The Housing Market(s) of San Diego*

Tim Landvoigt
Stanford

Monika Piazzesi
Stanford & NBER

Martin Schneider
Stanford & NBER

July 2011

Abstract

This paper uses a quantitative assignment model of the housing market to understand the cross section of house prices within a metro area. In the model, equilibrium house prices are determined to assign indivisible houses that differ by quality to movers who differ by age, income and wealth. We measure distributions of house prices, house qualities and mover charateristics from micro data on the San Diego Metro Area over the recent housing boom. The model suggests that cheaper credit for poor agents was important in generating higher capital gains at the low end of the market.

*Preliminary and incomplete. Comments welcome! Email addresses: timl@stanford.edu, piazzesi@stanford.edu, schneidr@stanford.edu. We thank Lars Ljunqvist, Francois Ortalo-Magne, Sergio Rebelo, and seminar participants at Boston University, Carnegie-Mellon University, ECB, Federal Reserve Board, Kellogg, Munich, NYU, Princeton, UNC as well as conference participants at CASEE 2010 at Arizona State, ERID Duke Conference 2010, Duke Macoeconomics Conference 2011, EEA Congress 2010 in Glasgow, HULM Conference 2010 in St. Louis, Richmond Fed 2009 PER Conference, SED 2010 in Montreal, SITE 2010, the 2010 USC Macro Conference, and the 2011 Jerusalem Conference.

1 Introduction

During the recent housing boom, there were large differences in capital gains across houses, even within the same metro area. Figure 1 illustrates the basic stylized fact for San Diego County, California. Every dot corresponds to a home that was sold in both 2000 and 2005. On the horizontal axis is the 2000 sales price. On the vertical axis is the annualized real capital gain between 2000 and 2005. The solid line is the capital gain predicted by a regression of capital gain on log price. It is clear that capital gains during the boom were much higher on low end homes. For example, the average house worth \$200,000 in the year 2000 appreciated by 17% (per year) over the subsequent five years. In contrast, the average house worth \$500,000 in the year 2000 appreciated by only 12% over the subsequent five years.

This paper considers a quantitative model of the housing market in the San Diego metro area over the boom period. Its key feature is that houses are indivisible and movers are assigned, in equilibrium, to one of a large number of house types. We use the model to

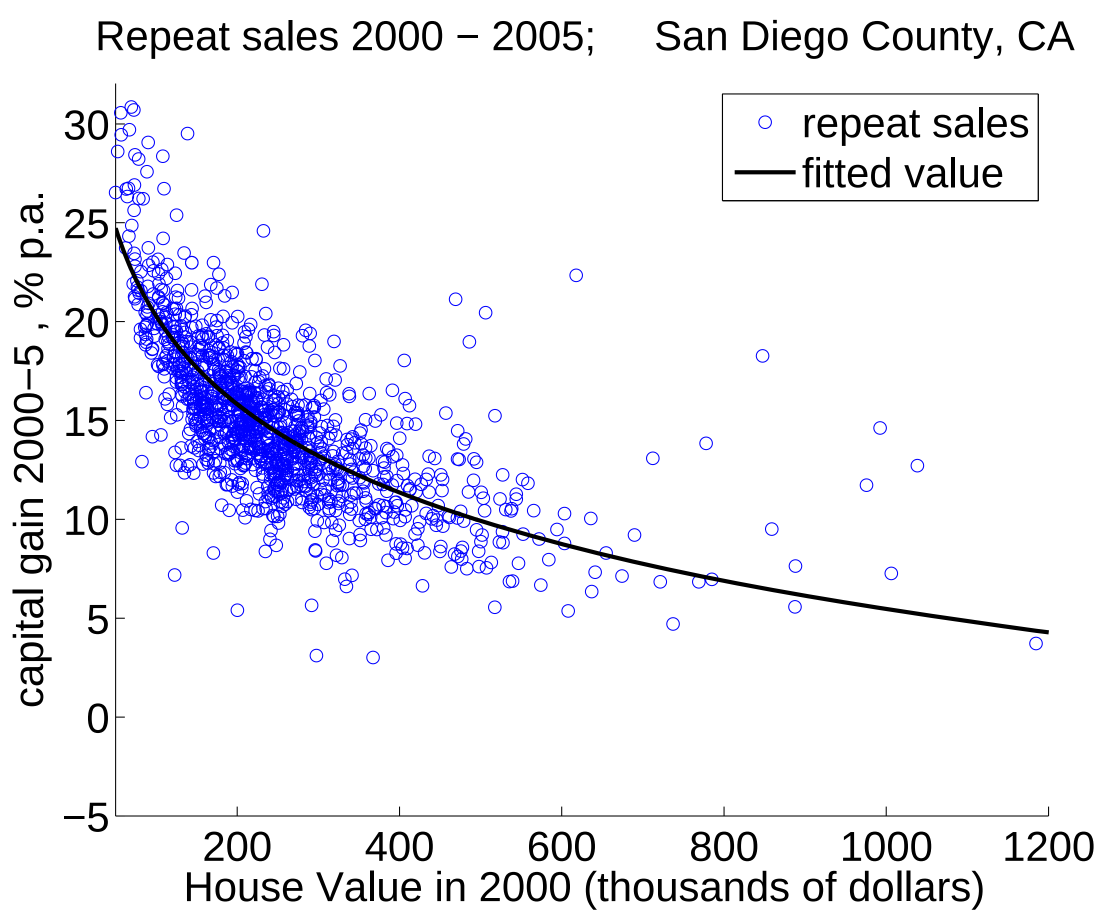

Figure 1: Every dot represents a residential property in San Diego County, CA, that was sold in 2000 and had its next sale in 2005. The horizontal axis shows the sales price in 2000. The vertical axis shows the real capital gain per year (annualized change in log price less CPI inflation) between 2000 and 2005.

connect various changes in the San Diego housing market with the cross section of capital gain there. In particular, we look at changes in the composition of houses that were transacted, shifts in the distribution of movers' characteristics, and the availability of cheap credit. The main finding is that the availability of cheap credit has larger effects on housing demand at the low end of the market, with strong implications for relative prices. In addition, shifts in the distributions of houses and movers are important for understanding the cross section of prices. For example, the relatively larger number of low quality houses transacted during the boom led to richer marginal investors at the low end, which also contributed to higher capital gains there.

In the model, movers meet houses. Houses differ by quality: there is a continuum of indivisible houses that differ in the flow of housing services provided. Movers differ by age, income, and wealth: their demand for housing is derived from an intertemporal savings and portfolio choice problem with transaction costs and collateral constraints. In equilibrium, prices adjust to induce agents with lower demand to move into lower quality houses. The distribution of equilibrium prices thus depends on the distribution of movers' characteristics as well as on the distribution of house qualities.

To implement the model quantitatively, we use micro data to measure the distribution of movers' characteristics and the quality distribution of transacted houses. We do this both for the year 2000 and for the year 2005 – the peak of the boom. We then compute model predictions for equilibrium prices in these two years and thus the cross section of capital gains by quality. We compare those predictions to those from a repeat sales model estimated on transaction data. We also explore model-implied capital gains for different assumptions on changes in the environment; in particular, we capture cheaper credit by lower interest rates and downpayment constraints.

Our exercise fits into a tradition that links asset prices to fundamentals through household optimality conditions. For housing, this tradition has given rise to the "user cost equation": the per-unit price of housing is such that all households choose their optimal level of a divisible housing asset. With a *single* per-unit price of housing, capital gains on all houses are the same. In our model, there is no single user cost equation since there many types of indivisible houses, with marginal investors different across house types.[1] Instead, there is a *separate* user cost equation for every house type, each reflecting the borrowing costs, transaction costs and risk premia of only those movers who buy that house type. Changes in the environment thus typically give rise to a nondegenerate cross section of capital gains.

Our results reflect this family of pricing equations in two ways. First, changes in the environment that more strongly affect a subset of movers will more strongly affect prices of houses which those movers buy. For example, lower downpayment constraints more strongly affect poor households for whom the constraint is more likely to be binding. As a result, lower downpayment constraints lead to higher capital gains at the low end of the market.

Second, higher moments of the quality and mover distributions matter. For example, we show that the quality distribution of transacted homes in San Diego County at the peak of

[1] In contrast, the user cost equation determines a unique price per unit of housing, so every investor is marginal with respect to every house.

the boom had fatter tails than at the beginning of the boom. This implied, in particular, that relatively lower quality homes had to be assigned to relatively richer households than before the boom. For richer households to be happy with a low quality home, homes of slightly higher quality had to become relatively more expensive. The price function thus had to become steeper at the low end of the market, which contributed to high capital gains in that segment.

The paper proceeds as follows. Section 2 presents evidence on prices and transactions by quality segment in San Diego County. Section 3 presents a simple assignment model to illustrate the main effects and the empirical strategy. Section 4 introduces the full quantitative model and results.

Related Literature

Assignment models with indivisible heterogeneous goods and heterogeneous agents have been used in several areas of economics, most prominently to study labor markets where firms with different characteristics hire workers with different skill profiles (for an overview, see Sattinger, 1993.) In the context of housing, an early reference is Kaneko (1982). Caplin and Leahy (2010) characterize comparative statics of competitive equilibria in a general setting with a finite number of agents and goods. Stein (1995) and Ortalo-Magne and Rady (2006) consider models with two types of houses and credit constraints. Rios-Rull and Sanchez-Marcos (2008) study a quantitative business cycle model with large and small houses.

A number of empirical studies has compared house price dynamics across price segments within a metro area and linked them to changes in amenities or changes in demographics: for example, Poterba (1991), Case and Mayer (1996), Case and Shiller (2005), Guerrieri, Hartley, and Hurst (2010). In our setup, amenities and demographics are captured as part of the quality distribution of transacted homes and the distribution of mover characteristics, respectively. The assignment equilibrium says how changes in those distributions translate into changes in the price distributions, as well as how quality and mover characteristics should be correlated in the cross section.

Since we measure the quality distribution of transacted homes directly from the data, we do not take a stand on where the supply of houses comes from. A more elaborate model might add an explicit supply side, thus incorporating effects of the availability of land to developers (as in Glaeser, Gyourko, and Saks, 2005), gentrification (as in Guerrieri, Hartley, and Hurst, 2010) or sellers' choice of when to put their house on the market (Piazzesi and Schneider, 2009.) At the same time, any model with a supply side also gives rise to an equilibrium distribution of transacted homes that has to be priced and assigned to an equilibrium distribution of movers. To the extent that the distribution of transacted homes and movers in the more elaborate model match the data, the assignment and pricing "module" would thus be the same as the one we study.

Empirical studies have suggested that credit constraints matter for house prices at the regional level. Lamont and Stein (1999) show that house prices react more strongly to shocks in cities where more households are classified as "borrowing constrained". Mian and Sufi (2010) show for the recent US boom that house price appreciation and borrowing were correlated across zip codes. Mian and Sufi (2009) show that areas with many subprime borrowers saw a lot of borrowing even though income there declined. Our focus is on the relationship between credit, income, and house prices at the property level.

