal Price Cutting Behavior under Demand Uncertainty”, *The Review of Economics and Statistics*, 1988, Vol. 70(2), p. 336-339.

Sinai, Todd, “Taxation, User-Cost, and Household Mobility Decisions”, University of Pennsylvania mimeo, 1997.

Stein, Jeremy, “Prices and trading volume in the housing market: A model with down-payment effects”, *Quarterly Journal of Economics*, 1995, Vol. 110(2), p. 379-406.

Vaupel, J. W., K. G. Manton and E. Stallard, "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality", *Demography*, 1979, Vol. 16, p. 439-54.

Wallace, Nancy, "The market effects of zoning undeveloped land: Does zoning follow the market?" *Journal of Urban Economics*, 1988, Vol. 23, p.307-326.

Washington Center for Real Estate Research, *Buying or Selling a Home in Washington: 2002 Survey of Recent Home Buyers*, November 2003.

Wheaton, William C., "Vacancy, Search and Prices in a Housing Market Matching Model", *Journal of Political Economy*, Vol. 2(6), p. 1270-92.

Williams, Joseph, "Pricing Real Assets with Costly Search", *Review of Financial Studies*, 1995, Vol. 8(1), p. 55-90.

Zorn, Thomas and William Sackley, "Buyers' and Sellers' Markets: A Simple Rational Expectations Search Model of the Housing Market*", Journal of Real Estate Finance and Economics*, 1991, Vol. 4(3), p.315-32.

Table 1: Sample Summary Statistics

Variables	Seller Sample		Buyer Sample		Joint Sample	
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Seller Time on Market	7.32	5.60			7.60	5.73
ln (Seller Time on Market)	1.73	0.77			1.77	0.77
Buyer Time on Market			8.18	4.42	8.14	4.25
ln(Buyer Time on Market)			2.02	0.38	2.02	0.37
Numbers of Homes Visited			9.96	4.34	9.94	4.27
ln (Homes Visited)			2.17	0.57	2.17	0.57
ln (Average Income)	3.45	0.32	3.43	0.33	3.45	0.32
ln (Population)	14.24	0.94	14.22	0.94	14.24	0.94
Changes in Average Income	0.041	0.025	0.043	0.024	0.041	0.025
Changes in Population	0.012	0.012	0.013	0.013	0.012	0 .012
ln Theta					0.25	0.82
# of Observations	1894		2372		1636	
# of Responses per MSA X Year	7.5	15.3	21.6	55.8	38.2	79.5
	Samples with OFHEO Price Information					
Changes in Price Index	0.065	0.053	0.061	0.057	0.063	0.060
% markets with Price Depreciation	0.051	0.22	0.074	0.26	0.068	0 .25
# of observations	1721		2183		1464	

Note: The main data source is the micro-level data from NAR home buyer and seller survey 1987-2007, aggregated to the MSA/year level. Annual population and income data are obtained from the Bureau of Economic Analysis. Annual house price indices are from the Office of Federal Housing Enterprise Oversight. All changes indicate annual changes in terms of percentage growth. Theta indicates buyer to seller ratio.

Table 2: Impact of Housing Demand Shocks on Search Behavior

	Seller Time on Market (*STOM*)		Buyer Time on Market (*BTOM*)		Homes Visited (*NHVIS*)	
	(1)	(2)	(3)	(4)	(5)	(6)
Population	-1.19	-1.14	-0.44	-0.43	-0.50	-0.47
	(0.49)	(0.46)	(0.15)	(0.15)	(0.18)	(0.18)
Average Income	-1.43	-0.42	-0.09	-0.001	-0.44	-0.25
	(0.77)	(0.78)	(0.26)	(0.29)	(0.34)	(0.38)
Δ Population		-14.44		-1.88		-4.01
		(3.53)		(1.17)		(1.79)
Δ Average Income		-6.98		-0.46		-0.98
		(1.27)		(0.42)		(0.81)
			Short Run Effects			
Population		-15.58		-2.31		-4.48
		(3.55)		(1.14)		(1.83)
Average Income		-7.40		-0.46		-1.23
		(1.20)		(0.40)		(0.74)
F-Stat*		0.03		0.10		0.09
(p-value)		(0.87)		(0.75)		(0.77)
# of observations	1894	1894	2372	2372	2372	2372

Note 1: The estimation sample is the NAR home buyer and seller surveys (1987-2007). The regressions are at the MSAXyear level, weighted by the number of individual responses in a given MSA for a given year. Robust standard errors are reported in brackets and are adjusted for the intra-MSA correlation. All specifications include MSA fixed effects and year dummies (year of survey and year of transaction). All variables are in logs, with Δx indicating annual changes $lnx_t - lnx_{t-1}$. Note 2: F-stat reports the F-statistic for the test of the null hypothesis that the ratio of the coefficient on average income growth to that on the average income equals the corresponding ratio for population.

