Panel Data Analyses of Urban Economics and Housing Markets

By

Nai Jia Lee
Prior Masters in Real Estate
National University of Singapore
Singapore (2006)

Submitted to the Department of Urban Studies and Planning
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Urban and Regional Planning

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2009

Author____
Department of Urban Studies and Planning
(date of signing)

Certified by ____
Professor William Wheaton
Department of Economics
Dissertation Supervisor

Accepted by____
Professor Eran Ben-Joseph
Chair, Ph.d Committee
Department of Urban
Studies and Planning

Panel Data Analyses of Urban Economics and Housing Markets

By

Nai Jia Lee
Submitted to the Department of Urban Studies and Planning
on Sep 2009in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Urban and Regional Planning

ABSTRACT

The thesis looks three pertinent issues in Housing Market and Urban Economics literature with panel data- home sales and house price relationship, efficiency of housing market and commercial property taxation. For the first part, I examine the strong positive correlation that exists between the volume of housing sales and housing prices. I develop a simple model of these flows which suggests they generate a negative price-to-sales relationship. This runs contrary to a different literature on liquidity constraints and loss aversion. Our results from both are strong and robust. Higher sales "Granger cause" higher prices, but higher prices "Granger cause" both lower sales and a growing inventory of units-for-sale. These relationships together provide a more complete picture of the housing market – suggesting the strong positive correlation in the data results from frequent shifts in the negative price-to-sales schedule.

For the second part, I tested the hypothesis whether the housing market is efficient and whether "bargains" can be found in the market or not. According to the User cost model, house price appreciation is positively correlated to price. Nevertheless, such correlation between price and appreciation can be caused by productivity differences, behavioral reasons or high transaction costs. Using 4 unique sets of panel data at zip code level, I am able to test the efficiency hypothesis without worrying about productivity reasons and transaction costs. In addition, I tested the efficiency hypothesis by removing influences caused by changes in buyers' preferences over time. The results show that appreciation and house price is positively correlated in San Diego, Boston and Phoenix. However, appreciation and house price is negatively correlated in Chicago.

For the last part, I examine an unusual phenomenon in Massachusetts, where some municipals impose a high property tax on commercial properties and low tax on residential properties. Unlike past studies, we treat the tax on firms as an entrance fee or compensation for the negative externalities the firms generate. This approach fits our context better because we are dealing with municipals- most of the individuals don't work where they live, and the firms are unlikely to provide them employment or other benefits. I develop a simple model to capture the firms' location decision and residents' demand for services and aversion to firms. The model suggests that rich neighborhoods tend to impose high commercial and residential property tax, as they try to reduce their reliance on firms for services. In addition, the municipals will impose a high commercial property tax rate if the number of firms in municipal is large. I assembled a panel data base covering 351 municipals over a period from 1975-2007. The empirical results strongly support the model, suggesting rich municipals rely less on firms.

Thesis Supervisor: William Wheaton

Title Professor of Economics

Thesis Committee Member: David Geltner

Title Professor of Real Estate Finance, Director of Research, Department of Urban Studies & Planning

Thesis Committee Member: Lynn Fisher

Title Associate Professor of Real Estate, Department of Urban Studies & Planning

Acknowledgement

I am grateful to many people that make this possible. I like to thank National University of Singapore for their generous financial support. I also like to thank Prof Ong Seow Eng, Assoc. Prof Sing Tien Foo, Assoc Prof Fu Yuming, Assoc Prof Sim Loo Lee and others who have helped me greatly in the interview and securing the scholarship.

I am extremely fortunate to work with Professor William Wheaton, my thesis supervisor, and I thank for his support and help over the past 5 years. He has taught me how to approach a research problem systematically, and how to think as a researcher. I also like to thank Professor David Geltner, my academic supervisor, for his help in settling in DUSP. Despite his heavy schedule as the director of Real Estate Center, he will always have time to listen to my problems and help me whenever he can. I also like to thank Prof Karen Polenske for generously letting me be her teaching assistant, which help financed my stay here to complete my thesis.

In addition, I like to thank Sandy and the administrative staff in DUSP for their help, especially when I have missed the deadlines of submission.

Lastly, I like to thank my family and my girl friend for their support. Even though they had problems of their own, they always make time available to talk to me and strengthen my belief to finish the whole course.

Table of Contents

Lists of Tables

List of Figures

Chapter 1: Introduction

Research in housing market at the macroeconomic level has been hindered by the availability of data- the time series data for house price, income and sales volume are short. The advent of panel data analysis has helped us overcome such problems, as long as we have had sufficiently many observations for each cross-section. Yet, when we like to find the "granger causality" between two variables using panel data, we are not able to get consistent estimates if we use the usual fixed effects regression model. By construction, the lagged dependent variable is correlated to the fixed effects. To correct such effects, Holtz-Eakin et al (1988) made use of the historical values as instrumental variables. Arellano Bond (1991) further extended the techniques. With such techniques at hand and the greater availability of data, we are able to better understand the housing markets at the regional level. Not only can we test some the theories postulated in the housing and urban economics literature, we can learn from the empirical results to come up with new theories in understanding the housing market in U.S. My thesis uses these panel data and techniques to explore three areas in the urban economics literature: a) the price and sales relationship in housing markets , b) the relationship between commercial property taxes and wages and employment at municipal level and c) the efficiency of housing market at the zip code level.

The first part of my thesis examines the co-movement between sales and prices in the housing market. I first study the gross housing flows in U.S., which can be divided into two types- owner-to-owner churn and sales involving changing of tenure. Much of the literature emphasizes on the positive schedule between own-to –own sales and price, but the latter type of sales represent the majority of sales. I developed a simple model of the flows and the model suggests that price negatively affect price. The implication runs against the loss aversion and liquidity constraints theory. We assemble two data sets and test our hypotheses of our model.

In the second part of my thesis, I study the efficiency of housing markets at zip code level. According to the user cost model, the costs that a homeowner incurred from owning a house includes the expected appreciation of house price. Hence, how much the home owner is willing to pay depends on the expected appreciation. If the market is efficient, home owners will pay a high (low) price for a housing unit with low (high) user costs and high (low) expected appreciation. Empirically, we will observe high (low) price locations experiencing high (low) appreciation if the market is efficient. Nevertheless, such persistent price dispersion could be caused by productivity differences, high transaction costs and other behavioral reasons. We assemble a unique set of data that traces changes in price at zip code level within the city, that allows us to examine price dispersion without some of the problems in previous empirical studies.

For the final part of my thesis, I investigate why local municipals select property classification for property tax purposes, and how and why the tax rates for different types of properties differ using Massachusetts as a case study. From anecdotal evidence, we found that the commercial property tax rates are usually set higher than residential tax rates. We look at how the rates affect the employment and income of the municipal and in turn how the employment and income affect the setting of property tax rates. With large number of municipals and the availability of tax and economic data over 30 years, we are able to get some new results and contribute to the housing and public finance literature.

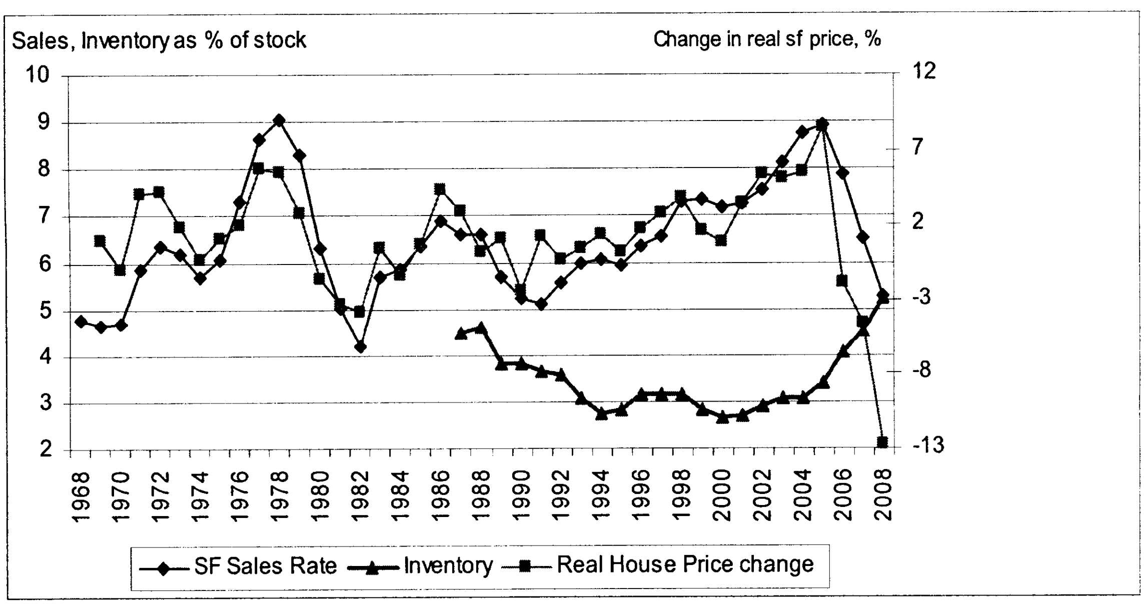
The thesis will comprise of 5 chapters. I will dedicate a chapter for each of the three research questions I proposed. Within each chapter, I will go through the motivation, the research question, literature review, empirical strategy, and analysis of results. In the last chapter, I will conclude with some ideas for future research in this field.

Chapter 2 The Co-Movements of Housing Sales and House Price: Theory and Empirics

2.1 Introduction.

As shown in Figure 1 below, there is a strong positive correlation between housing sales (expressed as a percent of all owner households) and the movement in housing prices (R^2=.66). On the surface the relationship looks to be close to contemporaneous. There is also a somewhat less obvious negative relationship between prices and the shorter series on the inventory of owner units for sale (R^2=.51). A number of authors have offered explanations for these relationships, in particular that between prices and sales.

Figure 1:US Housing Sales, Prices , Inventory



Source: Census, National Association of Realtors

In one camp, there is a growing literature of models describing home owner "churn" in the presence of search frictions [Wheaton (1990), Berkovec and Goodman (1996), Lundberg and Skedinger (1999)]. In these models, buyers become sellers – there are no entrants or exits from the market. In such a situation the role of prices is complicated by the fact that if participants pay higher prices, they also receive more upon sale. It is the transaction cost of owning 2 homes (during the moving period) that grounds prices. If prices are high, the transaction costs can make trading expensive enough to erase the original gains from moving. In this environment Nash-bargained prices move almost inversely to sales duration – equal to the vacant inventory divided by the sales flow. In these models, both the inventory and sales churn are exogenous. Following Pissarides (2000) if the matching rate is exogenous or alternatively of specific form, the sales time will be shorter with more sales churn and prices therefore higher. Hence greater sales cause higher prices. Similarly greater vacancy (inventory) raises sales times and causes lower prices.

There are also a series of papers which propose that negative changes in prices will subsequently generate lower sales volumes. This again is a positive relationship between the two variables, but with opposite causality. The first of these is by Stein (1995) followed by Lamont and Stein (1999) and then Chan (2001). In these models, liquidity constrained consumers are again moving from one house to another ("churn") and must make a down payment in order to purchase housing. When prices decline consumer equity does likewise and fewer households have the remaining down payment to make the lateral move. As prices rise, equity recovers and so does market liquidity. Relying instead on behavior economics, Genesove and Mayer (2001) and then Englehardt (2003) show empirically that sellers who would experience a loss if they sell set higher reservations than those who would not experience a loss. With higher reservations, the market as a whole would see lower sales if more and more sellers experience loss aversion as prices continue to drop.

In this section I try to unravel the relationship between housing prices and housing sales, and in addition, the inventory of housing units for sale. I accomplish the following:

1). First, I carefully examine gross housing flows in the AHS for the 11 (odd) years in which the survey is conducted and find there are more purchases of homes by renters or new households than there are by existing owners. Hence the focus on own-to-own trades does not characterize the majority of housing sales transactions.

2). I also examine which flows add to the inventory of for-sale units (called LISTS) and which subtract (called SALES). Own-to-own moves, for example do both. I show with a simple model that higher prices should generate more LISTS, lower SALES, and hence a larger inventory. When prices are low, the reverse happens.

3). This leads me to hypothesize a very specific joint causality between sales and prices. Own-to-own churn generates a positive schedule between sales and prices as suggested by frictional market theory. At the same time, inter-tenure transitions should lead to a negative schedule. In equilibrium, the overall housing market should rest at the intersection of these two schedules.

4). To test these ideas I first assemble a US panel data base of 33 MSA from 1999-2008. The shortness of the panel is due to limited data on the for-sale inventory. An estimated panel VAR model perfectly confirms our hypothesized relationships. Sales positively drive subsequent prices while prices negative drive subsequent sales and positively increase the inventory.

5). I also assemble a longer panel of 101 MSA from 1980 to 2006 on just sales and prices. Using a wide range of model specifications and tests of robustness we find again that sales positively "Granger cause" subsequent housing price movements, while prices negatively "Granger cause" subsequent housing sales. These joint relationships are exactly as my model suggests when owner churn is combined with inter-tenure transitions.