A number of papers have recently used asset pricing models to assess the role of cheap credit for house prices over the 2000s boom (for example, Himmelberg, Mayer, and Sinai 2005, Kiyotaki, Michaelides, and Nikolov 2010, Favilukis, Ludvigson, and Van Nieuwerburgh 2010, Glaeser, Gottlieb, and Gyourko 2010.) These models assume that houses are homogeneous and determine a single equilibrium house price per unit of housing capital. As a result, capital gains on all houses are identical and the models cannot speak to the effect of cheap credit on the cross section of capital gains, the focus of our study.

2 Facts

In this section we present facts on house prices and the distribution of transacted homes during the recent boom. We study the San-Diego-Carlsbad-San-Marcos Metropolitan Statistical Area (MSA) which coincides with San Diego County, California.

2.1 Data

We use several data sources provided by the US Census Bureau. The 2000 Census contains a count of all housing units in San Diego County. We also use the 2000 Census 5% survey sample of households that contains detailed information on house and household characteristics for a representative sample of about 25,000 households in San Diego County. We obtain information for 2005 from the American Community Survey (ACS), a representative sample of about 6,500 households in San Diego Country. A unit of observation in the Census surveys is a dwelling, together with the household who lives there. The surveys report household income, the age of the head of the household, housing tenure, as well the age of the dwelling, and a flag on whether the household moved in recently. For owner-occupied dwellings, the census surveys also report the house value and mortgage payments. From county deeds records we have the complete set of closing prices for housing transactions in San Diego County between 1997 and 2008. We use this sample to build a data set of 71,000 repeat sales over the period 1997-2008. Details are in the appendix.

2.2 The cross section of house prices and qualities

Suppose there exists a one-dimensional quality index that households care about. At any point in time, quality is then reflected one-for-one in the house price. It is thus natural to pick a base year and measure quality in terms of base year price. The base year quality distribution can be read off the cross section of base year transaction prices. We now describe how we measure price changes over time conditional on quality, as well as changes over time in the quality distribution.

Between distinct points in time, the price of an individual house can change because of either (i) idiosyncratic shocks that only affect the given house or (ii) common shocks that equally affect all houses of the same initial quality. Common shocks can affect prices because they bring about common changes in quality (such as depreciation or improvements). They can also lead to revaluation at constant quality if, for example, more houses of similar quality become available, or if the typical buyer in some quality segment obtains cheaper credit.

Our structural model below speaks to the common component in house prices that affects all houses of the same initial quality. The model takes idiosyncratic shocks to be exogenous. We isolate the quality-specific common component by estimating a repeat sales model that allows expected capital gains to depend on quality. Formally, let p_t^i denote the log price of a house i at date t. We assume that capital gain on house i between dates t and $t+1$ is

$$p_{t+1}^i - p_t^i = a_t + b_t p_t^i + \varepsilon_{t+1}^i, \tag{1}$$

where the idiosyncratic shocks ε_{t+1}^i are independent over time with identical variance σ_{t+1}^2, and where a law of large numbers if assumed to hold in the cross section of houses.

If housing were a homogenous capital good, then it should not be possible to forecast the capital gain using the initial price level p_t, that is, $b_t = 0$. The expected capital gain on all houses would be the same (and equal to a_t), much like the expected capital gain is the same for all shares of a given company. More generally, a nonzero coefficient for b_t indicates that quality matters for capital gains. For example $b_t < 0$ means that prices of low quality houses that are initially cheaper will on average have higher capital gains. In contrast, $b_t > 0$ says that expensive houses are expected to appreciate more (or depreciate less). The loglinear functional form (1) for the common component works well for describing house prices in our sample. In particular, we have checked the implied loglinear forecasting equations for capital gains at all horizons. Nonparametric regressions of capital gains from repeat sales on log price reveal only small deviations from linearity.

The statistical model (1) does not distinguish between different types of common shocks that hit all houses of the same quality. Between any two years, one would typically expect both common changes in quality and revaluation. At this point, the distinction is not important. However, in our structural model below we will be able to compare different sources of shocks – in the structural model any given experiment will be taking a stand on quality changes versus revaluation.

Estimation

Equation (1) differs from a typical time series model for returns in that the coefficients are time dependent. It is possible to identify a separate set of coefficients for every date because we have data on many repeat sales. Suppose $\left(p_t^i, p_{t+k}^i\right)$ is a pair of log prices on transactions of the same house that took place in years t and $t+k$, respectively. Equation (1) implies a conditional distribution for the capital gain over k periods,

$$p_{t+k}^i - p_t^i = \hat{a}_{t,t+k} + \hat{b}_{t,t+k}\; p_t^i + \varepsilon_{t,t+k}^i, \tag{2}$$

where the coefficients $\hat{a}_{t,t+k}$, $\hat{b}_{t,t+k}$ as well as $\mathrm{var}\left(\varepsilon_{t,t+k}^i\right)$ are derived by iterating on equation (1). We estimate the parameters (a_t, b_t) by GMM; the objective function is the sum of squared prediction errors, weighted by the inverse of their variance.

Table 1 reports point estimates based on the 71,000 repeat sales in San Diego County during 1997-2008. The first row shows the sequence of estimates for the intercept a_t. The intercept is positive for expected capital gains during the boom phase 2000-2005, and negative during 2006-2008, reflecting average capital gains during those two phases. The second row shows the slope coefficients b_t. During the boom phase, the slopes are strongly negative. For example, for the year 2002 we have $b_t = -.09$, that is, a house worth 10% more in 2002 appreciated by .9% less between 2002 and 2003. During the bust phase, the relationship was reversed: positive b_ts imply that more expensive houses depreciated relatively less.

In a second step, we estimate the variances of the residuals (σ_t^2) by maximum likelihood, assuming the shocks ε_{t+1}^i are normally distributed.The results are reported in the third row of Table 1. Volatility is around 9% on average, slightly higher then the idiosyncratic volatility of 7% reported by Flavin and Yamashita (2001).[2] Another interesting pattern is that idiosyncratic volatility increased by more than half in the bust period.[3]

TABLE 1: ESTIMATED COEFFICIENTS FROM REPEAT SALES MODEL FOR SAN DIEGO

year	97	98	99	00	01	02	03	04	05	06	07
a_t	0.13	0.35	0.76	1.29	1.41	1.30	0.87	0.60	-0.56	-1.09	-3.18
	(0.06)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)	(0.07)	(0.10)	(0.12)
b_t	-0.001	-0.02	-0.05	-0.093	-0.10	-0.09	-0.05	-0.04	0.04	0.07	0.22
	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.01)	(0.01)	(0.01)
σ_t^i	10.6	9.2	8.8	8.3	8.6	8.2	8.0	8.4	9.7	11.4	13.8

Note: This table reports estimates for coefficients a_t and b_t in equation (1) for the indicated years. The numbers in brackets are standard errors. The data are 71,000 repeat sales in San Diego Country from 1997-2008.

Quality distributions

Let Φ_0 denote the cumulative distribution function (cdf) of log transaction prices in the base year 0, which we take to be the year 2000. With base year price as the quality index, the cdf of house qualities in the base year 0 is $G_0(p_0) = \Phi_0(p_0)$. The repeat sales model describes price movements of houses that exist in both years 0 and in later years. In particular, between years 0 and $t > 0$, say, common shocks move the price of the average house that starts at quality p_0 in year 0 to the forecasted price

$$\widehat{p_t} = \hat{a}_{0,t} + (1 + \hat{b}_{0,t})p_0,$$

in year t, where the coefficients $\hat{a}_{0,t}$ and $\hat{b}_{0,t}$ are as in (2). From Table 1 we have that $1+\hat{b}_{0,t} > 0$ for all $t > 0$ and the mapping from year 0 quality p_0 to year t price is monotonic. Since

[2] Table 1A in Flavin and Yamashita (2001) reports a 14% return volatility for individual houses. Their Table 1B reports a 7% volatility for the Case-Shiller city index for San Francisco, which is comparable to San Diego. The difference between these two numbers is a 7% idiosyncratic volatility.

[3] Foreclosures (which are not included in our sample of repeat sales) are associated with 27% price discounts (Campbell, Giglio, and Pathak 2010.) Including these observations into our sample would thus further increase the estimated idiosyncratic volatility in the bust period.

price again reflects quality one for one in year t, we have that common shocks do not upset the relative ranking of house qualities. Of course, the quality ranking of individual houses may change because of idiosyncratic shocks – for example, some houses may depreciate more than others. These shocks average to zero because of the law of large numbers.

We now turn to the quality distribution in year t. Let Φ_t denote the cdf of log transaction prices in year $t > 0$. We know that the average house that starts at quality p_0 in year 0 trades at the price $\hat{p}_t$ in year t. We define the fraction of houses of quality lower than p_0 as

$$G_t(p_0) = \Phi_t\left(\hat{a}_{0,t} + (1 + \hat{b}_{0,t})p_0\right).$$

By this definition, the index p_0 tracks relative quality across years. If the same set of houses trades in both years 0 and t, then the quality distributions G_t and G_0 are identical. More generally, G_t can be different from G_0 because different sets of houses trades at the two dates. For example, if more higher quality houses are built and sold in t then G_t will have more mass at the high end.

We emphasize that the quality index p_0 tracks only relative quality – it does not provide a measure of absolute quality (or service flow, which is the flow utility derived from owning a house) that is comparable across years. For example, it could be that the housing service flow grows between 0 and t, possibly at different rates depending on the initial quality level. If the same set of houses trades in both years, it would still be true that the distributions G_t and G_0 are identical. Conversely, exit and entry of houses could make G_t different from G_0 even if the absolute quality of all houses remains unchanged over time. For example, all houses close to the median of the cross sectional distribution of qualities at date 0 might be well below the median at date t because higher quality houses are built between dates 0 and t, and also lower quality houses are torn down. The distinction between quality changes and revaluation will be taken up below in the context of our structural model.

Figure 2 shows the cumulative distribution functions G for the base year 2000 as well as for $t = 2005$. The cross sectional distribution of prices Φ are taken from Census and ACS data, respectively. The key difference between the two quality distributions is that there was more mass in the tails in the year 2005 (red line) than in 2000 (blue line.) In other words, the year 2005 saw more transactions of low and high quality homes compared to the year 2000. As a robustness check, we have also considered cdfs based directly on county deeds records, with similar results.

Improvements

A candidate explanation for the negative slope in Figure 1 are improvements. To see whether cheap homes experience more improvements than expensive homes, we use data from the American Housing Survey, which collects data on both improvements and routine maintenance expenses. The houses in California differ from the national average, because their value consists to a large part of land. Therefore, we use data from the 2002 AHS, which contains observations from San Diego County.