Table 3: Impact of Housing Demand Shocks on Matching Function Variables

	Seller Contact Hazard (lnq)		Buyer Contact Hazard (lnh)		Buyer-seller ratio ($ln\theta$)	
	(1)	(2)	(3)	(4)	(5)	(6)
Population	0.67	0.60	-0.04	-0.02	0.69	0.60
	(0.50)	(0.48)	(0.21)	(0.20)	(0.48)	(0.44)
Average Income	1.36	0.45	-0.47	-0.30	1.82	0.73
	(0.82)	(0.86)	(0.45)	(0.47)	(0.80)	(0.79)
Δ Population		11.57		-2.96		14.61
		(3.96)		(2.03)		(3.61)
Δ Average Income		6.45		-1.19		7.64
		(1.50)		(0.92)		(1.35)
			Short Run Effects			
Population		12.17		-2.98		15.21
		(4.02)		(2.05)		(3.61)
Average Income		6.90		-1.49		8.37
		(1.45)		(0.85)		(1.31)
F-Stat*		0.02		.01		0.66
(p-value)		(0.90)		(0.91)		(0.20)
# of observations	1636	1636	1636	1636	1636	1636

Note 1: The estimation sample is the NAR home buyer and seller surveys (1987-2007). This table shows ordinary least squares regressions at the MSAXyear level, weighted by the number of individual responses for each observation. Robust standard errors are reported in brackets and are adjusted for the intra-MSA correlation. All specifications include MSA fixed effects and year dummies (year of survey and year of transaction). All variables are in logs, with Δx indicating $lnx_t - lnx_{t-1}$. Note 2: F-stat reports the F-statistic for the test of the null hypothesis that the ratio of the coefficient on average income growth to that on the average income equals the corresponding ratio for population.

Table 4: Impact of Housing Demand Shocks on House Prices

	OFHEO Price Index (1)	Transaction/List Price (2)
Population	0.154	0.039
	(0.021)	(0.020)
Average Income	0.063	0.005
	(0.028)	(0.026)
Δ Population	1.391	0.408
	(0.375)	(0.129)
Δ Average Income	0.437	0.120
	(0.067)	(0.050)
Lagged OFHEO House Price Index	0.938	
	(0.011)	
# of observations	2535	2705

Note: The estimation sample is the NAR home buyer and seller surveys (1987-2007). The regressions are at the MSAXyear level, weighted by the number of individual responses in a given MSA for a given year. Robust standard errors are reported in brackets and are adjusted for the intra-MSA correlation. All specifications include MSA fixed effects and year dummies (year of survey and year of transaction). All variables are in logs, with Δx indicating annual changes $\ln x_t - \ln x_{t-1}$.

Table 5: Impact of Housing Demand Shocks on Search Behavior, Controlling for Age Distribution

	STOM: Seller Time on the market				*BTOM*: Buyer Time on the Market				*NHVIS*: Number of Homes Visited			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Population	-0.25	-0.58	-0.18	-0.36	-0.24	-0.10	-0.22	-0.15	-0.38	-0.28	-0.34	-0.39
	(0.66)	(0.78)	(0.64)	(0.76)	(0.24)	(0.25)	(0.25)	(0.26)	(0.21)	(0.24)	(0.20)	(0.24)
Average Income	-1.37	-1.15	-0.26	-0.31	0.10	0.01	0.14	0.12	-0.83	-0.90	-0.79	-0.69
	(1.06)	(1.06)	(1.17)	(1.17)	(0.32)	(0.32)	(0.38)	(0.38)	(0.41)	(0.43)	(0.49)	(0.49)
% Pop. Aged		7.94		3.35		-3.52		-2.39		-2.78		0.001
25-40		(7.13)		(7.51)		(1.86)		(1.99)		(2.30)		(2.58)
Δ Population			-15.98	-17.12			-1.43	-0.85			-3.75	-2.48
			(4.84)	(4.95)			(1.47)	(1.49)			(2.15)	(2.14)
Δ Avg. Income			-5.77	-5.52			-0.19	-0.26			-0.22	-0.38
			(1.76)	(1.75)			(0.51)	(0.51)			(0.82)	(0.77)
Δ % Pop. Aged				31.98				-11.65				-28.04
25-40				(22.40)				(8.09)				(10.76)
# of obs.	1455	1455	1455	1455	1751	1751	1751	1751	1751	1751	1751	1751