The chapter is organized as follow. In section 2.1, I set up an accounting framework for more completely describing gross housing flows from the 2001 AHS. This involves some careful assumptions to adequately document the magnitude of all the inter tenure flows relative to within tenure churn and to household creation/dissolution. In Section 2.3, I develop a simple stylized model of the inter tenure flows to illustrate how they can generate a negative relationship between prices and sales and a positive relationship between prices and inventory. I also present my hypothesized pair of relationships between prices and the sales/inventory ratio. In section 2.4, we test these ideas with a short panel data base (33 MSA) that covers the inventory as well as prices and sales. In sections 2.5 through 2.7 I present an analysis of a longer panel data set between just sales and prices across 101 MSA covering the years from 1981-2006. Here again I find conclusive evidence that sales positively "Granger cause" prices and that prices negatively "Granger cause" sales. My analysis is robust to many alternative specifications and subsample tests. I conclude with some thoughts about future research as well as the outlook for US house prices and sales.

2.2 US Gross Housing Flows: Sales, Lists, and the Inventory.

Much of the theoretical literature on sales and prices investigates how existing homeowners behave as they try and sell their current home to purchase a new one. This flow is most often referred to as "churn". To investigate how important a role "churn" plays in the ownership market, we closely examine the 2001 American Housing Survey. In "Table 10"of the Survey, respondents are asked what the tenure was of the residence previously lived in – for those that moved during the last year. The total number of moves in this question is the same as the total in "Table 11" – asking about the previous status of the current head (the respondent). In "Table 11" it turns out that 25% of current renters moved from a residence situation in which they were not the head (leaving home, divorce, etc.). The fraction is a smaller 12% for owners. What is missing is the joint distribution between moving by the head and becoming a head. The AHS is not strictly able to identify how many current owners moved either a) from another unit they owned b) another unit they rented or c) purchased a house as they became a new or different household.

To generate the full set of flows, we use information in "Table 11" about whether the previous home was headed by the current head, a relative or acquaintance. I assume that all current owner-movers who were also newly created households - were counted in "Table 10" as being part of a previous owner household. For renters, we assume that all renter-movers that were also newly created households were counted in "Table 10" in proportion to renter-owner households in the full sample. Finally, we use the Census figures that year for the net increase in each type of household and from that and the data on moves we are able to identify household "exits" by tenure. Gross household exits occur mainly through deaths, institutionalization (such as to a nursing home), or marriage.

Focusing on just the owned housing market, the AHS also allows us to account for virtually all of the events that add units to the inventory of houses for sale (herein called LISTS) and all of those transactions that remove units from the inventory (herein called SALES). There are two exceptions. The first is the net delivery of new housing units. In 2001 the Census reports that 1,242,000 total units were delivered to the for-sale market. Since we have no direct count of demolitions[1] we use that figure also as net and it is counted as additional LISTS. The second is the net purchases of 2nd homes, which must be counted as additional SALES, but about which there is simply little data[2]. In theory, LISTS – SALES should equal the change in the inventory of units for

[1] The growth in stock between 1980-1990-2000 Censuses closely matches summed completions suggesting negligible demolitions over those decades. The same calculation between 1960 and 1970 however suggests removal of 3 million units.

[2] Net second home purchases might be estimated from the product of: the share of total gross home purchases that are second homes (reported by Loan Performance as 15.0%) and the share of new homes in total home purchases (Census,

sale. These relationships are depicted in Figure 2 and can be summarized with the identities below (2001 AHS values are included).

SALES = Own-to-Own + Rent-to-Own + New Owner [+ 2nd homes] = 5,281,000

LISTS = Own-to-Own + Own-to-Rent + Owner Exits + New homes = 5,179,000

Inventory Change = LISTS – SALES

Net Owner Change = New Owners – Owner Exits + Rent-to-Own – Own-to-Rent

Net Renter Change = New Renters – Renter Exits + Own-to-Rent – Rent-to-Own

(1)

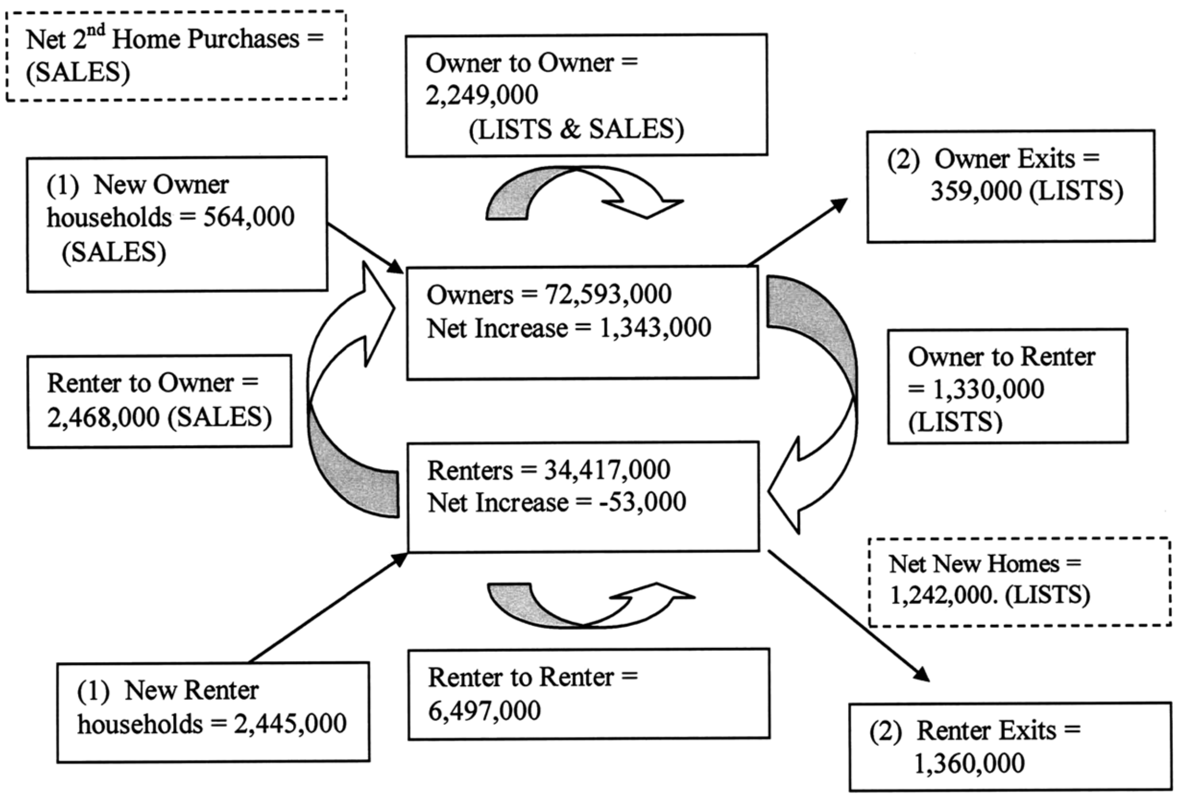
The only other comparable data is from the National Association of Realtors (NAR), and it reports that in 2001 the inventory of units for sale was nearly stable. The NAR however reports a slightly higher level of sales at 5,335,000 existing units.[3] This small discrepancy could be explained by repeat moves within a same year since the AHS asks only about the most recent move. It could also represent some 2nd home sales which again are not part of the AHS move data.

What is most interesting to us is that almost 60% of SALES involve a buyer who is not transferring ownership laterally from one house to another. So called "Churn" is actually a minority of sales transactions. The various inter-tenure sales also are the critical determinants of change-in-inventory since "Churn" sales leave the inventory unaffected.

25%). This would yield 3-4% of total transactions or about 200,000 units. There are no direct counts of the annual change in 2nd home stocks.

[3] The AHS is a repeat sample of housing units and excludes moves into new houses. Thus we compare its move number to NAR sales (both single and multi family) of existing units.

Figure 2: US Housing Gross Flows (2001)



Most inter-tenure SALES would seem to be events that one might expect to be sensitive (negatively) to housing prices. When prices are high presumably new created owner household formation is discouraged or at least deflected into new renter household formation. Likewise moves which involve changes in tenure from renting to owning also should be negatively sensitive to house prices. Both result because higher prices simply make owning a house less affordable. At this time we are agnostic about how net 2nd home sales are related to prices.

On the other side of Figure 2, many of the events generating LISTS should be at least somewhat positively sensitive to price. New deliveries certainly try to occur when prices are high, and such periods would be appropriate for any owners who wish or need to "cash out", consume equity or voluntarily choose to switch to renting. At this time I am still seeking a direct data source which investigates in more detail what events actually generate the own-to-rent moves. Thus the flows in and out of homeownership in Figure 2 suggest that when prices are high sales likely decrease, lists increase and the inventory grows.

The AHS has been conducted only semi-annually until very recently and also has used a consistent definition of moving only since 1985. In Appendix 2.3, I calculate the flows for each of the 11 AHS surveys between 1985 and 2007. The flows are remarkably stable, although there do exist interesting year to year variations. In all years, own-to-own moves ("churn") are less than the sum of new owners plus rent-to-own moves. Since the 2001 survey, the AHS values for LIST-SALES have increased significantly which is consistent with the growing national inventory reported in the NAR data.

2.3. A stylized model of inter-tenure flows.

Here I assume that the total number of households T is fixed with $H \leq T$ being home owners. Those not owning rent at some fixed (exogenous) rent – so I am largely ignoring the rental market. The total stock of units available for ownership U(p) is assumed to depend positively on price (long run supply) and with fixed rents we ignore rental supply. In this situation the inventory of units for sale is the difference between owner stock and owner households: $I = U(p) - H$.

Households flow out of ownership at some constant rate α which could represent unemployment, foreclosure, or other economic shock. Rental households purchase units out of the owner stock (become owners) at some rate s(p) which I presume depends negatively on price. High prices (relative to the fixed rent) make ownership less appealing, but in general renters wish to become owners because of some assumed advantage (the tax subsidy for example).

The equations below summarize both flows (time derivatives) and steady state values (denoted with *). Equation (2) simply says that the stable homeownership rate depends negatively on prices and the constant economic shock rate. When prices generate a sales rate equal to the economic shock rate, homeownership is a reasonable 50%. Equation (3) cleanly divides up the inventory change into the same two categories from our more detailed flow diagram: LISTS-SALES. Here LISTS are stock change (new construction) plus own-to-rent flows (economic shocks) while SALES are rent-to-own flows (whose equilibrium value is the last expression in (3).

$$dH / dt = s(p)[T - H] - \alpha H, \qquad H^* = \frac{s(p)T}{\alpha + s(p)} \tag{2}$$

$$dI / dt = dU / dt - dH / dt = [dU / dt + \alpha H] - s(p)[T - H], \tag{3}$$

$$I^* = U(p) - H^* = \frac{\alpha U(p) + s(p)U(p) - sT}{\alpha + s(p)} \geq 0$$

$$s(p)[T - H^*] = \frac{\alpha s(p)T}{\alpha + s(p)}$$

In equations (3) I must assume a non-negative inventory. Combining equations (2) and (3) I show below that as prices increase, the steady state value of the inventory grows and the steady state level of SALES decreases – as hypothesized about the flows diagramed in Figure 2.

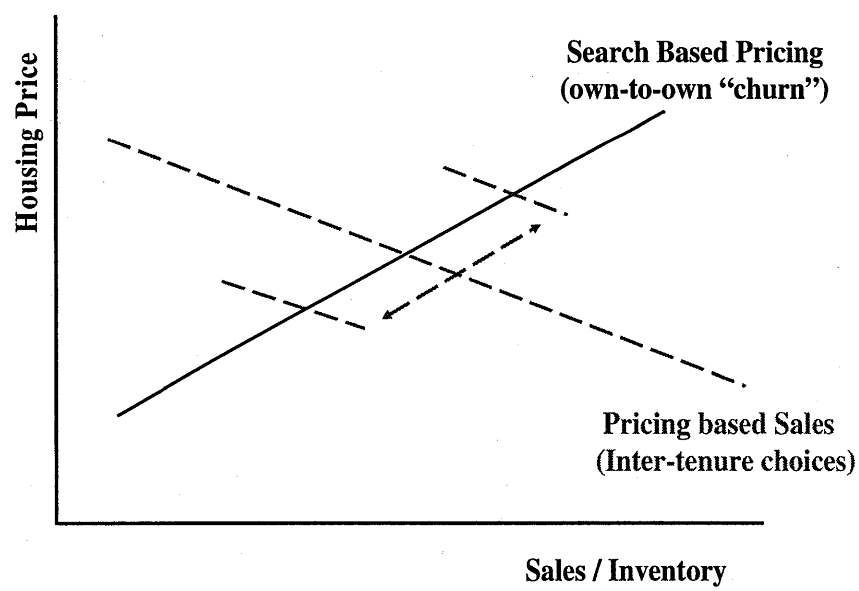
$$dI^* / dp = dU/dp\text{-}dH^*/dp \;=\; dU/dp\text{-}\frac{Tds/dp}{\alpha + s}[1\text{-}\frac{s}{\alpha + s}] \geq 0 \tag{4}$$

$$d\,(s(p)[T\text{-}H^*])/dp = \frac{\alpha Tds / dp}{\alpha + s}[1\text{-}\frac{s}{\alpha + s}] \leq 0$$

Again the conclusions above follow from the *assumptions* that long run stock is positively related to price and the sales *rate* is negative related to price. Thus this simple model of inter-tenure flows establishes a negative relationship between housing prices and subsequent Sales/Inventory ratios. Alternatively, there should be a positive relationship between prices and subsequent duration.