Table 2 shows improvements as a fraction of house value across the spectrum of house values (reported in 2005 Dollars, as in Figure 1), and the ratio of regular maintenance expenses as a fraction of house values. Except for houses worth less than \$50,000, the annual

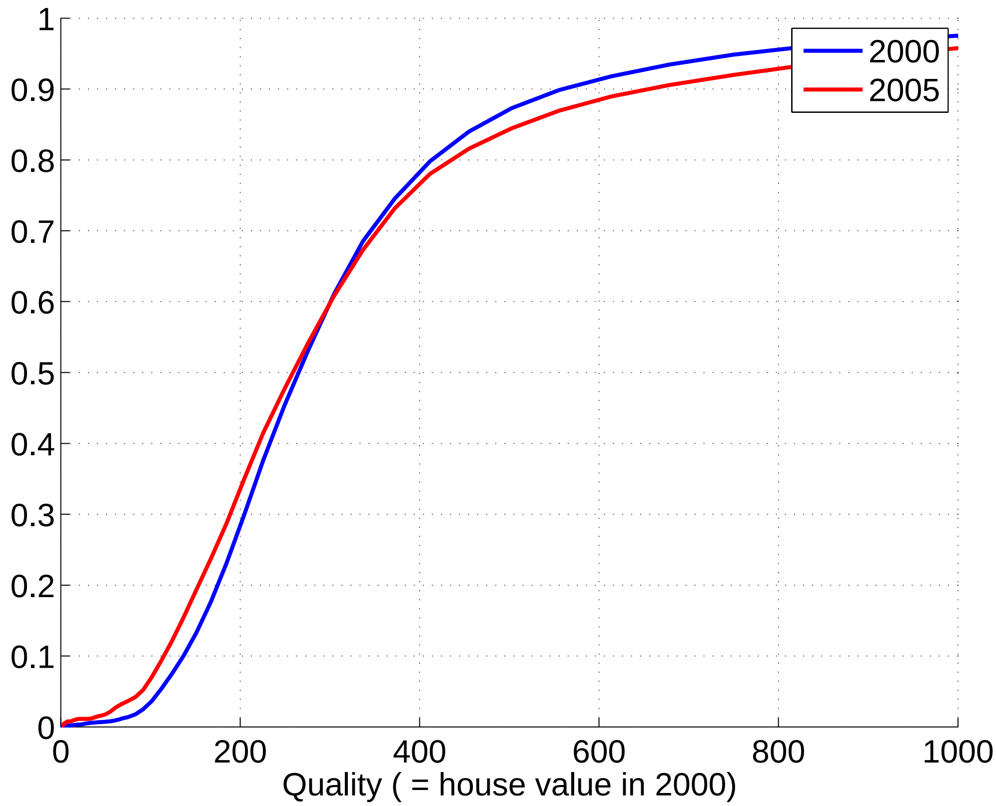

Figure 2: Cummulative distribution function of house qualities in 2000 and 2005

value of improvements is roughly a constant 1% of the house value. For houses worth less than \$50,000, this fraction is 3%. Alternatively, we run regressions of the log improvement fraction on log house values, and find a zero coefficient when we exclude houses worth less than \$50k and a small negative coefficient when we include these houses. From Figure 2, we can see that San Diego does not have many houses that are worth less than \$50,000; they represent 2.5% of the quality distribution in the year 2000. This evidence thus suggests that improvements cannot account for the higher capital gains between the years 2000 and 2005 on houses worth, say, \$100,000 in 2000.

San Diego homeowners spend about 0.3% of their house value on regular maintenance each year. This fraction is about a third lower than estimates for the national average, which are around 1%. In the national average, land is about a third of the house value, while in California, land is about two thirds of the house value. These numbers suggest that San Diego homeowners spend a similar amount on regular maintenance on their structure. Owners of houses that are worth less than \$50,000 spend more, almost 2% of the value of their house, on maintenance. Again, these differences in maintenance expenses across housing segments cannot explain the strong appreciation of cheaper homes (that are not below \$50,000.)

Table 2: Home Improvements and Regular Maintenance in San Diego

	House Value (in thousands, 2005 Dollars)							
	<50	50-100	100-150	150-200	200-300	300-500	500-1,000	>1,000
Home improvements per year (as a fraction of house value)								
mean	0.030	0.011	0.009	0.007	0.010	0.008	0.012	0.023
	(0.012)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.004)
Regular maintenance per year (as a fraction of house value)								
mean	0.017	0.008	0.005	0.003	0.003	0.002	0.002	0.001
	(0.003)	(0.002)	(0.001)	(0.001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)

Note: This table contains the estimated means and their standard errors (in brackets) of two variables: (i) the ratio of home improvement or replacement expenses over the house value, (ii) the ratio of annual routine maintenance costs over the house value. These statistics are computed for observations within the house price bins indicated on the top of the table. The data are the San Diego County observations of the 2002 American Housing Survey on 'rac' which measures the cost of replacements/additions to the unit and 'cstmnt' which measures routine maintenance last year. The 'rac' amount is divided by two, because the survey asks about expenses within the last two years.

2.3 Mover characteristics

We study the decisions by movers, and so we measure the characteristics of movers in San Diego in the years 2000 and 2005. Table 3 shows summary statistics regarding the three dimensions of household heterogeneity in our model: age, income, and wealth. Table 3 also shows the value of the house owned by these households.

The average mover is quite different from other households in San Diego, and so we also report the summary statistics for stayer households as comparison. The difference in mover versus stayer characteristics is particularly large in the year 2005, at the peak of the housing boom. These comparisons stress the importance of measuring the characteristics of movers, which are the households whose optimality conditions we want to evaluate.

First, Table 3 shows that movers tend to be younger than the average household. In San Diego, roughly 13% of household heads are below 35 years old, and 34% of movers in 2000 are in this age group. In 2005, the group of movers in this age group is even larger, 46%. Few old households move. Roughly 22% of households in San Diego are above 65 years old, and the percentage of movers in this age group is roughly 7%.

Young households have lower income and wealth than older households. The average annual income of households aged below 35 years is $89,600 in the year 2000 and $91,200 in the year 2005. (All these dollar numbers are reported in 2005 Dollars, and are thus comparable.) For example, the average annual income of households aged 50-65 is roughly $30k higher in 2000 and $65k higher in 2005. The wealth of young households in the year 2000 is $359,500 in 2000 and $443,300 in the year 2005. Older households have three times as much wealth in 2000, and five times as much wealth in 2005.

Young households then buy cheaper homes. The average value of a house bought by a household aged below 35 years is $272.200 in 2000 and almost twice that amount, $499,800, in 2005. Older households buy homes that are on average worth $391,500 in 2000 and twice that amount at the peak of the boom.

From Table 3, we can see that movers buy more expensive homes than what average San Diego households own in their age group. For example, in the year 2000, the average house value of 50-65 year old households is worth $339,400, while movers in this age group buy a house worth $391,500 on average. At the peak of the housing boom, the difference between the value of houses purchased by movers ($766,100) and average house owned by households in this age group ($657,400) is even larger for this age group.

Table 3: Characteristics of San Diego Movers and Stayers

	year 2000			year 2005		
	all	movers	top 10%	all	movers	top 10%
1. Fraction of households in each age group						
below 35	0.13	0.34		0.14	0.46	
35-50 years	0.37	0.42		0.36	0.34	
50-65 years	0.27	0.17		0.29	0.14	
above 65	0.23	0.08		0.22	0.07	
2. House value (in thousands, 2005 dollars)						
below 35	268.8	272.2	0.17	539.0	499.8	0.14
35-50 years	319.5	346.6	0.18	641.3	652.0	0.14
50-65 years	339.4	391.5	0.20	657.4	766.1	0.14
above 65	312.2	340.5	0.21	582.4	648.4	0.16
3. Income (in thousands, 2005 dollars)						
below 35	91.6	89.6	0.17	98.5	91.2	0.15
35-50 years	106.3	108.9	0.18	112.5	115.7	0.18
50-65 years	104.9	117.9	0.19	111.3	156.5	0.18
above 65	69.1	81.5	0.19	67.0	84.3	0.22
4. Wealth (in thousands, 2005 dollars)						
below 35	379.2	359.5	0.51	537.8	443.3	0.48
35-50 years	605.1	668.8	0.48	931.3	871.2	0.45
50-65 years	901.6	967.9	0.46	1,681.7	2,291.9	0.42
above 65	971.4	1,009.9	0.43	1,616.7	1,607.5	0.41

Note: This table reports summary statistics for average households ("all") and mover households ("movers") in San Diego County for the years 2000 and 2005. The table has four age bins for household heads: below 35 years, 35-50 years, 50-65 years, and above 65 years. The table breaks out the top 10% richest households and reports the fraction of the total value owned by the top 10%. For age, income, and house values we use age of the household head, income and house value reported in the 2000 Census and 2005 ACS. For wealth, we use imputed wealth with data from the Survey of Consumer Finances. The appendix explains the details.

Table 3 also shows that movers are better off than stayer households. They have higher incomes than the average household in San Diego, except for the youngest age group. Movers also have more wealth. In the year 2005, these differences were even more pronounced.

These averages mask substantial heterogeneity within each characteristic. For example, the top 10% richest households tend to make roughly 20% of the total income earned by their age group, which illustrates the fact that there is income inequality. The amount of inequality is within one percentage point among movers and stayers, and so we only report the average percentage of overall income that is attributable to the richest households. The top 10% households also own 20% of the housing wealth in their age group. This inequality is even more pronounced for wealth, where the top 10% households own 50% of the total wealth in their age group. The amount of inequality within these categories is roughly the same in 2000 and 2005.

3 Assigning houses to movers

We consider an assignment model of a city. A group of mover households faces an inventory of available houses. Houses are indivisible and come in different qualities indexed by $h \in [0,1]$. The one-dimensional quality index h summarizes various aspects of housing that households care about (for example square footage, location, views, other amenities such as schools.) The inventory of houses is described by a strictly increasing cumulative distribution function $G(h)$. A house of quality h trades in a competitive market at the price $p(h)$.

Every mover household buys exactly one house. Let $h^*(p,i)$ denote the housing demand function of household i. It depends on the house price function p as well as on household i's characteristics. In equilibrium, the markets for all house types clear. For every $h \in [0,1]$, the number of households who demand houses of quality less than h must therefore be equal to the number of such houses in the inventory:

$$\Pr\left(h^*(p,i) \leq h\right) = G(h)\,. \tag{3}$$

The price function $p(h)$ describes a set of house prices at which households are happy to be assigned to the available inventory of houses.

To illustrate the properties of this setup and compare it to other models of house prices, we consider a simple version. We assume that (i) housing demand h^* is derived from a frictionless, deterministic, two-period optimization problem and (ii) households differ by a single characteristic i, lifetime wealth. Households care about two goods: housing and other (numeraire) consumption. Below, all these assumptions will be relaxed in our quantitative model.[4]

The advantage of the simple version is that we can derive closed form expressions for equilibrium house prices. The quantitative version of the model has to be solved numerically, so there will not be formulas that help with the intuition. In the simple version, we can

[4]Below household demand will be explicitly derived from an intertemporal optimization program under uncertainty and with frictions (borrowing constraints and transaction costs.) Households will differ by multiple characteristics, including age, wealth, and income.

analyze the impact of two types of changes, which will also be important in the quantitative model later: changes in the distribution G and shocks to certain subpopulations (e.g., an increase in housing demand by poor households.)