Note: This table repeats estimation in Table 2 with additional controls on time-varying MSA-level age distribution. The main data source is the NAR home buyer and seller surveys (1987-2007). The demographic information is obtained from the Census. The regressions are at the MSAXyear level, weighted by the number of individual responses in a given MSA for a given year. Robust standard errors are reported in brackets and are adjusted for the intra-MSA correlation. All specifications include MSA fixed effects and year dummies (year of survey and year of transaction). All variables are in logs, with Δx indicating annual changes $lnx_t - lnx_{t-1}$.

Table 6: Impact of Housing Demand Shocks on Search Behavior, Controlling for Mean Respondent Characteristics

	STOM: Seller Time on the market				*BTOM*: Buyer Time on the Market				*NHVIS*: Number of Homes Visited			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Population	-0.84	-0.70	-0.52	-0.37	-0.77	-0.77	-0.75	-0.75	-0.88	-0.92	-0.72	-0.78
	(1.21)	(1.22)	(1.26)	(1.26)	(0.33)	(0.33)	(0.33)	(0.33)	(0.44)	(0.41)	(0.43)	(0.41)
Average Income	-1.68	-1.15	-0.93	-0.82	-0.02	-0.02	0.21	0.20	-0.55	-0.76	-0.57	-0.75
	(1.20)	(1.06)	(1.38)	(1.39)	(0.37)	(0.37)	(0.44)	(0.44)	(0.49)	(0.47)	(0.63)	(0.61)
Δ Population			-12.21	-12.77			-2.44	-2.44			-4.12	-3.82
			(5.50)	(5.48)			(1.46)	(1.46)			(2.50)	(2.37)
Δ Avg. Income			-3.64	-3.74			-0.71	-0.71			-0.11	-0.14
			(1.87)	(1.87)			(0.55)	(0.55)			(0.87)	(0.85)
Respondent		-0.19		-0.21		0.03		0.03		0.34		-0.34
Income		(0.14)		(0.14)		(0.06)		(0.06)		(0.08)		(0.08)
Respondent Age		0.02		0.05		-0.06		-0.05		0.10		0.10
(* 100)		(0.50)		(0.50)		(0.23)		(0.23)		(0.20)		(0.20)
Whether Couple		-0.18		-0.17		-0.0007		0.0025		0.13		0.13
		(0.12)		(0.12)		(0.08)		(0.08)		(0.07)		(0.07)
Whether have		-0.07		-0.09		-0.02		-0.023		0.02		0.02
children		(0.56)		(0.08)		(0.04)		(0.040)		(0.04)		(0.04)
F-Stat*		1.92		2.03		0.15		0.14		6.42		6.39
(p-value)		(0.10)		(0.09)		(0.96)		(0.97)		(0.00)		(0.00)
# of obs.	1368	1368	1368	1368	1622	1622	1622	1622	1622	1622	1622	1622

Note: This table repeats estimation in Table 2 with additional controls on the mean respondent characteristics from a given MSA in a given year. The main data source is the NAR home buyer and seller surveys (1987-2007). The regressions are at the MSAXyear level, weighted by the number of individual responses in a given MSA for a given year. Robust standard errors are reported in brackets and are adjusted for the intra-MSA correlation. All specifications include MSA fixed effects and year dummies (year of survey and year of transaction). All variables are in logs, with Δ□ indicating annual changes □□□$_{□}$ − □□□$_{□-1}$.