With this new schedule between prices and duration we are now ready to better describe the full set of relationships in the owner market between sales, prices and the inventory. We combine this new schedule with the positive schedule between prices and the Sale/Inventory ratio – created from the various models of own-to-own decisions. In these latter models it is sales that are determining prices, while with the model in (2)-(4) above it is prices that are determining sales. At a more complete equilibrium (in the ownership market) sales, prices and the inventory all rest at the intersection of the two schedules shown in Figure 3. Figure 3 presents a more complete picture of the housing market than the models of Stein, Wheaton, or Berkovec and Goodman – since it accounts for the very large role of inter-tenure mobility as well as for owner churn.

Figure 3: Housing Market Equilibrium(s)



The out-of-equilibrium dynamics of this model are also appealing and seem in line with economic intuition as well. Consider a permanent increase in economic shocks (α). Using (2) and (3), owner households decline, and the inventory increases. Sales however also increase and so the impact on duration is technically ambiguous. Within a wide range of reasonable parameter values however, we can show that the sale/inventory ratio declines with greater α – the net shift in the price-to-sales schedule is therefore inward.[4] The new equilibrium then has prices lower with a lower sales/inventory ratio (a higher duration). If we shift the s(p) schedule up (for example with a greater tax subsidy for ownership) the number of owners increases and the inventory drops, while sales increase. This leads to an unambiguous rightward shift in the price-to-sales schedule with a corresponding equilibrium rise in Sales/Inventory (drop in duration) and rise in prices.

The next task is to see if I can empirically identify the pair of relationships in Figure 3. For this, I turn to several panel data bases with different degrees of richness. The first data base is shorter and covers only 33 MSA. Its advantage is that it includes data on the inventory for sale by market – a series which the NAR has collected only recently. The second data base is much longer, covers 101 MSA, but includes only information on sales and prices.

2.4. A Short-Panel Analysis of Metropolitan Sales, Prices and Inventory.

Carefully constructed data series on Sales and Prices are available since the late 1970s or early 1980s and for a wide range of metropolitan areas. The price data we use is the deflated OFHEO repeat sales series [Baily, Muth, Nourse (1963)]. This data series has recently been questioned for not factoring out home improvements or maintenance and for not factoring in depreciation and obsolescence [Case, Pollakowski, Wachter (1991), Harding, Rosenthal, Sirmans

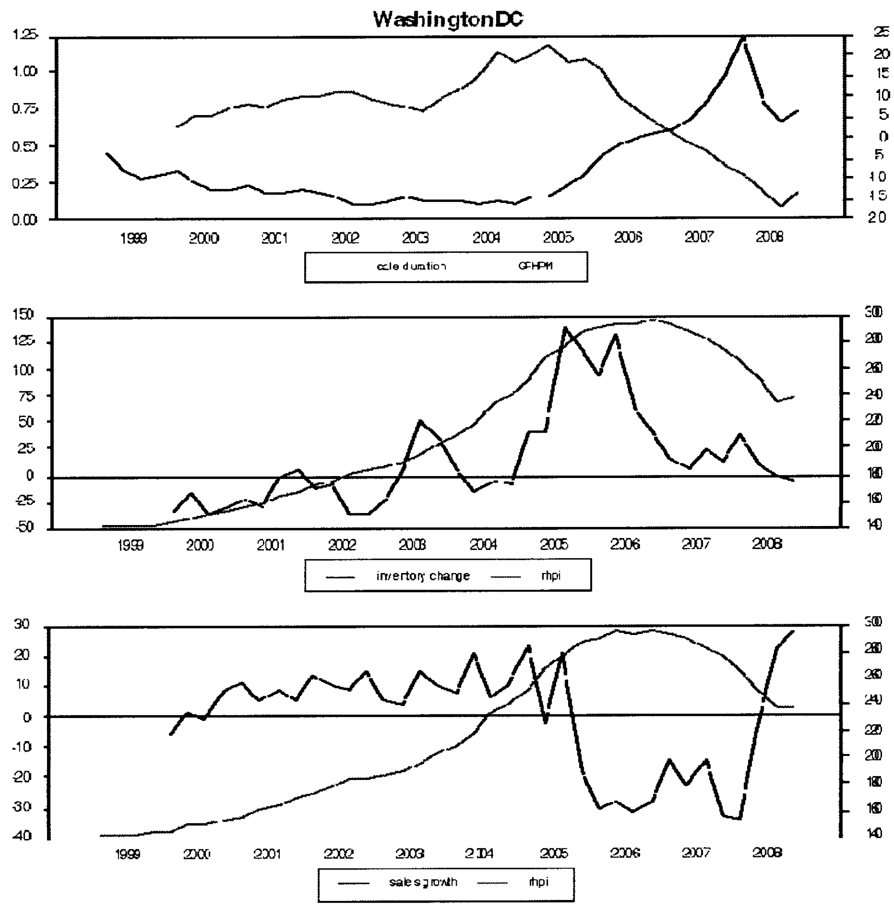
[4] A sufficient condition is for the number of renters [T-H] to exceed the for-sale inventory.

(2007)]. That said I am left with what is available, and the OFHEO index is the most consistent series available for most US markets over a long time period. The only alternative is CSW/FISERV, and it is available for far fewer metropolitan areas that in turn are disproportionately concentrated in the south and west.

In terms of sales, the only consistent source is that provided by the National Association of Realtors (NAR). The NAR data is for single family units only (it excludes condominium sales in the MSA series), but is available for each MSA over the full period from 1980 to 2006. The more limiting data series is that on the inventory of housing units for sale. Here the NAR distributes data only from 1999 (and for many MSA much later). I have been able to put together the three series since 1999 for 33 MSA, and Figure 4 illustrates the data for Washington D.C. In the first frame of Figure 4 we clearly see the strong negative relationship between duration and price (changes). The second frame shows a clear positive relationship between prices and inventory (change), while the final frame shows the negative relationship between prices and sales (changes). In the figures it seems clear that the relationships are not purely static – as modeled in (2)-(4). Rather, at high prices for example, sales ratchet down over time and the inventory ratchets upward.

These observations suggest that a panel VAR is the appropriate instrument to test the relationships. In the VAR I will have each variable depending on lagged values of itself and the other variables. If the panel is of order one, we also can use each coefficient as an effective test of "Granger causality". A final issue with this data involves some strong patterns of seasonality in the data. Prices seem smooth, but sales and the inventory exhibit strong seasonality. This is difficult to model directly since the patterns can be quite different for different MSA. In Figure 4 all changes are calculated as year-over-year, and for this reason we use a 4-period lag throughout the VAR.

Figure 4: Washington Sales, Prices, Inventory



The results of a simple 3 variable VAR are shown in Table 1.[5] In all respects, the results fully support equations (2)-(4) and the pair of relationships in Figure 3. In the first equation, duration is effectively measured with separate Sales and Inventory series. The signs here show duration negatively "Granger cause" subsequent prices. Prices then positively "Granger cause" the inventory to grow, and likewise for sales to decline. All coefficients are highly significant and the large coefficients on the lagged own variable suggest all the series have considerable momentum. The first VAR equation validates the upward schedule in Figure 3 while the second and third

[5] We do not report the cross section fixed effects, which in general are quite significant. Standard errors are shown in parenthesis

combine to yield the downward schedule. In Table 2, I present the results using the Holtz Eakin et al. approach with seasonal effects. The results are consistent with those in Table 1.

Table 1: Quarterly Panel VAR, 1999-2007, 33 MSA

$$P_t = \sum_{j=1}^{33} \alpha_j + \underset{(63.1)}{.833}\, P_{t-4} - \underset{(8.5)}{.0021}\, I_{t-4} + \underset{(9.4)}{.0012}\, S_{t\text{-}4}, \qquad R^2 = .885$$

$$I_t = \sum_{j=1}^{33} \alpha_j + \underset{(53.2)}{.921}\, I_{t-4} + \underset{(10.1)}{93.9}\, P_{t-4} + \underset{(3.3)}{.032}\, S_{t\text{-}4}, \qquad R^2 = .945$$

$$S_t = \sum_{j=1}^{33} \alpha_j + \underset{(52.3)}{.819}\, S_{t-4} - \underset{(6.7)}{102.6}\, P_{t-4} - \underset{(2.4)}{.072}\, I_{t\text{-}4}, \qquad R^2 = .98$$

Table 2: Quarterly Panel VAR Using Holtz Eakin et al Approach with Seasonal effects, 1999-2007, 33 MSA

$$P_t = \sum_{j=1}^{33} \alpha_j + \sum_{i=1}^{4}\sum_{j=1}^{33} \alpha_j quarter_i + \underset{(427.85)}{.9819115}\, P_{t-1} - \underset{(-22.31)}{.0000966}\, I_{t-1} + \underset{(19.79)}{.0006243}\, S_{t\text{-}1},$$

$$I_t = \sum_{j=1}^{33} \alpha_j + \sum_{i=1}^{4}\sum_{j=1}^{33} \alpha_j quarter_i + \underset{(57.69)}{.8847076}\, I_{t-1} + \underset{(4.14)}{64.37069}\, P_{t-1} - \underset{(-1.41)}{.2123897}\, S_{t\text{-}1} + \underset{(1.66)}{385.1749}\, year$$

$$I_t = \sum_{j=1}^{33} \alpha_j + \sum_{i=1}^{4}\sum_{j=1}^{33} \alpha_j quarter_i + \underset{(62.09)}{.8864376}\, I_{t-1} + \underset{(7.64)}{94.84853}\, P_{t-1} - \underset{(-0.57)}{.0774326}\, S_{t\text{-}1}$$

$$S_t = \sum_{j=1}^{33} \alpha_j + \sum_{i=1}^{4}\sum_{j=1}^{33} \alpha_j quarter_i + \underset{(76.60)}{.916015}\, S_{t-1} - \underset{(-1.84)}{2.864344}\, P_{t-1} - \underset{(-2.30)}{.0041445}\, I_{t\text{-}1},$$

2.5. A long-Panel of Metropolitan Sales and Prices.

To more carefully study the relationship(s) between housing sales and housing prices Ialso assembled a larger panel data base covering 101 MSA and the years 1981 through 2006. [6] This panel was purposely structured to be annual so as to avoid the seasonality of the shorter panel, while still maintaining plentiful degrees of freedom.

To better standardize the sales data in this panel raw sales were compared with annual Census estimates of the number of total households in those markets. Dividing single family sales by total households I get estimated sales rate for each market. In a similar manner I set the price level in each market to 100 in the base year. These re-scaling of the data will help make the cross section fixed effects smaller in the estimated VAR models.

In Figures 6 and 7 we illustrate the yearly NAR sales rate data, along with the constant dollar OFHEO price series – both in levels and differences - for two markets that exhibit quite varied behavior, Atlanta and San Francisco. Over this time frame, Atlanta's constant dollar prices increase very little while San Francisco's increased almost 200%. San Francisco prices, however, exhibit far greater price volatility. Atlanta's average sales rate is close to 4% and displays a strong trend over 1980-2006, while San Francisco's is almost half of that (2.6%) and increases by less. These trends illustrate the typical range of patterns seen across our sample of 101 metropolitan areas. In Appendix 2.1, we present the summary statistics for each market's price and sales rate series. In virtually all markets there is a long term positive trend in the sales rate, as well as in real house prices.

[6] There have been a few recent attempts test the relationship between movements in sales and prices. Leung, Lau, and Leong (2002) undertake a time series analysis of Hong Kong Housing and conclude that stronger Granger Causality is found for sales driving prices rather than prices driving sales. Andrew and Meen (2003) examine a UK Macro time series using a VAR model and conclude that transactions respond to shocks more quickly than prices, but do not necessarily "Granger Cause" price responses. Both studies are hampered by limited observations.

The data in Figure 4 and used in the short panel showed no obvious trends; prices, sales and the inventory generally rise and then fall. The longer term series in Figure 5 have more persistent trends and so it is important to test more formally for series stationarity. There are two tests available for use with panel data and in each, the null hypothesis is that *all* of the individual series have unit roots and are non stationary. Levin-Lin (1993) and Im-Persaran-Shin (2002) both develop a test statistic for the sum or average coefficient of the lagged variable of interest – across the individuals (markets) within the panel. The null is that all or the average of these coefficients is not significantly different from unity. In Table 1 we report the results of this test for both housing price and sale rate levels, as well as a 2nd order stationarity test for housing price and sales rate changes.