At date 0, households receive income y and buy a house of quality h at the price $p(h)$. A house of quality h yields housing services $s(h)$, where s is strictly increasing with $s(0) = 0$. Households also choose their consumption of numeraire c as well as an amount b borrowed at the gross riskless interest rate R (where negative b corresponds to the purchase of bonds). Date 1 variables are indicated by a tilde. At date 1, households receive income $\tilde{y}$, sell their house at a price $\tilde{p}(h)$, and pay or receive funds in the credit market so they are left with cash $\tilde{w}$.

A household maximizes utility

$$u(c, s(h)) + v(\tilde{w}) \tag{4}$$

subject to the budget constraints for dates 0 and 1

$$c + p(h) = y + b,$$
$$\tilde{w} = \tilde{y} + \tilde{p}(h) - Rb.$$

Define lifetime wealth by $w = y + \tilde{y}/R$. The household problem can be rewritten as choosing c, h and $\tilde{w}$ to maximize (4) subject to the single lifetime budget constraint

$$c + \tilde{w}/R + \underbrace{p(h) - \tilde{p}(h)/R}_{\text{user cost } \rho(h)} = w,$$

where the expression in braces is the user cost $\rho(h)$ of a house of quality h: the cost of purchasing the house less its discounted resale value.

Let $F(w)$ denote the strictly increasing cumulative distribution function of lifetime wealth w defined on the nonnegative real line. In this stylized version, wealth is the only dimension of household heterogeneity. An equilibrium consists of a consumption and house allocation together with a price function such that households optimize and markets clear. From the household problem, the price only matters to the extent that it affects the user cost ρ. It is thus natural to define equilibrium directly in terms of the user cost. We return to the relationship between price and user cost below.

The first order conditions for the household problem include

$$\begin{aligned} u_1(c, s(h)) &= Rv'(\tilde{w}) \\ \rho'(h) &= \frac{u_2(c, s(h))\, s'(h)}{u_1(c, s(h))} \end{aligned} \tag{5}$$

The first condition is standard: it equates the intertemporal marginal rate of substitution to the interest rate. The second condition says that the intratemporal marginal rate of substitution (MRS) between housing and numeraire consumption equals the *marginal* user cost of housing $\rho'(h)$ at the quality level h that the household chooses. Intratemporal Euler equations that equate MRS and user cost hold in many models of housing. What is special here is that the user cost does not need to be linear in quality. The MRS is thus equated to a marginal user cost that may differ across quality levels. In this sense, houses of different quality are priced by different marginal investors.

Consider an equilibrium such that optimal house quality is unique and strictly increasing in lifetime wealth. The assignment of houses to wealth levels is then given by a strictly increasing function $h^* : \mathbb{R}_0^+ \to [0,1]$. It is convenient to work with its inverse $w^*(h)$, which gives the wealth level of an agent who is assigned a house of quality h. The market clearing condition (3) now says that, for all h,

$$F(w^*(h)) = G(h) \Longrightarrow w^*(h) = F^{-1}(G(h)),$$

which works because F is strictly increasing. The assignment of wealth levels to house qualities depends only on the respective distributions, and is independent of preferences. Of course, prices will depend on preferences through the Euler equations.

The function w^* describes a QQ plot commonly used to compare probability distributions. Its graph is a curve in (h,w)-space that is parametrized by the common cdf value in $[0,1]$. The shape of the graph w^* is determined by the *relative* dispersion of house quality and wealth. If the relative dispersion is similar, the graph of w^* is close to the 45 degree line. In the case where the quality and wealth distributions are identical, the graph of w^* is exactly the 45 degree line. If the distribution of wealth F is more dispersed than the quality distribution G, the graph of w^* is steeper than the 45 degree line. Otherwise, w^* is flatter. For example, if all houses are essentially of the same quality, but there is some dispersion in wealth, then w^* must be close to a vertical line and is thus very steep.

To characterize equilibrium prices in closed form, we specialize further and assume a linear services production function $s(h) = h$ as well as separable log utility, that is, $u(c) = \log c + \theta \log s$ and $v(\tilde{w}) = \beta \log \tilde{w}$. From the Euler equations, the marginal user cost at quality h must be equal to the MRS between housing and wealth spent on other goods:

$$\rho'(h) = \frac{\theta}{1+\beta} \frac{w^*(h) - \rho(h)}{h}. \tag{6}$$

An agent with wealth $w^*(h)$ must be indifferent between buying a house of quality h and spending $w^*(h) - p(h)$ on other goods, or instead buying a slightly larger house and spending slightly less on other goods. An agent who already spends a lot on other goods is willing to pay more for a larger house (because of diminishing marginal utility of nonhousing consumption.) Therefore, if the house of quality h is assigned to an agent who spends more on nonhousing consumption per unit of house quality, then the slope of the user cost function must be steeper at the point h.

To obtain closed form solutions for equilibrium prices, we further assume that the distributions G and F are such that the assignment function is a polynomial

$$w^*(h) = \sum_{i=1}^{n} a_i h^i. \tag{7}$$

The lowest-quality house must have a zero user cost since it is purchased by the buyer who has zero wealth. The unique solution to the ordinary differential equation (6) that satisfies $\rho(0) = 0$ is given by

$$\rho(h) = \int_0^h \left(\frac{\tilde{h}}{h}\right)^\theta \frac{\theta w^*\left(\tilde{h}\right)}{\tilde{h}} \, d\tilde{h} = \sum_{i=1}^n a_i \frac{\theta}{\theta + (1+\beta)\, i} h^i. \tag{8}$$

If this solution is strictly increasing, then it is an equilibrium user cost function. The user cost for a house of quality h is the weighted average of MRS for all agents who buy quality *less* than h, with the MRS evaluated at total wealth.

Comparison to a model with divisible housing

It is interesting to compare the above model with indivisible housing to a standard model with divisible housing capital. The indivisible model takes as given cdfs $G(h)$ and $F(w)$ for houses and movers, respectively. It determines endogenously an assignment $w^*(h)$ and a user cost function $\rho(h)$. In contrast, a divisible model takes as given a distribution F and a total housing stock $\overline{H} = \int h \, dG(h)$. It also assumes a linear production technology that converts houses of different quality into each other.

Consider an equilibrium of the divisible model such that housing is monotonic in wealth. The divisible model endogenously determines an allocation that can also be represented by an increasing function $w^*(h)$, together with the slope of a linear user cost function $\rho(h) = \bar{\rho} h$ and an equilibrium distribution of house qualities, $\hat{G}(h)$ say. From market clearing, the *endogenous* distribution $\hat{G}$ must agree with G in the mean. All other changes from G to $\hat{G}$ can be achieved by costless transformation of houses.

In the divisible model, the per-unit user cost $\bar{\rho}$ enters the Euler equations of *all* households:

$$\bar{\rho} = \frac{u_2(c, s(h))}{u_1(c, s(h))} \tag{9}$$

The marginal (or, equivalently, the average, per unit) user cost can be read off the Euler equation of any household – in this sense *every* household is a marginal investor for *every* house. In contrast, in the indivisible model the marginal user cost at quality h can be read off only *one* Euler equation (5), that of the marginal investor with wealth $w^*(h)$.

The segment-specific pricing of houses implies that equilibrium prices under the indivisible model depend more on details of the distribution than under the divisible model. To see this, consider the case of separable log utility studied earlier. The average user cost becomes

$$\bar{\rho} = \frac{\theta}{\theta + 1 + \beta} \frac{E[w]}{\bar{H}} \tag{10}$$

It depends on the distributions F and G only through their respective means. In contrast, the user cost in the indivisible model (8) typically depends on details of the distributions through the parameters of the assignment (7). The two models produce identical results only in special cases. For example, suppose the distributions F and G are scaled version of each other. If $F(w) = G(w/k)$, then $w^*(h) = kh$ and $k = E[w]/\bar{H}$, so the price function for both models is given by (10).

User cost, expectations, and prices

So far we have only solved for the user cost. House prices depend on the equilibrium user cost as well as the households' price expectations. In this simple version of the model, we distinguish two scenarios for expectation formation. The first assumes that agents expect all prices to grow at a common gross rate μ, starting from the current equilibrium price function. Agents thus extrapolate forward the behavior of relative prices from what they currently observe.

Setting $\tilde{p}(h) = \mu p(h)$ in the definition of the user cost, it follows that the equilibrium price is given by

$$p(h) = \frac{\rho(h)}{1 - \mu/R}$$

In a frictionless model with these constant capital gain expectations, house prices are proportional to user costs, and log price changes are equal to log changes in user costs. This scenario is useful for analyzing an economy in normal times when households perceive it to be in steady state. It is also useful when analyzing a boom in which households believe changes in the price pattern to be permanent.

The second scenario assumes an exogenous price expectation function $\tilde{p}(h)$. It is of interest when thinking about a boom in which households expect prices to mean revert to some earlier price pattern, given by $\tilde{p}$. The house price is then

$$p(h) = \tilde{p}(h)/R + \sum_{i=1}^{n} a_i \frac{\theta}{\theta + (1+\beta)\, i} h^i$$

Under this scenario with mean-reverting capital gains expectations, relative house prices still depend on the distributions F and G via relative user costs, but relative price expectations now also play a role. The relative importance of the distribution term for house prices is increasing in the subjective and market discount rates β^{-1} and R. Intuitively, the current assignment matters more for prices if the holding period for houses is longer.[5]

Simple graphical example

Figure 3 compares equilibria of the indivisible and divisible model with separable log utility. The top left panel shows a lognormal wealth density $F'(w) = f(w)$. The top right panel shows a uniform house density $G'(h) = g(h)$. In both panels, the second and fourth quintile have been shaded for easier comparison. The bottom left panel shows prices. The solid line is the price function for the indivisible model. The dotted line is the price function for the divisible model. As shown above, it may equivalently be interpreted as the price function in an indivisible model with a uniform wealth distribution or a lognormal house quality distribution. Finally, the bottom right panel compares the price to the QQ plot $w^*(h)$, which provides the wealth level of an agent who buys a house of quality h.

The user cost function of the indivisible model is nonlinear. This reflects differences in the shapes of F and G that lead to a nonlinear assignment w^*. The bottom right panel shows

[5] In the simple model considered here, changing the discount rate corresponds to making the holding period exogenously longer for all agents. In the more general model below, transaction costs induce agents to endogenously choose long holding periods.

that the function $w^*(h) = F^{-1}(G(h))$ is relatively steep for low and high house qualities, and relatively flat in between. The shape of w^* is determined by the relative dispersion within quintiles in the top two panels. For the uniform distribution of house qualities, dispersion is the same within each quintile. In contrast, the dispersion of wealth is relatively high in the first and fifth quintile, but relatively low in the second and third quintile. To achieve an assignment with house quality increasing in wealth, wealth must thus rise more with house quality in the former quintiles than in the latter.

The price function is determined by the segment-specific Euler equation (6). For a given quality h, the user cost function is steeper if the marginal investor spends more on nonhousing consumption per unit of house quality. In particular, for a given $\rho(h)$, the user cost must rise more if the marginal investor is richer. The user cost function thus inherits from the assignment w^* the property that it steepens for very high and very low qualities. There is less steepening at high qualities where wealth per unit of quality responds less to w^*. In addition, the user cost function must be consistent with pricing by all marginal investors at qualities less than h. More dispersion in wealth relative to house quality in lower quintiles thus leads to higher prices in higher quintiles.