Table 7: Impact of Housing Demand Shocks on Search Behavior, Controlling for Internet Use

	Seller Time on Market *(STOM)*		Buyer Time on Market *(BTOM)*		Homes Visited *(NHVIS)*	
	(1)	(2)	(3)	(4)	(5)	(6)
Population	-1.29	-1.32	-.43	-.40	-.47	-.43
	(.48)	(.48)	(.15)	(.16)	(.18)	(.18)
Average Income	-.45	-.42	.002	-.01	-.24	-.26
	(.81)	(.80)	(.29)	(.29)	(.38)	(.38)
Δ Population	-14.15	-14.18	-1.88	-1.89	-4.02	-4.03
	(3.75)	(3.74)	(1.17)	(1.18)	(1.80)	(1.76)
Δ Average Income	-7.27	-7.30	-.46	-.43	-.98	-.95
	(1.35)	(1.35)	(.42)	(.42)	(.81)	(.80)
Internet Use		-0.27		0.24		0.30
		(0.25)		(0.11)		(0.12)
# of observations	1639	1639	2364	2364	2364	2364

Note: This table repeats estimation in Table 2 with additional controls on Internet use. The estimation sample is the NAR home buyer and seller surveys (1987-2007). The regressions are at the MSAXyear level, weighted by the number of individual responses in a given MSA for a given year. Robust standard errors are reported in brackets and are adjusted for the intra-MSA correlation. All specifications include MSA fixed effects and year dummies (year of survey and year of transaction). All variables are in logs, with Δ□ indicating annual changes $□□□_{□} - □□□_{□-1}$.

Table A1: Sample Statistics of Home Search Behavior from Alternative Surveys

	Federal Trade Commission		American Housing Survey		Anglin	Washington Center for Real Estate Research	
	Median	Mean	Median	Mean	Median	Median	Median
Seller Time on Market	One month	Two months					
Buyer Time on Market					9-10 weeks		12 weeks
Number of Home Visits	11	17	10.7	15.1	10-11	10	
Years	1979-80					2003 (check)	
Sample	U.S. consumer panel		U. S.		Windsor, ON Canada	Washington State	
Response Rates	83% in first stage		More than 90%[24]		About 60%	Unknown	

Note: Sources: (i) Federal Trade Commission: Federal Trade Commission, 1983. The survey population was recent movers that belonged to a standing, nationwide consumer panel used for marketing purposes. The nature of the sampling in the second stage was such that response rates can not be calculated. (ii) American Housing Survey: micro data from the American Housing Survey (National), 1999–2007. Since 1999, respondents are asked for the number of homes "looked at *before* choosing [the] one they bought" [our italics]. We thus add 1 to each response. The reported number are weighted (by number of recent movers) averages across MSAs. (iii) Anglin: Anglin, 1994, 1997. (iv) Washington Center for Real Estate Research: Washington Center for Real Estate Research, November 2003.

[24] See www.aspe.hhs.gov/hsp/06/catalog-ai-an-na/AHS-MS.htm and www.aspe.hhs.gov/hsp/06/catalog-ai-an-na/AHS.htm .