Figure 5: Atlanta Sales, Prices

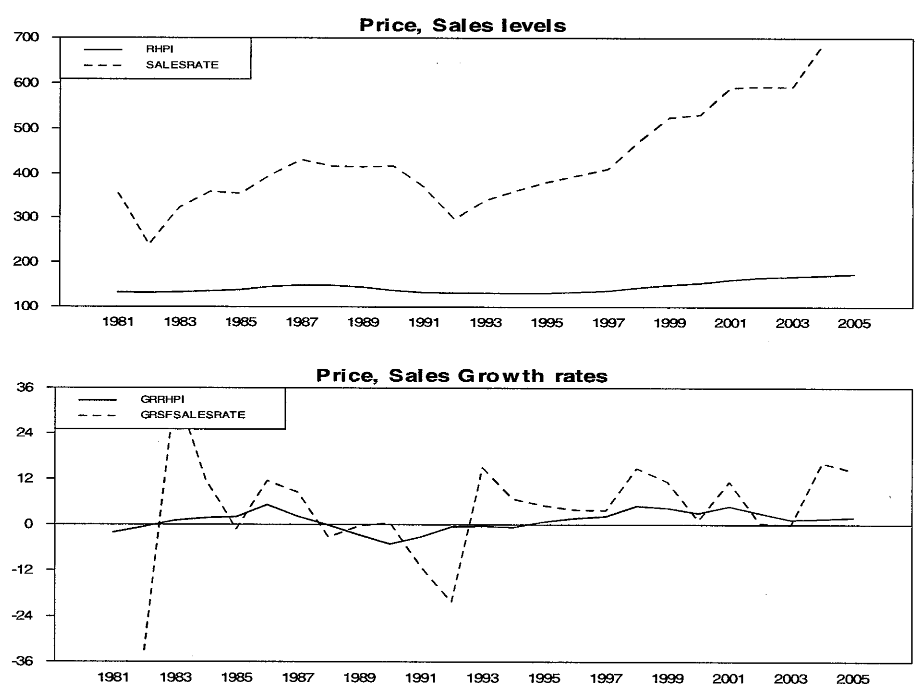


Figure 6: San Francisco Sales, Prices

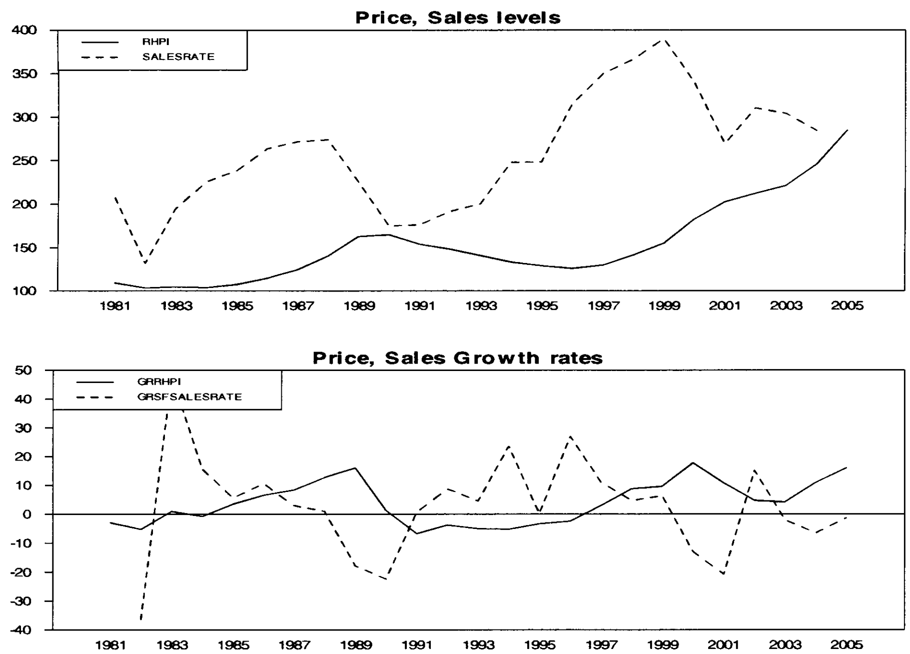


Table 3: Stationarity tests

RHPI (Augmented by 1 lag)

Levin Lin's Test	Coefficient	T Value	T-Star	P>T
Levels	-0.10771	-18.535	0.22227	0.5879
First Difference	-0.31882	-19.822	-0.76888	0.2210
IPS test	T-Bar	W(t-bar)		P>T
Levels	-1.679	-1.784		0.037
First Difference	-1.896	-4.133		0.000

SFSALESRATE (Augmented by 1 lag)

Levin Lin's Test	Coefficient	T Value	T-Star	P>T
Levels	-0.15463	-12.993	0.44501	0.6718
First Difference	-0.92284	-30.548	-7.14975	0.0000
IPS test	T-Bar	W(t-bar)		P>T
Levels	-1.382	1.426		0.923
First Difference	-2.934	-15.377		0.000

With the Levin-Lin test we cannot reject the null (non-stationarity) for either house price levels or differences. In terms of the sales, we can reject the null for differences in sales rate, but not for levels. The IPS test (which is argued to have more power) rejects the null for house price levels and differences and for sales rate differences. In short, both variables would seem to be stationary in differences, but levels are more problematic and likely non-stationary.

2.6. Long Panel Estimations.

My panel approach uses a well-known application of Granger-type analysis. I will ask how significant lagged sales are in a panel model of prices which uses lagged prices and then several conditioning variables. The conditioning variables we choose are market area employment, and national mortgage rates. The companion model is to ask how significant lagged prices are in a panel model of sales using lagged sales and the same conditioning variables. This pair of model is shown (5).

$$P_{i,T} = \alpha_0 + \alpha_1 P_{i,T-1} + \alpha_2 S_{i,T-1} + \beta' X_{i,T} + \delta_t + \varepsilon_{t,\mathrm{T}} \tag{5}$$

$$S_{i,T} = \gamma_0 + \gamma_1 S_{i,T-1} + \gamma_2 P_{i,T-1} + \lambda' X_{i,T} + +\eta_i + \varepsilon_{t,\mathrm{T}}$$

With a significant concern about the stationarity of both series in levels, we estimate (5) largely out of curiosity. More reliable estimates will come from estimating the model in differences – as specified in (6).[7]

$$\Delta P_{i,T} = \alpha_0 + \alpha_1 \Delta P_{i,T-1} + \alpha_2 \Delta S_{i,T-1} + \beta' \Delta X_{i,T} + +\delta_t + \varepsilon_{t,\mathrm{T}} \tag{6}$$

$$\Delta S_{i,T} = \gamma_0 + \gamma_1 \Delta S_{i,T-1} + \gamma_2 \Delta P_{i,T-1} + \lambda' \Delta X_{i,T} + +\eta_i + \varepsilon_{t,\mathrm{T}}$$

In panel VAR models with individual heterogeneity there exists a specification issue. The Equations in (5) or (6) can have an error term that is correlated with the lagged dependent variables [Nickell, (1981)]. OLS estimation will yield coefficients that are both biased and also that are not consistent in the number of cross-section observations. Consistency occurs only in the number of time series observations. Thus estimates and any tests on the parameters of interest (the α and γ) may not be reliable. These problems might not be serious in our case since we have 26 time series

[7] In (6) the fixed effects are cross-section trends rather than cross section levels as in (5)

observations (more than many panel models). To be on the safe side, however, we also estimated the equations following an estimation strategy by Holtz-Eakin et al. As discussed in Appendix 2.2, this amounts to using 2-period lagged values of sales and prices as instruments with GLS estimation.

From either estimates, we conduct a "Granger" causality test. Since I am only testing for a single restriction, the t statistic is the square root of the F statistic that would be used to test the hypothesis in the presence of a longer lag structure (Greene, 2003). Hence, I can simply use a t test (applied to the α_2 and γ_2) as the check of whether changes in sales "Granger cause" changes in price and vice versa.

In Table 3, I report the results of equations (5) and (6) in each set of rows. The first column uses OLS estimation, the second the Random Effects IV estimates from Holtz-Eakin et al. The first set of equations is in levels, while the second set of rows reports the results using differences. In all Tables, variable names are self evident and differences are indicated with the prefix GR. Standard errors are in parenthesis.

Among the levels equations, I first notice that the two conditioning variables, the national mortgage rate and local employment have the wrong signs in two cases. The mortgage interest rate in the OLS price levels equation and local employment in the IV sales rate equation are miss-signed. There is also an insignificant employment coefficient in the OLS sales rate equation (despite almost 2500 observations). Another troublesome result is that the price levels equation has excess "momentum" – lagged prices have a coefficient greater than one. Hence prices (levels) can grow on their own without necessitating any increases in fundamentals, or sales. I suspect that these two anomalies are likely the result of the non-stationary feature to both the price and sales series when measured in levels. Interestingly, the two estimation techniques yield quite similar coefficients – as might be expected with a larger number of time series observations.

When we move to the results of estimating the equations in differences all of these issues disappear. The lagged price coefficients are small so the price equations are stable in the 2nd degree, and the signs of all coefficients are both correct – and highly significant.

As to the question of causality, in every price or price growth equation, lagged sales or growth in sales is always significantly positive. Furthermore in every sales rate or growth in sales rate equation, lagged prices (or its growth) are also always significant. Hence there is clear evidence of joint causality, *and the effect of lagged prices on sales is always of a negative sign.* Holding lagged sales (and conditioning variables) constant, a year after there is an increase in prices – sales fall. This is the opposite of that predicted by theories of loss aversion or liquidity constraints, but fully consistent with the role played by tenure choices in Figure 2 and our simple model of these flows.

Table 4: Sales-Price VAR

	Fixed Effects	E Holtz-Eakin estimator
Levels		
Real Price		
(Dependent Variable)		
Constant	-25.59144**	-12.47741**
	(2.562678)	(2.099341)
Real Price (lag 1)	1.023952**	1.040663**
	(0.076349)	(0.0076326)
Sales Rate (lag 1)	3.33305**	2.738264**
	(0.2141172)	(0.2015346)
Mortgage Rate	0.3487804**	-0.3248508**

	(0.1252293)	(0.1209959)
Employment	0.0113145**	0.0015689**
	(0.0018579)	(0.0003129)
Sales Rate		
(Dependent Variable)		
Constant	2.193724**	1.796734**
	(0.1428421)	(0.1044475)
Real Price (lag 1)	-0.0063598**	-0.0059454**
	(0.0004256)	(0.0004206)
Sales Rate (lag 1)	0.8585273**	0.9370184**
	(0.0119348)	(0.0080215)
Mortgage	-0.063598**	-0.0664741**
	(0.0069802)	(0.0062413)
Employment	-0.0000042	-0.0000217**
	(0.0001036)	(0.0000103)
<u>**First Difference**</u>		
GR Real Price		
(Dependent Variable)		
Constant	-0.4090542**	-0.49122**
	(0.1213855)	(0.1221363)
GR Real Price (Lag 1)	0.7606135**	0.8008682**
	(0.0144198)	(0.0148136)
GR Sales Rate (Lag 1)	0.0289388**	0.1826539**
	(0.0057409)	(0.022255)
GR Mortgage Rate	-0.093676**	-0.08788**
	(0.097905)	(0.0102427)

GR Employment	0.3217936**	0.1190925**
	(0.0385593)	(0.048072)
GR Sales Rate		
(Dependent Variable)		
Constant	0.7075247	1.424424**
	(0.3886531)	(0.3710454)
GR Real Price (Lag1)	-0.7027333**	-0.8581478**
	(0.0461695)	(0.0556805)
GR Sales Rate (Lag 1)	0.0580555**	0.0657317**
	(0.0183812)	(0.02199095)
GR Mortgage Rate	-0.334504**	-0.307883**
	(0.0313474)	(0.0312106)
GR Employment	1.167302**	1.018177**
	(0.1244199)	(0.1120497)

** indicates significance at 5%.

I have experimented with these models using more than a single lag, but qualitatively the results are the same. In levels, the price equation with two lags becomes dynamically stable in the sense that the sum of the lagged price coefficients is less than one. As to causal inference, the sum of the lagged sales coefficients is positive, highly significant, and passes the Granger F test. In the sales rate equation, the sum of the two lagged sales rates is virtually identical to the single coefficient above and the lagged price levels are again significantly negative (in their sum). Collectively higher lagged prices "Granger cause" a reduction in sales. We have similar conclusions when two lags are used in the differences equations, but in differences, the 2nd lag is always insignificant.

As a final test, I investigate a relationship between the *growth* in house prices and the *level* of the sales rate. In the search theoretic models sales rates determine price levels, but if prices are slow to adjust, the impact of sales might better show up on price changes. Similarly the theories of loss aversion and liquidity constraints relate price changes to sales levels. While the mixing of levels and changes in time series analysis is generally not standard, this combination of variables is also the strong empirical fact shown in Figure 1. In Table 4 price changes are tested for Granger causality against the level of sales (as a rate).

Table 5: Sales-Price Mixed VAR

Differences and Levels	Fixed Effects	E Holtz-Eakin estimator
GR Real Price		
(Dependent Variable)		
Constant	-6.61475**	-1.431187**
	(0.3452743)	(0.2550279)
GR Real Price (lag 1)	0.5999102**	0.749431**
	(0.0155003)	(0.0141281)
Sales Rate (lag 1)	1.402352**	0.2721678**
	(0.0736645)	(0.0547548)
GR Mortgage Rate	-0.1267573**	-0.0860948**
	(0.0092715)	(0.0095884)
GR Employment	0.5059503**	0.3678023**
	(0.0343458)	(0.0332065)
Sales Rate		
(Dependent Variable)		
Constant	-0.0348229	0.0358686
	(0.0538078)	(0.0026831)
GR House Price (lag 1)	-0.0334235**	-0.0370619**
	(0.0024156)	(0.0026831)
Sales Rate (lag 1)	1.011515**	1.000989**
	(0.0114799)	(0.0079533)
GR Mortgage Rate	-0.0162011**	-0.0151343**
	(0.0014449)	(0.0014294)
GR Employment	0.0494462**	0.043442**
	(0.0053525)	(0.0049388)

** indicates significance at 5%

In terms of causality, these results are no different than the models estimated either in all levels or all differences. One year after an increase in the *level* of sales, the *growth* in house prices accelerates. Similarly, one year after house price *growth* accelerates the *level* of home sales falls (rather than rises). All conditioning variables are significant and correctly signed and lagged dependent variables have coefficients less than one.