Comparative statics

To investigate how equilibrium house prices respond to various changes in distributions, we assume for simplicity that households have constant capital gains expectations. With these expectations, house prices are equal to the user costs in the lower bottom panel (up to the factor $1/(1-\mu/R)$.)

Figure 4 shows what happens if the house quality distribution becomes fatter tailed; it has more mass at both the low and high end. The new density – the green line in the top right panel – comes from a beta distribution with mean one half. The mean house quality is thus unchanged from the uniform distribution. Therefore, the equilibrium price in the divisible model is the same as before. In contrast, prices change in the indivisible model to reflect the change in distribution. The green line in the bottom left panel is the new price function.

The bottom right panel shows capital gains by house value implied by the change in distributions: it plots the log house value in the "blue economy" (the log of the blue line in the bottom left panel) on the horizontal axis against the capital gain from blue to green (the log difference between the green and blue lines in the bottom left panel) on the vertical axis. The main result here is that capital gains are much higher at the low end than at the high end.

To understand the intuition, it is again helpful to consider the relative dispersion of wealth and quality within quintiles. In the bottom quintile, the dispersion of house qualities has now decreased, making wealth even more dispersed relative to quality. The assignment w^* must become steeper in this region as richer agents must buy lower quality houses. A steeper assignment in turn implies a steeper price function. Starting from the smallest house, prices rise faster to keep richer marginal investors indifferent. For higher qualities, for example in the third quintile, the effect is reversed: as the house distribution is more dispersed than wealth distribution, poorer marginal investors imply a flatter price function.

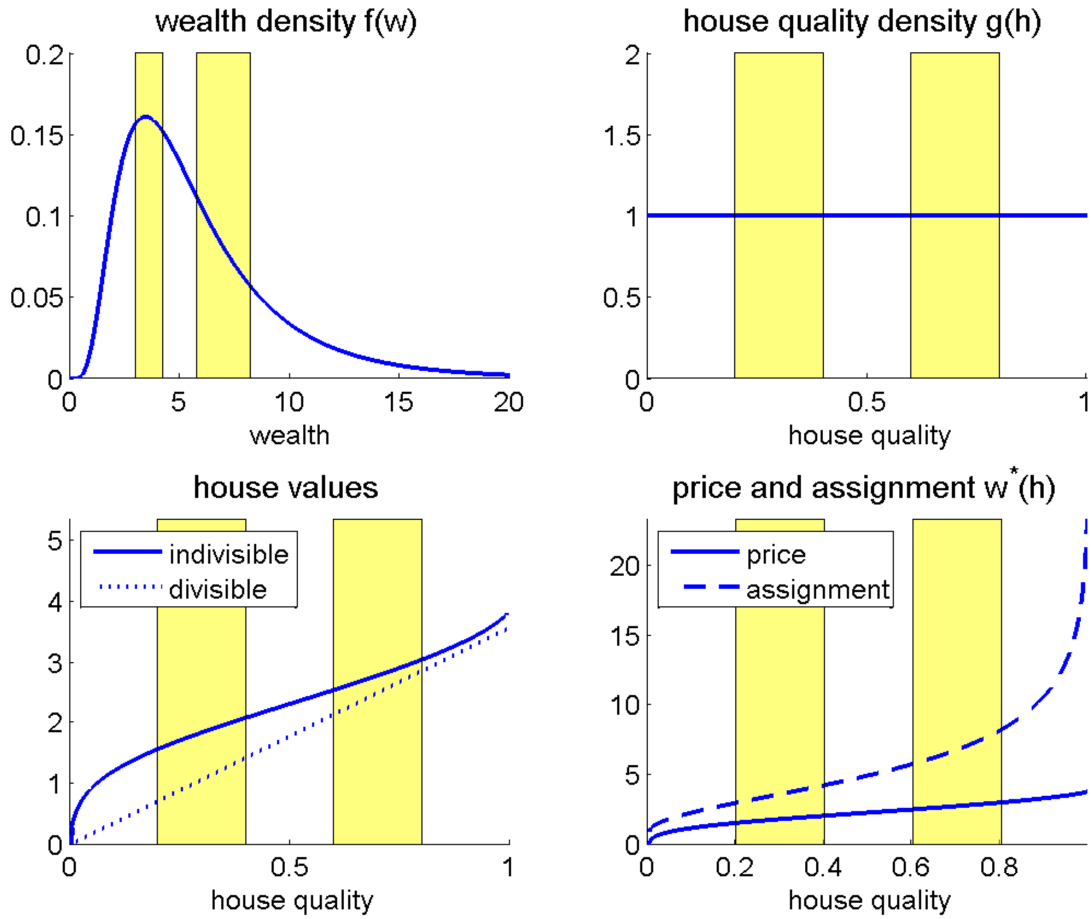

Figure 3: Equilibrium prices and assignment with lognormal wealth density $f(w)$ and uniform house quality density $g(h)$. Top left: wealth density $f(w)$. Top right: house quality density $g(h)$. Bottom left: the equilibrium price function in the indivisible model (solid line) and the divisible model (dotted line). Bottom left: assignment $w^*(h)$ and prices. Shaded areas indicate the quintiles of the distribution.

Figure 5 provides an example of a shock to a subpopulation. We assume that all agents with wealth less than 4 develop a higher taste for housing, as measured by the parameter θ. Over that range, we choose the increase in θ to be linearly declining in wealth, with a slope small enough such that the assignment is still monotonic in wealth. The wealth and quality distributions are the same as before. The bottom panels compare the price and capital gain effects. In the divisible model, the price rises to reflect higher demand for housing. In the indivisible model, the Euler equation predicts that the slope of the price function becomes steeper for low house qualities (where the poor households buy.)

Not surprisingly, higher demand leads to higher prices in both the indivisible and divisible models. The divisible model predicts that capital gains are the same for all qualities. Interestingly, the indivisible model implies a higher capital gain at the low end, reflecting the higher demand of households who buy low quality houses. At the high end, the capital

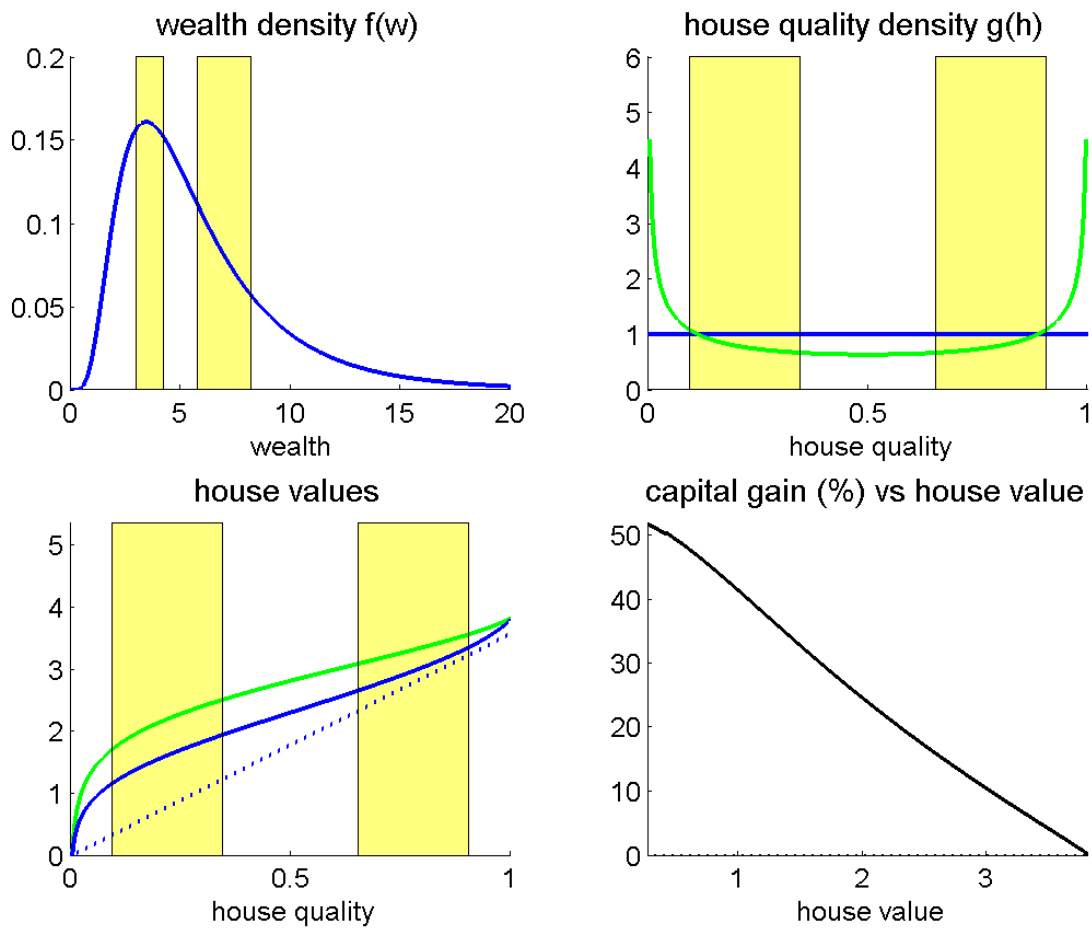

Figure 4: Changing the distribution of house qualities. Top left: wealth density $f(w)$. Top right: uniform quality density $g(h)$ in blue, beta quality density in green. Bottom left: blue equilibrium price function for unform distribution, green function for beta quality density. Bottom right: capital gain from blue prices to green prices.

gain in the indivisible model is actually lower than under the divisible model.

4 A Quantitative Model

For the stylized model in the previous section, housing demand was derived from a frictionless, deterministic, two-period optimization problem and households differed only in wealth. In this section, we describe a more general intertemporal problem for household savings and portfolio choice. This problem accommodates many features that have been found important in existing studies with micro data, and thus lends itself better to quantitative analysis. It differs from most existing models because there is a continuum of (indivisible) assets that agents can invest in.