Figure 1: Equilibrium in the Baseline Model

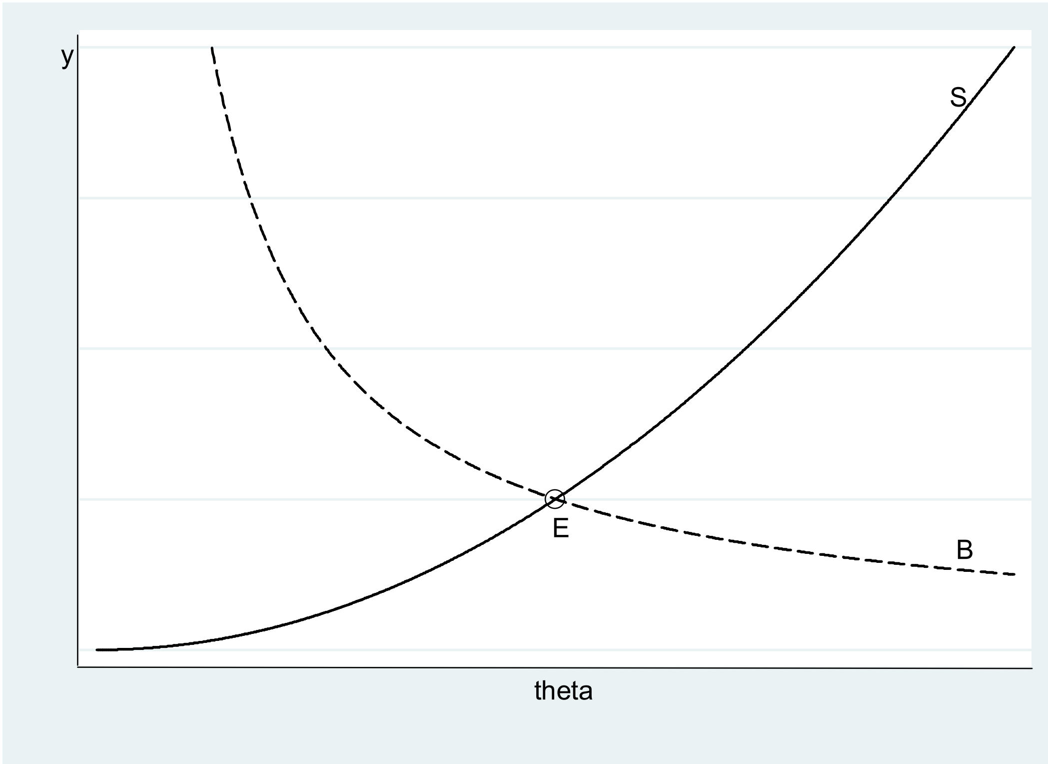

Figure 2: A Dynamic Extension of the Baseline Model

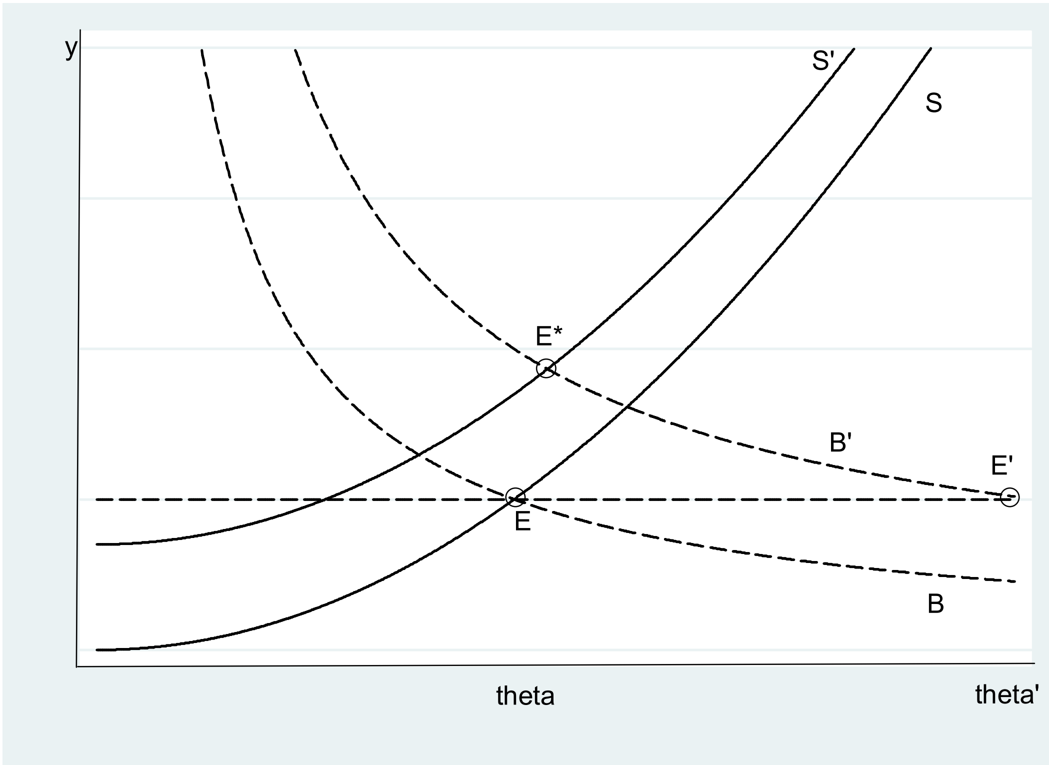

Figure 3: The Evolution of Online Home Search