2.7. Long Panel Tests of Robustness.

In panel models it is always a good idea to provide some additional tests of the robustness of results, usually by dividing up either the cross section or time series of the panel into subsets and examining these results as well. Here we perform both tests. First we divide the MSA markets into two groups: so-called "coastal" cities that border either ocean, and "interior" cities that do not. There are 31 markets in the former group and 70 in the latter. The coastal cities are often felt to be those with strong price trends and possibly different market supply behavior. These results are in Table 5. The second test is to divide the sample up by year – in this case we estimate separate models for 1980-1992 and 1993-2006. The year 1992 generally marks the bottom of the housing market from the 1990 recession. These results are depicted in Table 6. Both experiments use just the differences model that seems to provide the strongest results from the previous section.

Table 6:Geographic Sub Panels

	Fixed Effects		E Holtz-Eakin estimator	
	Coastal MSA	Interior MSA	Coastal MSA	Interior MSA
GR Real Price (Dependent Variable)				
Constant	-0.6026028 (0.2974425)	-0.274607** (0.1132241)	-0.543562 (0.3332429)	-.338799** .1054476
GR Real Price (Lag 1)	0.7661637** (0.0255794)	0.7731355** (0.0178884)	0.855731** (0.0351039)	.7834749** .0171874
GR Sales Rate (Lag 1)	0.0608857** (0.0141261)	.0094349* (0.0054047)	0.3475212** (0.0573584)	.0799289** .0198759
GR Mortgage Rate	-0.106036** (0.023653)	-.0866954** (0.0092136)	-0.112101** (0.0278593)	-.0776626** .008816
GR Employment	0.5717489** (0.0978548)	.1978858** (0.0359637)	-0.0434497 (0.153556)	.1617733** .0381004
GR Sales Rate (Dependent Variable)				
Constant	2.098906** (0.7412813)	0.0396938** (0.4541917)	3.03388** (0.7426378)	0.8084169* (0.4261651)
GR Real Price (Lag1)	-0.8320889** (0.0637485)	-0.5447358** (0.0637485)	-0.9763902** (0.0798291)	-0.8519448** (0.0919725)
GR Sales Rate (Lag 1)	-0.0004387 (0.0352049)	0.0770193** (0.0216808)	-0.0350817 (0.0402424)	0.1111637** (0.0251712)

GR Mortgage Rate	-0.2536587** (0.0589476)	-0.3772017** (0.0369599)	-0.2390963** (0.0595762)	-0.3323406** (0.036746)
GR Employment	1.265286** (0.2438722)	1.172214** (0.1442662)	1.102051** (0.2223687)	1.03251** (0.1293764)

Note:

a) *- 10 percent significance. **- 5 percent significance.

b) MSAs denoted coastal are MSAs near the East or West Coast (see Appendix I).

c) MSAs denoted interior are MSAs that are not located at the East or West Coast.

In Table 5, the results of Table 4 hold up remarkably strong when the panel is divided by region. The coefficient of sales rate (growth) on prices is always significant although so-called "costal" cities have larger coefficients. In the equations of price (growth) on sales rates, the coefficients are always significant, and the point estimates are very similar as well. The negative effect of prices on sales rates is completely identical across the regional division of the panel sample. It should be pointed out that all of the instruments are correctly signed and significant as well.

The conclusion is the same when the panel is split into two periods (Table 6). The coefficients of interest are significant and of similar magnitudes across time periods, and all instruments are significant and correctly signed as well. The strong negative impact of prices on sales clearly occurred during 1982-1992 as well as over the more recent period from 1993-2006. With fewer time series observations in each of the (sub) panels in Table 6, the Holtz-Eakin estimates are now sometimes quite different than the OLS results.

Table 7: Time Subpanels

	Fixed Effects		E Holtz-Eakin estimator	
	1982-1992	1993-2006	1982-1992	1993-2006
GR Real Price (Dependent Variable)				
Constant	-2.63937** (0.2362837)	-0.1053808 (0.1453335)	-1.237084** (0.2879418)	-0.2731544 (0.1943765)
GR Real Price (Lag 1)	0.5521216** (0.0271404)	0.9364014** (0.0183638)	0.6752733** (0.0257512)	0.9629539** (0.0196925)
GR Sales Rate (Lag 1)	0.0194498** (0.0073275)	0.0363384** (0.0097935)	0.1622147** (0.0307569)	0.0874362 ** (0.0307703)
GR Mortgage Rate	-0.2315352** (0.0193262)	-0.0707981** (0.0116032)	-0.1432255** (0.0244255)	-0.0812995** (0.0163056)
GR Employment	0.6241497** (0.063533)	0.4310861** (0.0501575)	0.157348* (0.0910416)	0.3441402** (0.0493389)
GR Sales Rate (Dependent Variable)				
Constant	-6.269503** (0.9018295)	4.398222** (0.447546)	-4.898023** (0.8935038)	3.00473** (0.4587499)
GR Real Price (Lag1)	-0.8795382** (0.1035874)	-0.5704616** (0.0565504)	-1.080492** (0.1243784)	-0.4387881** (0.066557)
GR Sales Rate (Lag 1)	0.0056823 (0.027967)	-0.025242 (0.0301586)	-0.0035275 (0.0350098)	0.066557 (0.029539)

GR Mortgage Rate	-0.5636095** (0.0737626)	-0.1934848** (0.0357313)	-0.550748** (0.0819038)	-0.2720118** (0.0420076)
GR Employment	2.608423** (0.2424878)	0.4856197** (0.154457)	2.026295** (0.2237316)	0.7631351** (0.1325586)

Note:

a) Column labeled under 1982-1992 refer to the results using observations that span those years..

b) Column labeled under 1993-2006 refer to the results using observations that span those years.

2.7. Conclusions

We have shown that the causal relationship from prices-to-sales is actually negative – rather than positive. Our empirics are quite strong. As an explanation, we have argued that actual flows in the housing market are remarkably large between tenure groups – and that a negative price-to-sales relationship makes sense as a reflection of these inter-tenure flows. Higher prices lead more households to choose renting than owning and these flows decrease SALES. Higher prices also increase LISTS and so the inventory grows. Conversely, when prices are low, entrants exceed exits into ownership, SALES increase, LISTS decline and so does the inventory.

Our empirical analysis also overwhelmingly supports the positive sales-to-price relationship that emerges from search-based models of housing churn. Here, a high sales/inventory ratio causes higher prices and a low ratio generates lower prices. Thus we arrived at a more complete description of the housing market at equilibrium – as shown with the two schedules in Figure 3.

Figure 3 offers a compelling explanation for why in the data, the simple price-sales correlation is so overwhelmingly positive. Over time it must be the "price based sales" schedule that is shifting up and down. Remember that this schedule is derived mainly from the decision to enter or exit the ownership market. Easy credit availability and lower mortgage rates, for example would shift the schedule up (or out). For the same level of housing prices, easier credit increases the rent-to-own flow, decreases the own-to-rent flow, and encourages new households to own. SALES expand and the inventory contracts. The end result of course is a rise in both prices as well as sales. Contracting credit does the reverse. In the post WWII history of US housing, such credit expansions and contractions have indeed tended to dominate housing market fluctuations [Capozza, Hendershott, Mack (2004)].

Figure 3 also is useful for understanding the current turmoil in the housing market. Easy mortgage underwriting from "subprime capital" greatly encouraged expanded homeownership from the mid 1990s through 2005 [Wheaton and Nechayev, (2007)]. This generated an outward shift in the price-based-sales schedule. Most recently, rising foreclosures have expanded the rent-to-own flow and shifted the "price based sales" schedule back inward. This has decreased both sales and prices. Preventing foreclosures through credit amelioration theoretically would move the schedule upward again, but so could any countervailing policy of easing mortgage credit. It is interesting to speculate on whether there might be some policy that would shift the "search based pricing" schedule upward. This would restore prices, although it would not increase sales. For example some policy to encourage interest-free bridge loans would certainly make it easier for owners to "churn". Likewise some form of home sales insurance might reduce the risk associated with owning two homes. That said, such policies would seem to be a less direct way of assisting the market versus some stimulus to the "price-based-sales" schedule.

Chapter 3 Do Bargains Exist in Housing Market? Efficiency of Housing Market Within Cities

3.1 Introduction

Using a unique panel data set of house prices of neighborhood, we attempt to study whether the market is efficient at the zip code level within the city. Given the availability of neighborhood price data over 10 years, we are able to observe the efficiency of housing markets without worrying about productivity and transaction costs that cost the market to be inefficient. Using the User Cost model as a framework, I study the price trends across neighborhoods using fixed effects and other techniques to understand whether prices across neighborhoods diverge over time or overlap each other periodically. If prices diverge-price of houses in expensive neighborhoods increasing faster than price of houses in cheap neighborhoods, the market is likely be efficient.

Looking for a cheap home with high growth potential in price is a national pastime of many home buyers. While most home buyers are looking for the great "bargain", academics , researchers and policy makers are wrangling over the issue whether such bargains actually exist- or whether the market is efficient. If the market is efficient, the price in the housing market should reflect the net present value of future cash flow accrued to or payable by the homeowner from the house. The cash flow, in this context, is usually referred to the user costs or imputed rent. The user cost of housing is simply the sum of maintenance costs, cost of foregone return to equity, cost of mortgage and net property taxes payable less expected growth in price. Examining whether the user costs exceed the actual rent helps us to identify which regions are underpriced and overpriced above the fundamentals (Himmelberg, Mayer and Sinai, 2005). The key problem arises from such tests come from finding a suitable households' expected growth rate in price. A change in the horizon to compute the expected growth rate will change the outcomes of the analysis.

Besides comparing the user costs with a benchmark, I expect the high house price is positively correlated to risk (Capozza, Hendershott, and Mack [2004], Cannon, Miller and Pandher [2006]) when house price is efficient. In addition, the user costs model implies that the high priced homes must reflect high appreciation in future if the markets are efficient. I also expect places with high priced homes will continue to be priced high and low priced homes to be priced low. Any change in the rankings of places in price will imply the market being efficient. Yet the divergence in prices between high and low rise housing could be due to several reasons- supply restrictions (Capozza, Hendershott, Mack and Mayer, 2002), interest rates movements (Himmelberg, Mayer and Sinai, 2005), behavioral reasons (Brunnermeier and Julliard, 2007 and Case and Shiller, 1988), or difference in productivities across regions (Van Nieuwerburgh and Weil, 2007, and Gyorko, Mayer and Sinai, 2006).

Given that most studies look at changes in prices at Metropolitan Statistical Area (MSA) or State level, it has been hard to disentangle the different forces affecting the differences in prices across regions. In addition, households are less likely to arbitrage given the high transaction costs to move from one MSA or State to another. In this chapter, I try to avoid the problems faced in past studies by examining the differences in house price within a city at zip- code level. With this unique panel data set, supply restrictions and difference in productivity across regions are unlikely to apply. This allows us to examine whether housing market is efficient or not.

In the following section, I begin with a review of the models and theories in explaining the differences in house prices across regions in Section 3.2. Section 3.3 lays out a simple theoretical model to ascertain whether there is inefficiency or not. This is followed by a presentation of our empirical results in Section 3.3 and a discussion on the limitations.

3.2 Literature Review

The standard user cost model is widely used to explain the dynamics of housing price market. Hendershott and Slemrod (1983) used it to examine whether homeowners are favored under property tax, and Poterba (1984) used it to explain how inflation caused the house price to go up via user costs. Together with an inter-temporal housing consumption framework, Poterba (1984) fitted the expectation rational framework in the equilibrium and showed that unanticipated shocks in housing demand will lead to anticipated changes in prices that will alter housing demand. Himmelberg, et.al (2005) used the user cost as a measure to test whether house prices relative to rents in 46 metro areas were over priced in 2004. They found that the user coat to actual rent ratio does not suggest widespread or historically large mispricing of owner-occupied housing in 2004. Dipasquale and Wheaton [1996] further showed how with anticipated growth, certain locations within a city would have both high price levels and high growth rates in Ricardian rent, while other areas would exhibit the reverse. In equilibrium, a "user cost" measure of rent-minus-appreciation would be exactly the same across locations. Wheaton, Seslen and Pollakowski (2007) further incorporated the user costs in a two period housing consumption model. The model further postulates that neighborhoods with high expected house price appreciation and high volatility in prices will have highly priced housing.

Varying liquidity constraints across regions and prohibitive transaction costs can lead to divergence in house prices across regions too. Stein (1995) and Ortalo-Magne and Rady (1999, 2001) argue that liquidity constraints could explain the seemingly sensitivity of house prices to income shocks. Lamont and Stein (1998) and Chan (2001) further shows that house price in places with less liquidity constraints are more volatile. Hence, when there are positive shocks, places with easier access to credit will have higher prices. Besides liquidity constraints, Grossman and Larqoue (1990) show that transaction costs can prevent arbitrage, and the high moving costs can make some areas persistently more costly to live in than others. Meese and Wallace (1994) found that pricing inefficiencies in San Francisco area are due to high transaction costs that limit arbitrage opportunities for rational investors.

Alternatively, several authors argue that psychological factors play a part in house price dynamics. Case and Shiller (1988) surveyed homebuyers in four cities regarding their expectation of future price growth and they found that home buyers are myopic. They reported that homeowners in areas with strong appreciation like Los Angeles in the 1980s, expect greater growth appreciation than those in Milwaukee, where prices are flat in the eighties. Hence, the divergence in prices of the markets is a result of "feedback mechanism or social epidemic that encourages housing as in important investment opportunity." Such feedback expectations can create a housing price bubble in regions with high price appreciation and lead to divergence in prices. Nevertheless, if there is a bubble, we expect the bubble to pop when expectations change in future (Siegel, 2003). As defined by Charles Kindleberger (1987), a bubble as a "sharp rise in price of an asset or range of assets in a continuous process with the initial rise generating expectations of further rises and attracting new buyers….. rise is usually followed by a reversal of expectations and a sharp decline in price often resulting in financial crisis." Empirical studies cannot agree whether the bubble exists or not. In the most recent study, Himmelberg, et.al (2005) examines whether house price bubble exists across different MSAs. They compute the imputed rent (user cost) and compare them with the actual rents. They found that house prices in most regions are not far off from the fundamentals.