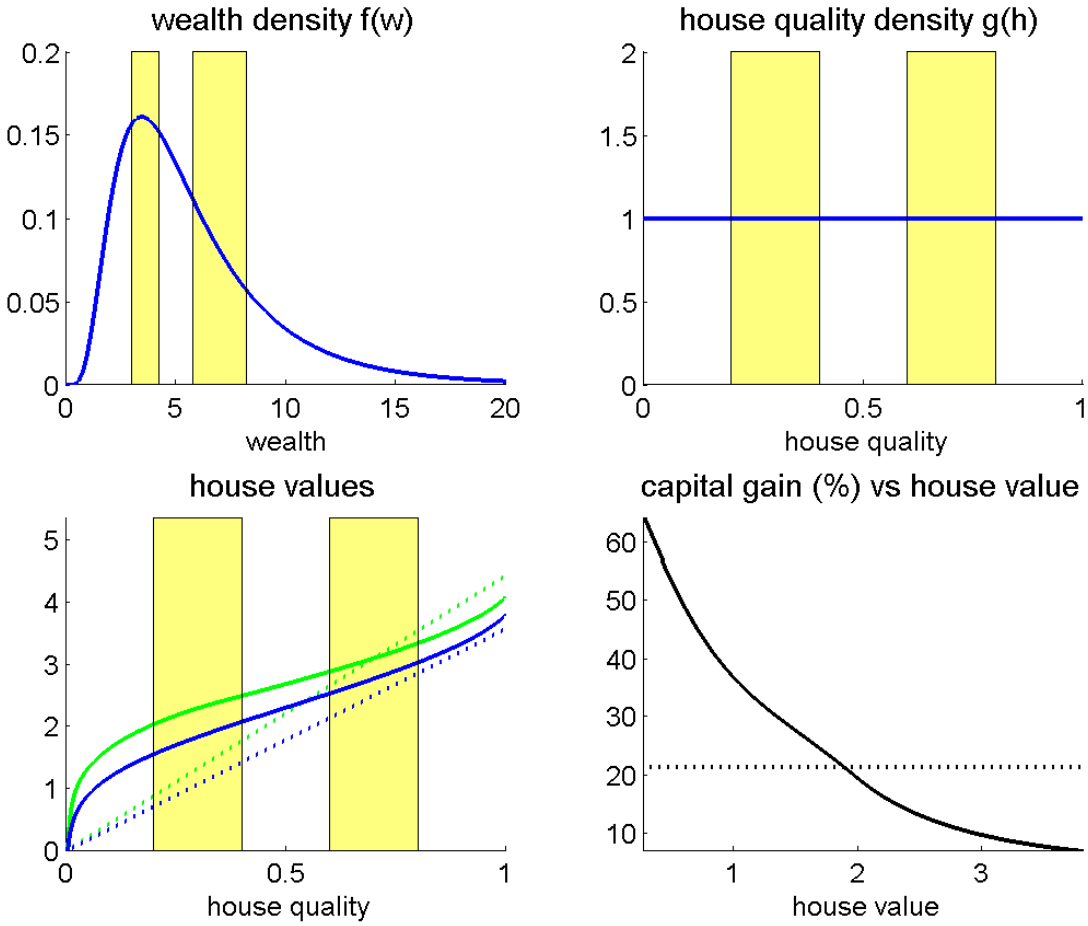

Figure 5: Changing poor households' preferences for housing.

The problem is solved for a distribution of households that differ in age, income, and cash on hand (that is, liquid resources). The distribution is then chosen to capture the set of movers in San Diego County in a given year. An equilibrium is then defined by equating the distribution of movers' housing demand derived from the dynamic problem to the distribution of transacted houses.

4.1 Setup

Households live for at most T periods and die at random. Let D_t denote a death indicator that equals one if the household dies in period t or earlier. Preferences are defined over streams of housing services s and other (numeraire) consumption c during lifetime, as well as the amount of numeraire consumption w left as bequest in the period of death. Conditional on period τ, utility for an agent aged a_τ in period τ is

$$E_\tau \left[\sum_{t=\tau}^{\tau+T-a_\tau} \beta^t \left[(1-D_t) \ u(c_t, s_t(h_t)) + (D_t - D_{t-1}) \ v(w_t) \right] \right] \tag{11}$$

Households have access to two types of assets. First they can buy houses of different qualities $h \in [0,1]$ that trade at prices $p_t(h)$. Owning a house is the only way to obtain housing services for consumption. A house of size h_t owned at the end of period t produces a period t service flow $s_t(h_t)$ where the function s_t is strictly increasing. It may depend on time to accommodate growth.

Households borrow b_t at the gross interest rate R_t between period t and $t+1$. The amount b_t measures *net* borrowing. A negative position b_t corresponds to bond purchases. We assume that a household can only borrow up to a fraction $1-\delta$ of the value of his house. In other words, the amount b_t must satisfy

$$b_t \leq (1-\delta)p_t(h_t). \tag{12}$$

The fraction δ is the downpayment requirement on a house.

We introduce three further features that distinguish housing from bonds. First, selling houses is costly: the seller pays a transaction cost that is proportional to the value the house. Second, every period an owner pays a maintenance cost ψ, also proportional to the value of the house. Finally, a household can be hit by a moving shock $m_t \in \{0,1\}$, where $m_t = 1$ means that they must sell their current house. Formally, the moving shock may be thought of as a shock to the housing services production function $s_t(\cdot)$ that permanently leads to zero production unless a new house is bought.

Households receive stochastic income

$$y_t = f(a_t)\, y_t^p y_t^{tr} \tag{13}$$

every period, where $f(a_t)$ is a deterministic age profile, y_t^p is a permanent stochastic component, and y_t^{tr} is a transitory component.

Our approach to incorporating the tax system is simple. We assume that income is taxed at a rate τ. So the aftertax income $(1-\tau)\,y_t$ enters cash on hand and the budget constraint. Mortgage interest can be deducted at the same rate τ. Interest on bond holdings is also taxed at rate τ. Therefore, the aftertax interest rate $(1-\tau)\,R_t$ enters cash on hand and the budget constraint. We assume that housing capital gains are sheltered from tax.

To write the budget constraint, it is helpful to define cash on hand net of transaction costs–the resources available if the household sells:

$$w_t = p_t(h_{t-1})(1-\nu) + (1-\tau)\,y_t - (1-\tau)\,R_t b_{t-1} \tag{14}$$

The budget constraint is then

$$c_t + (1+\psi)p_t(h_t) = w_t + 1_{[h_t=h_{t-1}\&M_t=0]}\nu p_t(h_{t-1}) + b_t \tag{15}$$

Households can spend resources on numeraire consumption and houses, which also need to be maintained. If a household does not change houses, resources are larger than w_t since the households does not pay a transaction cost. The household can also borrow additional resources.

Consider a population of movers at date t. A mover comes into the period with cash w_t, including perhaps the proceeds from selling a previous home. Given his age a_t, current house prices p_t as well as stochastic processes for future income y_τ, future house prices p_τ, the interest rate R_τ, and the moving shock m_τ, the mover maximizes utility (11) subject to the budget and borrowing constraints. We assume that the only *individual-specific* variables needed to forecast the future are age and the permanent component of income y_t^p. The optimal housing demand at date t can then be written as $h_t^*(p_t; a_t, w_t, y_t^p)$.

As in the previous section, the distribution of available houses is summarized by a cdf $G_t(h)$. The distribution of movers is described by the joint distribution of the mover characteristics (a_t, w_t, y_t^p). An equilibrium for date t is a price function p_t and an assignment of movers to houses such that households optimize and market clear, that is, for all h,

$$\Pr\left(h_t^*(p_t; a_t, w_t, y_t^p) \leq h\right) \leq G_t(h).$$

4.2 Numbers

We now explain how we quantify the model. It is helpful to group the model inputs into four categories

1. Preferences and Technology

   (Parameters fixed throughout all experiments.)

   (a) Felicity u, bequest function v, discount factor β

   (b) conditional distributions of death and moving shocks

   (c) conditional distribution of income

   (d) service flow function (relative to trend)

   (e) maintenance costs ψ, transaction costs ν

2. Distributions of house qualities and mover characteristics

3. Credit market conditions

   (a) current and expected future values for the interest rate R

   (b) downpayment constraint δ

4. House price expectations

Our goal is to compare different candidate explanations for house price changes during the boom. We thus implement the model for two different trading periods: once before the boom, in the year 2000, and then again at the peak of the boom, the year 2005.

Preferences and technology are held fixed across trading periods. For each period, we measure the distribution of house qualities and mover characteristics from the 2000 Census cross section (this implementation is labeled $t = 2000$) and then again for the peak of the boom, using the 2005 ACS ($t = 2005$).

We measure credit market conditions in 2000, and assume that households in 2000 were expecting constant capital gains; households were expecting the 2000 relative prices to remain unchanged, and absolute prices to grow at a constant rate with income. The service flow function is chosen to match 2000 house prices at these expectations.

After solving the model for the year 2000, we compute the model for the year 2005 under different scenarios for credit market conditions and house price expectations. We compare the predictions for 2005 equilibrium house prices with 2005 data. Below, we describe all elements in more detail.

Preferences

Felicity is given by power utility over a Cobb-Douglas aggregator of housing services and other consumption:

$$u(c, s) = \frac{[c^{\rho}\, s^{1-\rho}]^{1-\gamma}}{1-\gamma}, \tag{16}$$

where ρ is the weight on housing services consumption, and γ governs the willingness to substitute consumption *bundles* across both time period and states of the world. We work with a Cobb-Douglas aggregator of the two goods, with $\rho = .2$. If divisible housing services are sold in a perfect rental market, the expenditure share on housing services should be constant at 20%. This magnitude is consistent with evidence on the cross section of renters' expenditure shares (see for example, Piazzesi, Schneider, and Tuzel 2007.) We also assume $\gamma = 5$, which implies an elasticity of substitution for consumption bundles across periods and states of 1/5.

The period length for the household problem is three years. Households enter the economy at age 22 and live at most 23 periods until age 91. Survival probabilities are taken from the 2004 Life Table (U.S. population) published by the National Center of Health Statistics. To define utility from bequests, we compute the utility from receiving an L-year-annuity if a house of average quality can be rented at a rent to price ratio of 7%, the long run average in the US. We select $L = 7.5$ to match to match the age-profile of housing expenditure for households over 65.

The moving shocks are computed based on two sources. First, we compute the fraction of households who move by age, which is about a third per year on average. The fraction is higher for younger households. To obtain the fraction of movers who move for exogenous reasons, we use the 2002 American Housing Survey which asks households in San Diego about their reasons for moving. A third of households provides reasons that are exogenous to our model (e.g., disaster loss (fire, flood etc.), married, widower, divorced or separated.)

We assume that improvement and maintenance expenses cover the depreciation of the house. The evidence in Table 2 suggests that these expenses ψ are roughly 1.3% of the house value per year, which is the value we use. The transaction costs ν are 6% of the value of the house, which corresponds to real estate fees in California.

Conditional distribution of income

We estimate the deterministic life-cycle component $f(a_t)$ in equation (13) from the income data by movers. The permanent component of income is a random walk with drift

$$y_t^p = \mu y_{t-1}^p \exp(\eta_t), \tag{17}$$

where μ is a constant growth factor and η_t has mean $-\sigma_\eta^2/2$. The transitory component y_t^{tr} of income is iid. The standard deviation of permanent shocks η_t is 6% and the standard deviation of the transitory component is 22% per year, consistent with Campbell and Cocco (2010).

Distribution of mover characteristics

The problem of an individual household depends on characteristics (a_t, y_t^p, w_t). For age and income, we use age of the household head and income reported in the 2000 Census (for $t = 2000$) and 2005 ACS (for $t = 2005$). The Census data does not contain wealth information. Therefore, we impute wealth using data from the Survey of Consumer Finances. The appendix contains the details of this procedure.