Some researchers argue that there exists some form of first nature advantage (Krugmen,) or productivity differences across regions that resulted the price dispersion. Gyourko, Mayer and Sinai (2006) examines why prices in some “Superstar Cities” consistently lead faster house price growth for 50 years and suggest that households rationally expect future prices to rise faster in these “Superstar Cities”. Van Nieuweburg and Weil (2007) examine the long run price dispersion and postulate that productivity differences among regions cause the long run dispersion across cities. By using some actual parameters in their calibration of their model, they found that even without supply constraints, house prices among region will diverge.

Given the numerous forces acting in the housing market, it is difficult to disentangle them to determine whether the market is efficient using the user cost model. My study makes use of a unique set of panel data that comprises of the price indices for each zip code level. I further derive a standard house price using the indices and observe whether there is efficiency in the housing market. Since I am looking at prices changes within the city at zip code level, productivity differences, high transaction costs and liquidity constraints will not explain why house prices in one neighborhood is higher than the other in the same city. After controlling for the locational and physical attributes of each zip code, I can test whether bargains can be found in the market.

3.3 Theory

Suppose each owner faces the user cost, UC, which has the following equation.

$$UC_{i,t} = rP_{i,t} - \frac{\dot{P}_i}{P_{i,t}} \tag{1}$$

where $\dot{P} = E(P_{t+1} - P_t)$.

$P_{i,t}$ refers to the price of housing in location i and at time t. $\dot{P}$ is the expected appreciation price, given above. Unlike the specification in Himmelberg et. al (2006), I only focus on mortgage costs and expected appreciation, and assume the other costs remain costs over time. Equation (1) shows that the higher the expected appreciation, the lower the user costs assuming all things remain the same.

Rearranging the user cost function gives us the following:

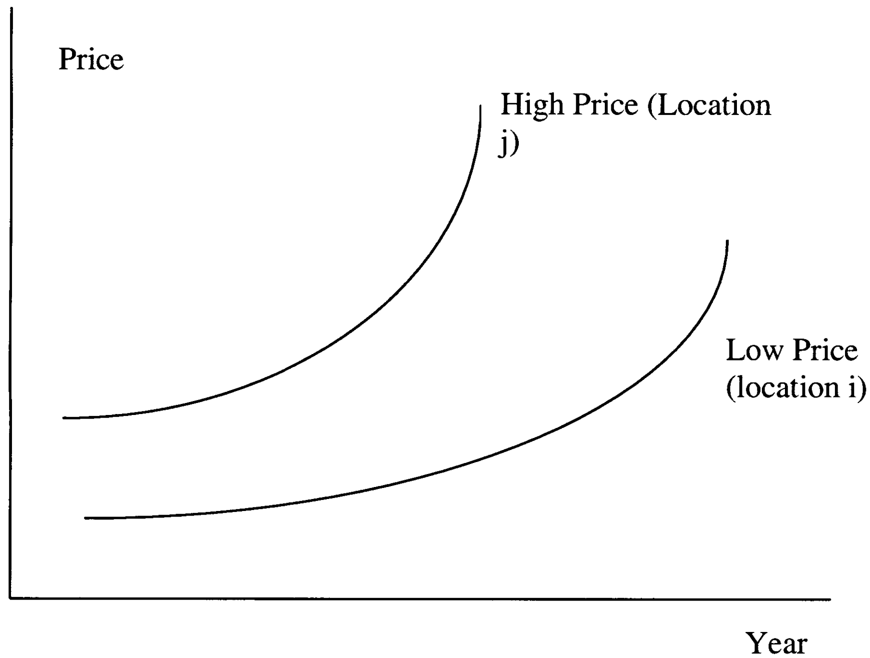
$$\frac{\dot{P}_i}{P_{i,t}} = rP_{i,t} - UC_{i,t} + m \tag{2}$$

If the market is efficient, the user costs for every location at equilibrium are the same (Dispaquale and Wheaton ()). We remove the subscript i as shown below.

$$\frac{\dot{P}}{P_t} = rP_t - UC_t + m \tag{3}$$

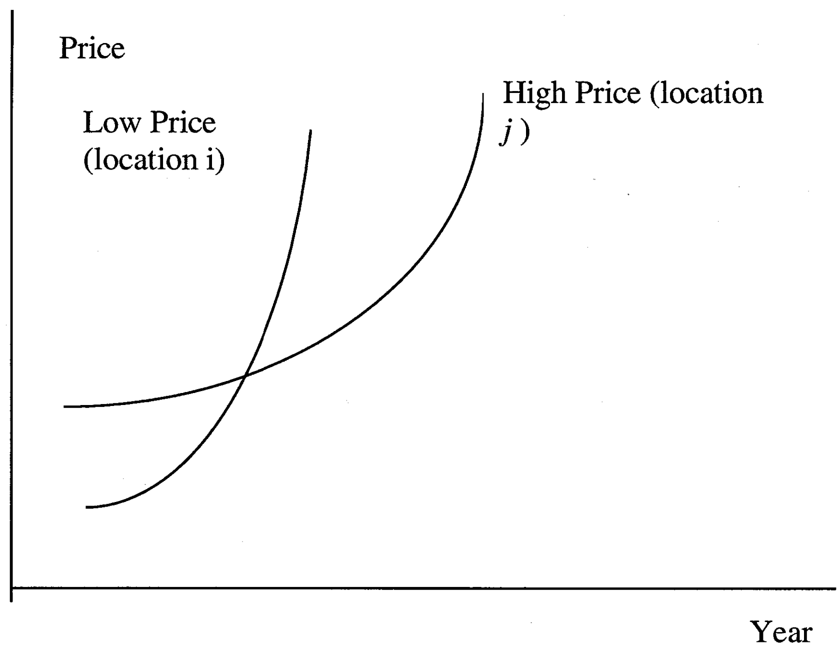
Equation (3) indicates that expected price appreciation and price should be positively correlated. Given maintenance costs and user costs across regions are constant, if price of house in location j is higher than that in location i, then the expected appreciation in location j must be higher. In that case, house price in location j will either move parallel or away from the house price in location i. The price trends are further showed in Fig 7.

Figure 7: Efficient Housing Market



Alternatively if the market is inefficient, we should observe the price trends to converge or overlap. Home buyers are able to get homes cheaply and sell them at a high price at a later date, while home owners of highly priced homes do enjoy similar appreciation in their house price. The price trends between location *i* and *j* will converge and overlap, as shown in Figure 8.

Figure 8: Inefficient Housing Market

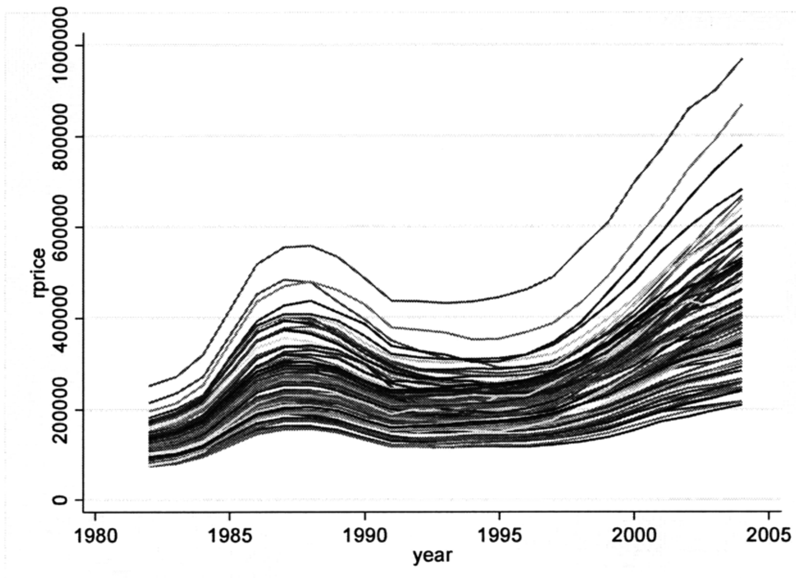


3.4 Empirical Methodologies

The data consists of repeat-sales housing price indices provided to us by Case Shiller Weiss/ FISERV at the zip code level. The data cover four MSAs: Boston, Chicago, Phoenix and San Diego. In choosing these MSAs, I attempt to create a sample representing a diverse set of demographic, geographic and housing market related conditions. The Boston metropolitan area covers 249 ZIP codes from 1982 through 2004, Chicago comprises 152 ZIP codes from 1987 through 2004; Phoenix includes 164 ZIP codes spanning 1998 to 2004 and San Diego covers 86 ZIP codes starting from 1975 to 2004. After dropping the ZIP series that did not meet the length of time series, the final sample consists of 109 ZIP observations for Boston, 51 for Chicago, 80 for Phoenix and 42 for San Diego.

I first conduct a hedonic regression of all 4 MSAs, taking account of all the location and physical attributes of the house. From the hedonic regression, we obtain the price for a benchmarked single family house for each ZIP observation. The hedonic results are presented in the Appendix 3.1. The benchmarked house consists of the following attributes- 2 bedroom unit, built in 1960, and has 1.5 bathrooms. We took the mean lot size and apartment area (of the city) as our benchmark. Similarly, we take the median income and density as our bench mark. From the hedonic regression, we derive the price of benchmarked house in 1998. With the indices, we plot the house prices over the years.

Figure 9: Boston Price Trends

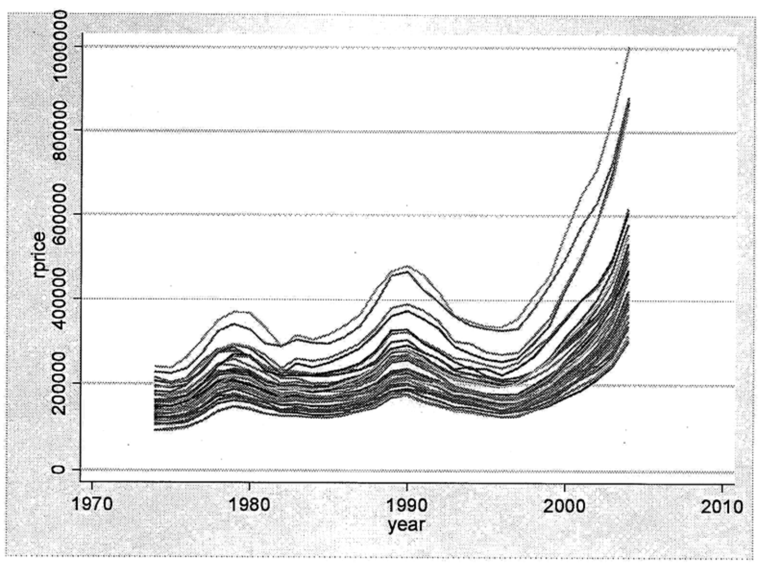


Note: rprice refers to the real price of 2 bed room house with certain features.

Source: Author computation

As shown in Figure 9, Boston price trends have been trending up, peaking once around 1986-1988. The highly priced neighborhoods are located at West Roxbury, MA, and those neighborhoods with the lowest priced housing are located at Wayland, MA. The prices for all the neighborhoods seem to move parallel to one another. For the higher price neighborhoods, their prices seem to diverge over time from the prices of houses in the low price neighborhoods. There is also not much over taking occurring among the prices of neighborhoods that fall in the middle band. Not only do the price diverge, we also find that the ranking in prices of houses do not change much. The top most expensive neighborhoods remain the same for past 20 years. If there is a bubble forming in the expensive neighborhoods, we expect the ranking in house price to change in the future. But for the ranking to stay consistently the same for 20 years means there are more forces in play. The divergence is unlikely caused by productivity and liquidity constraints because they are likely to be same across the neighborhoods within the city. One possible cause for the divergence may be the change in preference over time. Home buyers may find river view or sea view more attractive. Homebuyers may find proximity to good schools more important, as the labor market becomes more competitive.

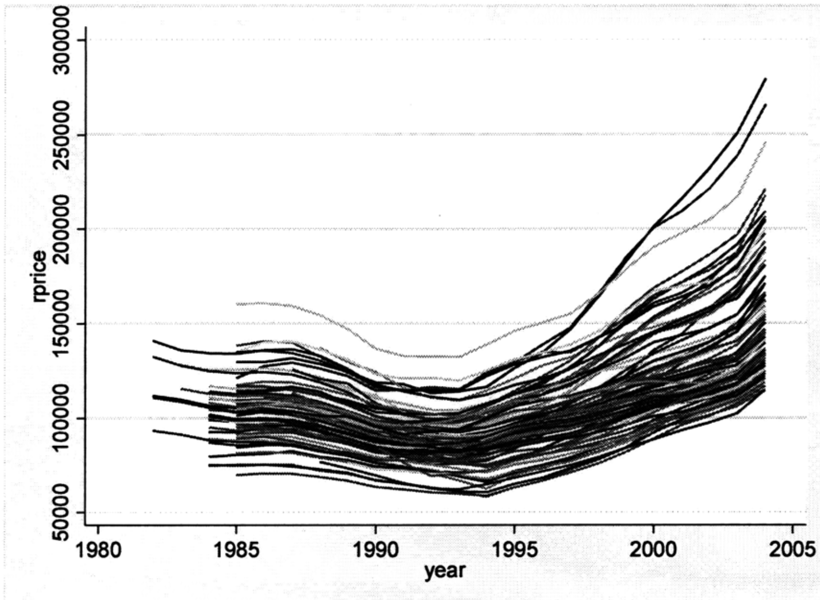
Figure 10: Price trends in San Diego



Note: rprice refers to the real price of 2 bed room house with certain features.

Source: Author computation

Figure 11: Phoenix Price Trends

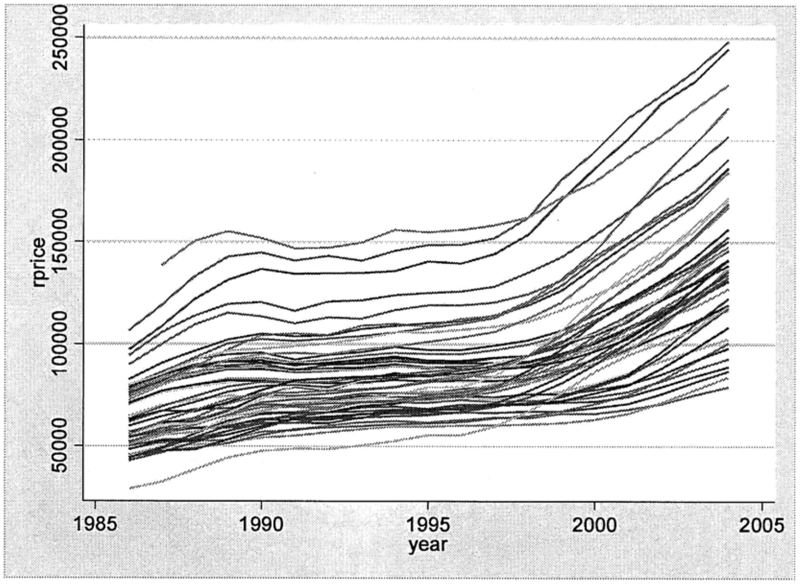


Note: rprice refers to the real price of 2 bed room house with certain features.

Source: Author computation

The house price trends of San Diego and Phoenix are shown in Figures 9 and 10. The price (real) trends up in San Diego, like Boston. In addition the prices in the top three most expensive neighborhoods have a steeper gradient than those prices of the three cheapest neighborhoods in San Diego. In Phoenix, the prices first fell to the lowest in 1993 and picked up again thereafter. The most expensive neighborhood in 1985, shown in the yellow line, occupied the top position until 1998. Prices in two neighborhoods (zip code 85003 and 850018) surpassed the top neighborhood around 1998. The steep rise in prices for both neighborhoods began in 1993, when Phoenix local authorities implemented several initiatives to rejuvenate the economy. The economy also switched from a mining and other primary industries oriented economy to one that emphasized on financial services and research and development.

Figure 12: Chicago Price Trend



Note: rprice refers to the real price of 2 bed room house with certain features.

Source: Author computation

The price trends in Chicago (shown in Figure 12) appear different from the price trends in other 3 MSAs. While the top three most expensive neighborhoods stay the same over the period of observation, there are many changes in ranking from the bottom to mid-level priced neighborhoods. The change in price trends are likely caused by the tax classification system implemented in some towns/cities in Cook County. The higher tax on commercial properties led to a flight of firms from city central, and residential prices in the affected towns are likely to fall.

We tested the change in price again the price level in the following equation of each city. We need to find out how low price locations perform against high price locations. Using a normal pooled regression will not answer what we are looking for. This is because the coefficient captures how appreciation vary against price over time within each neighborhood. To correct this bias, we demean the change in price and lagged prices for each year and city. This is akin to the fixed effects test, except using the years as fixed effects.

$$C_{i,t,k} - C_{.,t,k} = \tilde{\beta}(P_{i,t-k} - P_{.,t-k}) + e_{i,t} \qquad (4)$$

We denote C as the change in price, $\dot{P}$, and β the coefficient of the lagged price, f the year fixed effect and e the error term and *k* represents the number of years in our time window. We conduct the tests for yearly, 2-yearly and 3-yearly and 5-yearly window. In order to avoid auto correlation, we do not include years that overlapped with previous windows. The results are shown in the following tables.

Table 8: Demeaned Change in Price Against Demeaned Prices

3.4.1 Yearly Change in Price against One Year Lagged Price

Boston: Dependent Variable: Yearly Change in Price (demeaned)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 1	5.31e-08	7.17e-09	7.40	0.000	3.90e-08	6.72e-08
R Square	0.0230					
Adjusted R Square	0.0226					

Phoenix: Dependent Variable: Yearly Change in Price (demeaned)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 1	1.48e-07	2.47e-08	6.00	0.000	9.97e-08	1.97e-07
R Square	0.0222					
Adjusted R Square	0.0216					

San Diego: Dependent Variable: Yearly Change in Price (demeaned)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 1	2.83e-08	1.29e-08	2.20	0.028	3.11e-09	5.36e-08
R Square	0.0039					
Adjusted R Square	0.0031					

Chicago: Dependent Variable: yearly Change in Price (demeaned)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 1	-5.06e-08	3.89e-08	-1.30	0.194	-1.27e-07	2.58e-08
R Square	0.0019					
Adjusted R Square	0.0008					

3.4.2 2-Yearly Change in Price against Two Year Lagged Price

Boston: Dependent variable: Change in Price (Demeaned, 2 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 2	1.07e-07	2.32e-08	4.61	0.000	6.15e-08	1.53e-07
R Square	0.0197					
Adjusted R Square	0.0188					

Phoenix: Dependent variable: Change in Price (Demeaned, 2 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 2	2.62e-07	7.08e-08	3.69	0.000	1.23e-07	4.01e-07
R Square	0.0182					
Adjusted R Square	0.0169					

San Diego: Dependent variable: Change in Price (Demeaned, 2 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 2	1.23e-07	3.92e-08	3.13	0.002	4.58e-08	2.00e-07
R Square	0.0138					
Adjusted R Square	0.0123					

Chicago: Dependent variable: Change in Price (Demeaned, 2 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 2	-2.11e-07	1.29e-07	-1.64	0.102	-4.64e-07	4.20e-08
R Square	0.0068					
Adjusted R Square	0.0043					

3.4.3 3-Yearly Change in Price against Three Year Lagged Price

Boston: Dependent variable: Change in Price (Demeaned, 3 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 3	1.88e-07	4.22e-08	4.45	0.000	1.05e-07	2.71e-07
R Square	0.0261					
Adjusted R Square	0.0248					

Phoenix: Dependent variable: Change in Price (Demeaned, 3 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 3	4.67e-07	1.27e-07	3.68	0.000	2.18e-07	7.17e-07
R Square	0.0268					
Adjusted R Square	0.0248					

San Diego: Dependent variable: Change in Price (Demeaned, 3 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 3	7.10e-08	6.89e-08	1.03	0.303	-6.44e-08	2.06e-07
R Square	0.0020					
Adjusted R Square	0.0001					

Chicago: Dependent variable: Change in Price (Demeaned, 3 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 3	-3.64e-07	2.14e-07	-1.71	0.089	-7.85e-07	5.62e-08
R Square	0.0099					
Adjusted R Square	0.0065					

3.4.4 5-Yearly Change in Price against FiveYear Lagged Price

Boston: Dependent variable: Change in Price (Demeaned, 5 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 5	2.86e-07	1.38e-07	2.07	0.039	1.48e-08	5.57e-07
R Square	0.0101					
Adjusted R Square	0.0077					

Phoenix: Dependent variable: Change in Price (Demeaned, 5 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 5	4.85e-07	3.87e-07	1.25	0.212	-2.78e-07	1.25e-06
R Square	0.0063					
Adjusted R Square	0.0023					

San Diego: Dependent variable: Change in Price (Demeaned, 5 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 5	5.85e-07	2.27e-07	2.58	0.010	1.39e-07	1.03e-06
R Square	0.0227					
Adjusted R Square	0.0192					

Chicago: Dependent variable: Change in Price (Demeaned, 5 years return)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned)						
Lag 5	-8.75e-07	6.11e-07	-1.43	0.154	-2.08e-06	3.33e-07
R Square	0.0139					
Adjusted R Square	0.0071					

Source: Author computation

I found that changes in prices for San Diego, Boston and Phoenix are positively influenced by one year lag in price. For Boston, a house that costs one million dollars more than another will experience an additional 0.5% appreciation rate for a year. If a house worth one million dollars more than others in Phoenix, the owner of the house will experience an additional 1.5% appreciation rate compared to the other cheaper homes. The same owner will enjoy less added appreciation if his house is in San Diego- only an additional 0.2% to the appreciation rate of cheaper homes. Chicago, alternatively, has prices in the lagged period negatively correlated to change in prices. A home that costs one million dollars more than the rest will experience 0.5% less appreciation rate than the other cheaper homes. This could be due to the flight of firms caused by the property tax classification. When we broaden the window, the signs remain but most of the coefficients become insignificant. This is not surprising because there are more noises in the broader window.

3.4.1 Correction of preference changes in physical attributes

In my empirical specifications in previous section, we assume no changes in preferences for locations over time. But in reality, people may have increasing preferences over certain physical features of the neighborhood. For instance, home buyers may value sites facing the sea more over time. This created an omission variable problem as the empirical specifications fail to include such changes. In fact, we can regard the change in preference a change in fixed effects. In this section, we try to correct the omission variable specification.

The error term, with changing preferences of fixed effects, can be rewritten as

$$e_{i,t} = \theta_t \alpha_i + \varepsilon_{i,t}\,, \qquad (5)$$

where □ is the individual effect arising from each zip and $\varepsilon_{i,t}$ is the error term for each zip. In year 1998, there is no fixed effect, i.e. $e_{i,t} = \varepsilon_{i,t}$. This is because we derive the prices from the hedonic equations in the same year. Hence, θ_t represents the change in preference for location characteristics with respect to the year 1998.

I first remove the time fixed effects by demeaning the price and return variables, which gives us the following:

$$C_{i,t} - C_{.,t} = \beta(P_{i,t-1} - P_{.,t-1}) + \theta_t(\alpha_i - \alpha_.) + (\varepsilon_{i,t} - \varepsilon_{.,t}). \tag{6}$$

The second term of equation (7) represents the how much more the consumers value the individual effect of the MSA at time t relative to the rest of MSAs in the region on average.

I then take the lagged of one period, multiplied with rho and differenced it with current period.

$$\begin{aligned}(C_{i,t} - C_{.,t}) - r(C_{i,t-1} - C_{.,t-1}) \\ = \beta[(P_{i,t-1} - P_{.,t-1}) - r(P_{i,t-2} - P_{.,t-2})] + [\theta_t(\alpha_i - \alpha_.) - r\theta_{t-1}(\alpha_i - \alpha_.)] \\ + [(\varepsilon_{i,t} - \varepsilon_{.,t}) - r(\varepsilon_{i,t-1} - \varepsilon_{.,t-1})].\end{aligned}$$

Since $r = \theta_t/\theta_{t-1}$, the above can be simplified as

$$\begin{aligned}(C_{i,t} - C_{.,t}) - r(C_{i,t-1} - C_{.,t-1}) \\ = \beta((P_{i,t-1} - P_{.,t-1}) - r(P_{i,t-2} - P_{.,t-2})) + (\varepsilon_{i,t} - \varepsilon_{.,t}) \\ - r(\varepsilon_{i,t-1} - \varepsilon_{.,t-1}))\end{aligned}$$

I first find *r* by running a Prais-Winsten GLS for each zip code, with the demeaned change in price as the dependent variable and the demeaned price as the independent variable. After obtaining *r* - the correlation between two errors spaced by a lag, we then transform the dependent variables (the demeaned change in price) and independent variables (demeaned price). In other words, I assume that the correlation in attributes is consistent over the years. There are no changes in preferences back and fro over time. Since $\varepsilon_{i,t-1}$ and $\varepsilon_{i,t}$ are not correlated by construction, *we* can run a simple OLS to obtain estimates of beta. The results are shown below.

Table 9: Quasi-Demenaned Tests

Boston: Dependent Variable: Yearly Change in Price (demeaned and quasi-differenced)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned and quasi differenced)						
Lag 1	4.40e-08	6.96e-09	6.33	0.000	3.04e-08	5.77e-08
R Square	0.0177					
Adjusted R Square	0.0173					

Phoenix: Dependent Variable: Yearly Change in Price (demeaned and quasi-differenced)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned and quasi differenced)						
Lag 1	1.33e-07	2.73e-08	4.88	0.000	7.97e-08	1.87e-07

R Square	0.0156					
Adjusted R Square	0.0149					

San Diego: Dependent Variable: Yearly Change in Price (demeaned and quasi-differenced)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned and quasi differenced)						
Lag 1	2.93e-08	1.28e-08	2.28	0.023	4.12e-09	5.44e-08
R Square	0.0043					
Adjusted R Square	0.0035					

Chicago: Dependent Variable: yearly Change in Price (demeaned and quasi-differenced)

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Price (Demeaned and quasi differenced)						
Lag 1	-8.19e-08	5.03e-08	-1.63	0.104	-1.81e-07	1.68e-08
R Square	0.0032					
Adjusted R Square	0.0020					

Like the previous test, Boston, Phoenix and San Diego have positive and significant coefficients for price. The coefficients, however, are smaller than those in previous tests. Phoenix has the largest coefficient and San Diego has the smallest. Chicago, alternatively, has a negative and insignificant coefficient for lagged price. The results are consistent with the results in previous specifications. The difference in signs is likely caused by the increasingly property tax on commercial properties that led firms to relocate to the suburbs Dye et. al (2001). The results have showed that the housing market seems to be efficient for Boston Phoenix and San Diego housing markets. In particular, I expect owners of high price neighborhoods to enjoy high appreciation and owners of low price neighborhoods enjoy low user costs. The behavior of Chicago house prices is different, and I believe that it is partially caused by the flight of firms from center of MSA to the suburbs. Nevertheless, more data is needed to ascertain the claim.