Credit market conditions

The interest rate R is set to 3% in 2000, which we measure from three-year interest rate data on TIPS (since our model period is three years.) In 2005, the three-year TIPS rate fell to 1%, so we use that value. Households expect future interest rates to stay at 3%, their 2000 value. For the downpayment constraint δ, we use 20% to describe conditions before the housing boom in 2000, and 10% for the peak of the boom in 2005.

House price expectations

We specify house price expectations for the years 2000 and 2005 separately. For the base year 2000, we assume that households expect constant capital gains. Specifically, households expect all house prices to grow at the same rate μ as labor income:

$$p_{t+1}(h) = p_t(h)\,\mu \exp(u_{t+1}(h)), \tag{18}$$

where the shock $u_{t+1}(h)$ has mean $-\sigma_u^2/2$. We set the volatility of $u_{t+1}(h)$ to 14%, to capture the volatility of idiosyncratic house price shocks estimated in Table 1 (which is roughly 9%) together and the volatility of city-wide shocks to house prices (estimated from a San Diego city house price index.) Under this specification, house price expectations are endogenous, since they depend on current equilibrium house prices $p_t(h)$, as in Grandmont (1977).

For the peak of the housing boom in 2005, we consider two specifications. First, we assume that households expect house prices to revert back down to their 2000 levels $p_0(h)$, up to a common growth rate:

$$p_{t+1}(h) = p_0(h)\,\mu \exp(u_{t+1}(h)). \tag{19}$$

In this assumption, house price expectations are exogenous. Since equilibrium prices for low quality houses are relatively high in the year 2005, households expect cheap houses to fall more in price. These expectations are thus consistent with the regression evidence in Table 1.

We also consider a second specification in which households expect house prices to stay high at their 2005 levels and go up further with the common growth rate. This assumption corresponds to equation (18), where again expectations are endogenous; they depend on equilibrium 2005 levels. The survey evidence suggests that 20% of households in 2005 were indeed expecting house prices to go up further (Piazzesi and Schneider, 2009.)

Service flow function

The housing services produced by a house of quality h grow at the same rate μ as income. Starting from an initial service flow function $s_t(h)$, households expect

$$s_{t+1}(h) = \mu s_t(h). \tag{20}$$

As discussed above, the initial service flow function s_0 is backed out so that the model exactly fits the 2000 price distribution.

4.3 Results

In this section we compare pricing by quality segment in our base year 2000 and at the peak of the boom in 2005. We describe a number of different experiments to examine the role of distributional shifts (movers or houses) and cheaper credit.

Prices and service flow in the base year 2000

The first step in our quantitative analysis considers the base year 2000. Here we take as given (*i*) the distribution of household characteristics and (*ii*) the distribution of prices. We then determine a service flow function $s_0(h)$ such that the model exactly fits prices. This exercise is performed under the assumption of constant capital gains expectations.

Figure 6 shows the resulting service flow function, which is strictly increasing and concave in the quality index. For the base year 2000, the quality index is simply the current price. It follows that the price of a house is convex in the amount of housing services the house provides. Some intuition can be obtained from our analysis of the stylized model in Section 3. In that model, a price function linear in service flow obtains if the distribution of service flow across houses is a scaled version of the wealth distribution. In contrast, if the dispersion of the wealth distribution relative to the service flow distribution is larger over a particular quantile range, say the top quintile, than elsewhere, then we would expect the price function to be steeper over that range.

This logic rationalizes the shape of the backed out service flow function. Indeed, the wealth distribution is more dispersed in the top quintile than the distribution of house values. If the service flow were, say, linear, then rich households would all try to buy the most expensive houses. For markets to clear, some rich households must be induced to choose cheaper houses. This requires a steep increase in the price per unit of service flow.

Change in house prices from 2000 to 2005

Figures 7 and 8 compare prices and capital gains relative to 2000, respectively, for different experiments. The black lines show what happened in the data, while the blue line shows our benchmark results for 2005. The benchmark describes our preferred case for how house

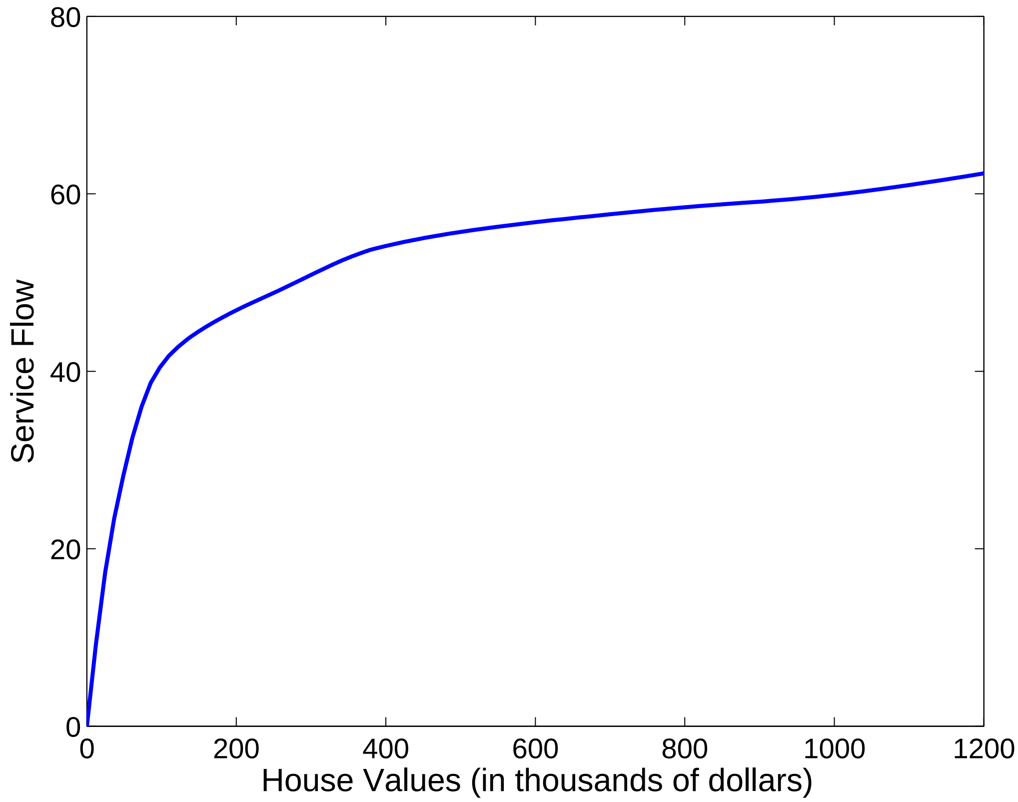

Figure 6: Service flow as a function of house quality (= 2000 house value)

prices repond to changes in the distributions of movers and house qualities, as well as to the credit market conditions in 2005. The exercise here is to assign the 2005 distribution of houses to the 2005 distribution of movers. Our preferred results are derived under house price expectations that are given by (18) with next period prices p_{t+1} equal to 2000 prices, up to a growth factor. In other words, whatever equilibrium prices are in 2005, agents expect relative prices to revert back to what they were *before* the boom. These expectations are consistent with the regression evidence in Table 1.

Low interest rates versus lower downpayment constraints

The green lines in Figures 7 and 8 compute the 2005 equilibrium prices under benchmark assumptions, except that we assume a high 2000 value for the interest rate. In other words, the green lines isolate the importance of lower downpayment constraints. A comparison of the green line and the blue line in Figure 8 shows that downpayment constraints are responsible for most of the quantitative effect of a change in credit market conditions on house prices between 2000 and 2005. Moving from the green line to the blue line adds the effect of low interest rates, which is smaller quantitatively.

Lower downpayment constraints relax borrowing constraints and increase the housing demand of mostly poor households. As a consequence, the house prices of low quality houses increase. Intuitively, the relaxation of borrowing constraints acts like the increase in housing demand by poor households in the simple model illustrated in Figure 5.

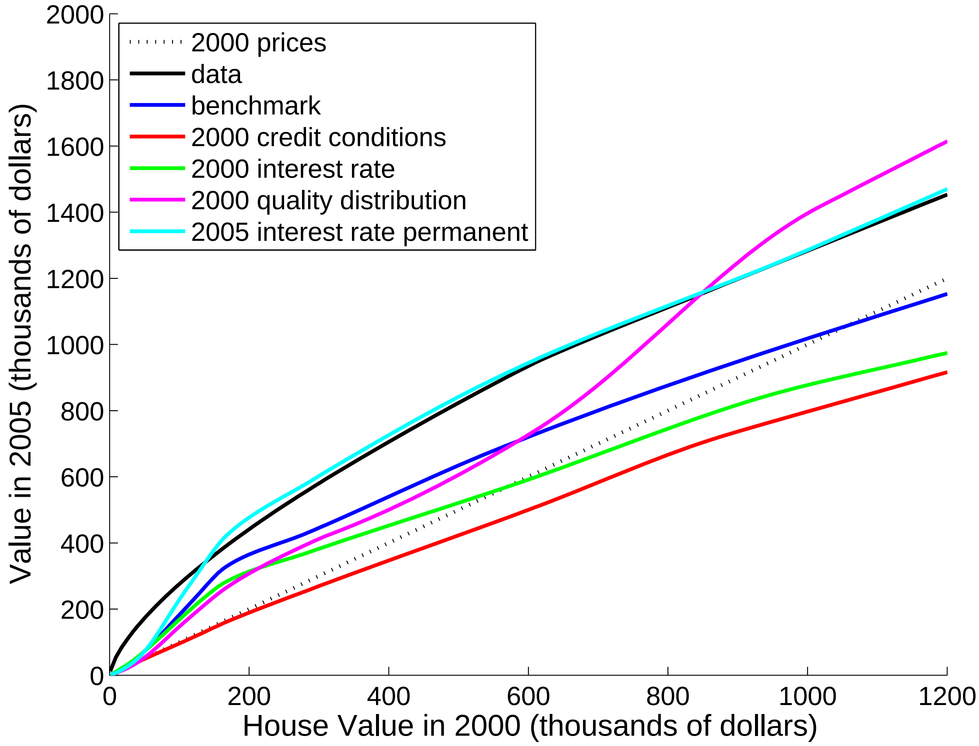

Figure 7: Price function for 2005, different experiments. The solid black line is the price function from the data. The blue line uses 2005 distributions of movers and house qualities as well as 2005 credit conditions. The other lines are computed on the same benchmark assumptions with the departure from the benchmark indicated in the legend. The red line uses 2000 credit conditions. The green line uses 2000 interest rates. The magneta line uses the 2000 quality distribution. The cyan line assumes that households expect interest rates to stay permanently low at 2005 levels.

Changes in mover and house quality distributions

The red lines in the figures compute the 2005 equilibrium prices under benchmark assumptions, except that we assume 2000 credit conditions. This experiment highlights that changes in credit conditions are critical for house prices. The changes in distributions of mover characteristics and house qualities, however, affect the relative pricing of houses across qualities. In particular, the red line in Figure 8 shows higher capital gains for low end homes than for high end homes.

The changes in relative house prices is at least in part a conquence of the fatter tails of the 2005 house quality distribution in Figure 2. Richer households move into low quality homes and push up their prices relative to high quality homes. The intuition for this change of the house quality distribution is thus the same as in Figure 4 of the simple model.

The magneta lines compute the model under benchmark assumptions, but with the 2000 house quality distribution. The associated equilibrium capital gains in Figure 8 show that changes in the quality distribution matter quantitatively for the relative change of house

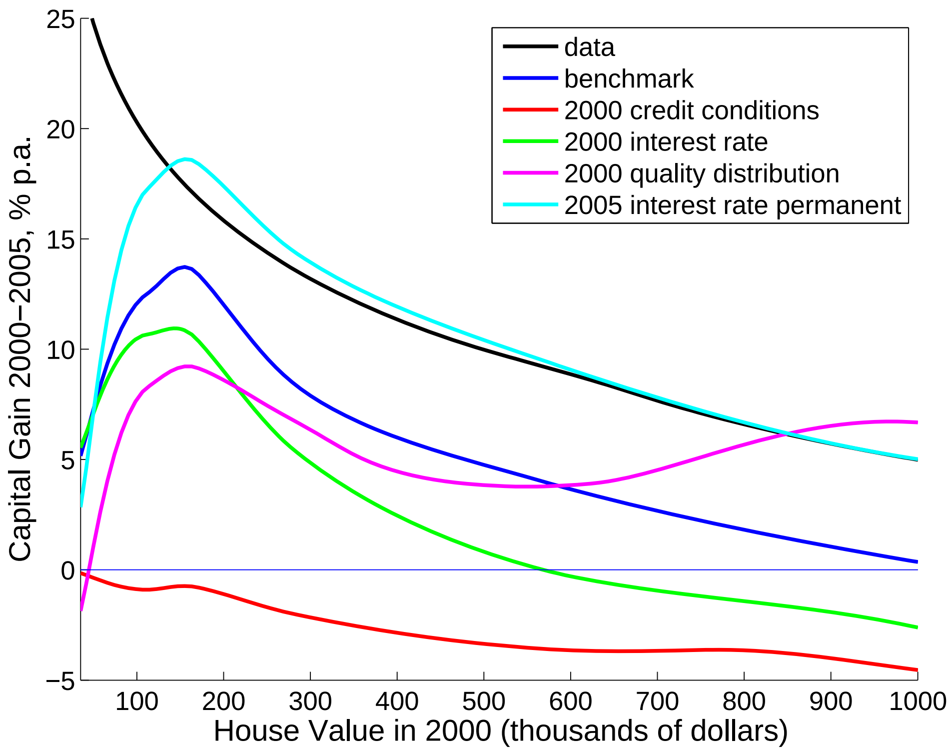

Figure 8: Capital gains per year, 2000-2005, for different experiments

prices. With the 2000 quality distribution, fewer houses at the low end lead to a less steep price curve. These houses are instead in the middle range of the distribution, where the price curve needs to become steeper to induce the assigment.

Interest rate expectations

The blue benchmark lines assume that households expect interest rates to stay at 2005 levels for one period, and then to revert back to higher 2000 levels. The cyan lines shows what happens when households expect interest rates to stay at 1% permanently. With lower expected future interest rates, future cash flows of houses are discounted less, and so current house prices are higher than in the benchmark case.

References

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo 2011. "Understanding Booms and Busts in Housing Markets." *NBER Working Paper* 16734.

Case, Karl E. and Christopher J. Mayer 1996. "House Price Dynamics within a Metropolitan Area." *Regional Science and Urban Economics* 26, pp. 387-407.

Campbell, John and Joao Cocco 2010. "A Model of Mortgage Default." Working Paper, Harvard.

Campbell, John, Stefano Giglio, and Parag Pathak 2010. "Forced Sales and House Prices." Forthcoming *American Economic Review.*

Case, Karl E. and Robert J. Shiller 2005. "Is there a Bubble in the Housing Market?" *Brookings Papers on Economic Activity* 2: 2003.

Caplin, Andrew and John Leahy 2010. "A Graph Theoretic Approach to Markets for Indivisible Goods." Working Paper, NYU.

Caplin, Andrew, Sumit Chopra, John Leahy, Yann LeCun, and Trivikraman Thampy 2008. "Machine Learning and the Spacial Structure of House Prices and Housing Returns." Working Paper, NYU.

Chatterjee, Satyajit and Burcu Eyigungor 2009. "Foreclosures and House Price Dynamics: A Quantitative Analysis of the Mortgage Crisis and the Foreclosure Prevention Policy." Working Paper, Philadelphia Fed.

Cocco, Joao F., Francisco J. Gomes, and Pascal J. Maenhout 2005. "Consumption and Portfolio Choice over the Life Cycle." *Review of Financial Studies* 18(2), pp. 491-533.

Favilukis, Jack, Sydney Ludvigson, and Stijn van Nieuwerburgh 2010. "The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk-Sharing in General Equilibrium." Working Paper, NYU.

Flavin, Marjorie and Takashi Yamashita 2002. "Owner-Occupied Housing and the Composition of the Household Portfolio." *American Economic Review* 92(1), p. 345-362.

Glaeser, Edward L., Joseph Gyourko, and Raven E. Saks 2005. "Why Have Housing Prices Gone Up?" *American Economic Review* 95 (2), pp. 329-333.

Glaeser, Edward L., Joshua Gottleb, and Joseph Gyourko 2010. "Can Cheap Credit Explain the Housing Boom?" Working Paper, Harvard.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst 2010. "Endogenous Gentrification and Housing Price Dynamics." Working Paper, Chicago Booth.

Himmelberg, Charles, Christopher Mayer, and Todd Sinai 2005. "Assessing High House Prices: Bubbles, Fundamentals and Misperceptions." *Journal of Economic Perspectives* 19(4), pp. 67-92.

Jovanovic, Boyan 2009. "The Technology Cycle and Inequality." *Review of Economic Studies* 76, pp. 707-729.

Kahn, James A. 2008. "What drives House Prices?" Working Paper, Federal Reserve Bank of New York.

Kiyotaki, Nobuhiro, Alexander Michaelides, and Kalin Nikolov 2010. "Winners and Losers in Housing Markets." Working Paper, Princeton.

Kaneko, Mamoru 1982. "The Central Assignment Game and the Assignment of Markets". *Journal of Mathematical Economics* 10, pp. 205-232.

Lamont, Owen and Jeremy Stein 1999. "Leverage and House-Price Dynamics in U.S. Cities." *Rand Journal of Economics*, 30(3), pp. 498-514.

Lustig, Hanno and Stijn van Nieuwerburgh 2005. "Housing Collateral, Consumption Insurance and Risk Premia: an Empirical Perspective." *Journal of Finance* 60 (3), pp. 1167-1219

Mayer, Christopher J. 1993. "Taxes, Income Distribution, and the Real Estate Cycle: Why all Houses do not Appreciate at the Same Rate." *New England Economic Review* 39

Mian, Atif and Amir Sufi 2009. "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis." *Quarterly Journal of Economics* 124(4), 1449-1496.

Mian, Atif and Amir Sufi 2010. "House Prices, Home Equity-Based Borrowing, and the U.S. Household Leverage Crisis." Forthcoming, *American Economic Review*.

Ortalo-Magne, Francois and Sven Rady 2006. "Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints." *Review of Economic Studies*, 73:459–485, April 2006

Piazzesi, Monika, Martin Schneider, and Selale Tuzel 2007. "Housing, Consumption and Asset Pricing." *Journal of Financial Economics* 83, pp. 531-569.

Piazzesi, Monika and Martin Schneider 2009. "Momentum Traders in the Housing Market: Survey Evidence and a Search Model." *American Economic Review* 99(2), pp. 406-411.

Piazzesi, Monika and Martin Schneider 2010. "Inflation and the Price of Real Assets." Working Paper, Stanford.

Poterba, Jim 1991. "House Price Dynamics: The Role of Tax Policy and Demography." *Brookings Papers on Economic Activity* 2, pp. 143-202

Rios-Rull, Jose-Victor and Virginia Sanchez-Marcos 2008. "An Aggregate Economy with Different House Sizes." *Journal of The European Economic Association*, 6(2-3),pp. 705-14.

Sattinger, Michael 1993. "Assignment Models of the Distribution of Earnings." *Journal of Economic Literature* 31(2), pp. 831-880.

Stein, Jeremy 1995. "Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects." *Quarterly Journal of Economics* 110(2), pp. 379-406.

van Nieuwerburgh, Stijn and Pierre-Olivier Weill 2010. "Why Has House Price Dispersion Gone Up?" *Review of Economic Studies* 77(4), pp.1567-1606.

Appendix

A Census and Wealth Data

The Census data reports house prices in ranges and top codes them. This appendix explains how we deal with these issues. Moreover, Census data does not contain wealth information. This appendix explains how we use data from the Survey of Consumer Finances to impute household wealth.

Price ranges. The Census data does not contain actual prices but rather price ranges. Let p^c denote the set of prices that delineate the census prices ranges. Apart from the actual census range cutoffs, it contains a lower bound of zero for the lowest (bottom-coded) range and an upper bound of \$5mn for the highest, top-coded range. The upper bound \$5mn is consistent with our transactions data. Sensitivity analysis has shown that the price function and quality distribution are not sensitive to changing this upper bound to e.g. \$3mn or \$7mn, except obviously in the topcoded range, which contains 9.6% of houses in 2005 and 1.8% of houses in 2000.

We assume that the house quality distribution is continuous. Let $G_0(p^c)$ denote the vector obtained by evaluating G_0 at every element of p^c (that is, a vector of cumulative probability masses assigned to each Census price range). To obtain a continuous distribution, we fit cubic splines through $(p^c, G_0(p^c))$. The same procedure is also applied to the quality distribution in the later year $t > 0$, where we have a set of points $(p^c, G_t(p^c))$.

Top coding. We have replaced the top coded value with the mean of the distribution of homes transacted in San Diego with transaction prices higher than the top coded value. Moreover, we have used data on property tax payments from the Census. These payments are also reported in ranges and top-coded, but the ranges are finer and the top-coding is higher than the top-coded house value. Therefore, we have run interval-regressions of house values on property tax payments, and investigate the predicted values.

Wealth. For age and income, we use age of the household head and income reported in the 2000 Census (for $t = 2000$) and 2005 ACS (for $t = 2005$). We are thus given age and income, as well as a survey weight, for every survey household. However, Census data do not contain wealth. To correct for this, we construct a conditional distribution of wealth using data from the Survey of Consumer Finances (SCF). We use the 1998 and 2004 SCF to build the conditional distributions for 2000 and 2005, respectively.

We use a chained equations approach to perform imputations. The estimation is in two steps. In the first step, we use SCF data to run regressions of log net worth on log housing wealth, a dummy for whether the household has a mortgage and if yes, the log mortage value, and log income for each age decade separately. In the second step, we use a regression switching approach that draws regression coefficients to generate a distribution of wealth for a given set of regressors.

For each original household in the census sample, we then create three households with the same income and age, but with different wealth levels given by each of the three possible realizations for wealth using our imputation method. A survey weight for each new household is obtained by dividing the original survey weight by one third.