Conclusion

I found evidence that support the market is quite efficient according to the user cost model framework, using the neighborhood panel set of data. The data allows us to remove possible reasons arising from productivity differences and liquidity constraints. We also found that the high priced neighborhoods are consistently ranked high among others over 10-20 years, ruling out possible bubble theories. Yet, we cannot control for possible changes in preferences and major policy shifts. We make use of the quasi-differencing method to remove possible correlation in the fixed effects, since the correlation reflects changes in preferences over time. We still find strong positive correlation between price and appreciation for Boston, San Diego and Phoenix Housing markets. This implies that the housing markets for these three cities are efficient, and buyers are unlikely to make a bargain. Nevertheless, the Chicago market appears inefficient, with price and appreciation negatively correlated. Such negative correlation is likely caused by the flight of firms.

Even after adopting quasi-differencing method, I cannot capture all possible changes that happen over time. We have to assume that the correlation in preferences over the neighborhood attribute is consistent over time. In addition, exogenous changes like property tax policies are not captured in the model either. In future research, a panel set of information on events that exogenously change preferences could be used to better control the correlation in preferences of the physical attributes of the neighborhood.

Chapter 4 Commercial and Residential Property Tax Rates: Why Do Firms Pay More

4.1: Introduction

The housing market crisis, coupled with the shortage of local tax revenue, brought new interest in how to raise revenue for local governments. For instance, Massachusetts Budget and Policy Center reported 3.1b US dollars gap in budget for the 2010 financial year for Massachusetts alone. To help shoulder the tax burden, local authorities can do the following:

1) Attract more firms to help support the fiscal expenditure by relaxing the zoning ordinances or reducing tax rates for firms.
2) Impose high tax rates on firms that cannot cheaply relocate.
3) Impose high property tax rates on both commercial and residential properties.

In the 1970s, to help relieve the tax burden of the residents, some States have allowed the municipals to classify properties for property taxation purposes. In most cases, the commercial properties are levied at a much higher rate than the residential properties. Several studies studied the incidence of property tax, the impact on the classification has on the residents and the economic development of the jurisdiction. In the spirit of Mieszkowski's approach, many studies treat property tax rates on commercial properties as tax on capital. In addition, firms are courted by the municipals because they provide employment and income to the local residents (For instance, Sonestelie (1979) and DiMasi (1988). Yet in the context of the municipals, not many residents work where they stayed. Rather than providing income and employment to the municipals, the firms generate negative externalities to the community. In this context, Fischel's approach (1975) is more suitable. Rather than treating property tax as a capital tax, we deem property tax on commercial properties as compensation to residents for the negative externalities the firms generated.

Using Massachusetts as a case study, we examine why some jurisdictions replace the uniform property tax system with the classified property tax system. In addition, we examine why some municipals place most of the burden on firms in the Massachusetts Taxpayers Report. In the next section, we review the past literature on property tax classification and property tax incidence. Section 3 covers the theory and Section 4 examines the theory empirically using tax, employment and other variables.

4.2 Literature Review

The classification of real property in the United States was first instituted in Minnesota in 1913. There are altogether 26 States and District Columbia have property tax classification systems as of now, according to information from the University of California at Davis' Institute of Governmental Affairs (See Table 9 below). The property tax classification system adopted by each State varies from each other- some have fewer classifications and others demand different valuation ratios.

Table 10: States with Property Tax Classification Systems

State	**Number of Classes**	**Different Ratios**	**Different Rates**
Alabama	7	X	
Arizona	9	X	
Colorado	3	X	
D. C.	3		X
Georgia	2	X	
Hawaii	7		X
Illinois	2 (Cook County 6)	X	

Kansas	13	X	
Kentucky	14		X (state rates)
Louisiana	5	X	
Massachusetts	4		X
Minnesota	12	X	
Mississippi	5	X	
Missouri	8	X	
Montana	11	X	
Nebraska	2	X	
New Hampshire	2		X
New York	Local option		
North Dakota	2	X	
Oklahoma	4	X	
Rhode Island	Local option		
South Carolina	11	X	
South Dakota	3		X
Tennessee	4	X	
Utah	2	X	
West Virginia	4		X
Wyoming	3	X	

Source: http://www.orange-ct.gov/govser/PROPERTY%20TAX%20OLR.htm

Some towns, even though they are required by statutes to impose a uniform rate, have a tax structure that is similar to the tax structure of cities that allowed property classification for tax purposes. For instance, Hartford credits homeowners a portion of their tax bill. The town can do this if its post-revaluation effective tax rate on residential properties exceeds 1. 5%. The credit continues for five years, including the year that the revaluation took effect.

The shift towards the property classification tax system begun in the late 1970s, and it has garnered much attention. Fischel (1975), adopting the Tiebout model approach, suggests firms maximizes their profits by scouring for the municipal with cheapest property tax rates. Since the firms generate negative externalities, they need to compensate the residents by subsidizing the residents' expenditure on public services. Firms that pollute more pay higher taxes. Fischel argued that using property tax rates as a pricing mechanism, instead of a direct payment scheme, for the negative externalities generated by firms is second best solution. The residents may be better off using the direct payments to consume private goods and allocate themselves than enjoying the better services the authorities got from the firms. For instance, elderly residents will be better off if the firms pay direct payments to the residents than to pay for the provision of schools.

Fischel (1975) further did an studied the tax rates on Bergen County, New Jersey, and found that 70% of commercial and 52% of industrial property tax levies benefitted residents by either lowering household tax payments or by increasing local spending (Fischel (1975,p 155).]. Erickson and Wollover, (1987), further provided evidence that increasing the tax base helps reduce the tax burden of households- $1000 more industrial valuation reduce tax burden by $3.1 per capita. Oakland and Testa (1995) too find similar results in their study of Philadelphia.

Sonestelie (1979) examines the long term tax incidence of classified property tax system from a different approach. He adopts Mieszkowski's view (1972) that property tax on firms is akin to capital tax. In addition, the approach postulates that firms provide local residents employment and income and does not generate any negative externalities. In his model, there are three zone; the inner ring is made up of commercial land users, the second ring is a mix between commercial and residential and the outer ring only comprise of residential properties. His model suggested that a higher tax rate will tend to shift the burden of property tax from residents of a jurisdiction to the customers of the commercial establishments and to the jurisdiction's landowners. The former will bear much of the burden if the demand for commercial real estate is inelastic and vice versa. Wheaton(1984) shows that the differences in property tax rates among jurisdictions cannot explain the differences in office rent between jurisdictions, implying that capital owners of land and commercial properties absorbing the office rent.

DiMasi (1988), using a similar approach, further examines some of the structural and welfare effects of a change from a tax system in an urban area that classifies property by use for tax assessment to one that taxes all properties at the same uniform rate. He develops a Computational General Equilibrium model that adopts a monocentric circular city where the households travel to the CBD to work. By providing suitable parameters, DiMasi measures the welfare effects if Massachusetts switch from the classified tax system to the uniform tax system. He found that the uniform tax system is more superior unless the elasticity of substitution is especially low.

Another related strand of literature examines whether classification property tax system drives business away. Four studies in the eighties found that tax rates matter. Wasylenko (1980) examines the effect of property taxes on the number of firms relocating to the suburbs of Milwaukee from 1964 to 1974, and found property tax is a significant factor. Fox (1981) studies the effect of taxes and spending on the amount of industrial land in municipalities in the Cleveland metropolitan area in 1970 and arrives at the same conclusion. Charney (1983) test the effect of local taxes on new firm locations in zip code areas in Detroit from 1970 to 1975 and found property tax rates a contributing factor to where new firm locates. McGuire (1985) examines the effect of property taxes on the location of business building permits in the Minneapolis-St. Paul metropolitan area from 1976 to 1979, and concludes they matter.

More recently, a 1999 analysis of Hartford's tax structure by the Connecticut Center for Economic Analysis (CCEA) at University of Connecticut found "the cap/surcharge structure seems to have damaged the City of Hartford economically, creating a hostile environment for businesses and apartments by distorting the tax burdens of different classes of property" (The Economic Effects of Revaluation and Tax Policy on the City of Hartford, December 1999). Similarly, The Massachusetts Taxpayers Foundation reported in 1998 that 102 communities used the classification system to shift an estimated $ 600 million of property taxes from residential to business taxpayers in FY 1998. The foundation also cited a "troubling trend" in the 200% growth of the disparity between the residential and business property tax burden between 1984 and 1998 (Unequal Burdens: Property Tax Classification in Massachusetts, November, 1998). The organization's president remarked that the higher business tax burden "adds to the already high costs of doing business in the Commonwealth, placing many companies at a further competitive disadvantage. Nevertheless, a study shows that tax classification alone could not have caused firms to relocate. Dye et al. (2001) examined the claim that classification is responsible in driving away business in Cook County. They regress the measures of business establishment in Cook County against tax rates and classification in 1993 and other factors. While they found evidence that property taxes deter firms, they could not find sufficient evidence that the classification caused the flight of firms.

In my study, I relaxed the assumption that all municipals have same population. In the real world, the size of each municipal varies a lot. For example, Boston has a population of more than half a million but Boxborough only has a population of 5000. This warrants the use of Asymmetric Nash Equilibrium framework. The Asymmetric Nash Equilibrium is been used by Bucovetsky (1991) to study tax competition among different sizes of jurisdictions. He found that residents in smaller jurisdictions pay a lower levy and enjoy better welfare than those staying in large jurisdictions. The small jurisdictions are greatly affected if firms relocate. These small jurisdictions have high transfer elasticity, making them to charge at low rates. The large jurisdictions are unlikely affected by moves caused by a raise in tax rate; as each unilateral increase in tax loses them less per capital. Wheaton (2000) also adopts the Asymmetric Nash Equilibrium in studying the underprovision of welfare. He found that large states have higher benefits level and payments because large states face lower migration elasticity. Similarly, in our context, if we implement the Asymmetric Nash Equilibrium framework, we expect large municipals have higher commercial tax rates. This is because like the above, these firms in such large states face a low elasticity to move compared to firms in small states.

4.3 Case Study- Massachusetts

Since 1780, the Massachusetts Constitution has required uniform assessment of all real properties subject to taxation. However, this provision has been rarely honored. Not only are different classes of properties within a single municipality, but properties within the same class are assessed differently as well. These disparities are illegal and precipitated a landmark court case (Bettigole vs Assessors of Springfield) in 1961. The result of the case was a judicial mandate for statewide 100 percent assessment. Some years later, in response to growing pressure to enforce statewide 100 percent assessment, politicians began to lobby for property tax classification as a way to legalize existing inter-class disparities. In Massachusetts, classification required a state constitutional amendment which had to be passed by two sessions of the State Legislature and a majority of voters in a subsequent state wide referendum. In 1968 and 1969, although the State legislature passed the referendum, only 36% of the voters supported it and the classification amendment was not passed.

In 1974, a second case concerning 100 percent assessment (Town of Sudbury vs Commissioner of Corporations and Taxation) was brought before the Massachusetts Supreme Judicial Court. The outcome of the case generated pressure to reform assessment practices and renewed interest in property classification again. In 1975 and 1976, the State Legislature passed a new classification amendment and placed it on the ballot in 1978. The classification was successful given much publicity. In 1978, the citizens of the Commonwealth of Massachusetts adopted a constitutional amendment authorizing the General Court to classify real property into as many as four separate and distinct classes and thereafter to tax such classes differently. In 1979, the General

Court adopted an act which implemented the amendment. Bloom (1979) argues that local citizens support property tax classification in Massachusetts because many local residents are not aware of the alternative measures and their implications. In addition, the continuous promotion of the advantages of property tax classification and the "good" analyses by media affect the citizens' perception of monetary self-interest and their votes.

The Classification does not raise additional dollars from the property tax, but serves to redistribute how much levy will be raised from each class. Preferential tax treatment for any class of property is not mandated, but the choice of distributing the levy burden among the various classes remains a local option. The Commissioner of Revenue supervises the implementation of property classification. After the Commissioner has determined that a city or town's assessed values represent full and fair cash values, the Assessor classifies all real property according to use. Local elected officials are then permitted to determine, within limits calculated by the Commissioner, what percentage of the tax burden is to be borne by each property class. The determination how to allocate the tax burden by class is made annually by the City Council, generally in November.

Massachusetts law provides for the implementation of the classification process through three phases: first, every city and town must value all taxable property at full and fair cash value; second, each city and town must classify every parcel of property according to use; third, each city and town which has revalued and classified may allocate its tax levy among the various classes of property. The first and second steps are mandatory. The third stage, determining whether to allocate the tax burden by class, is optional with each community.

Since 1780, the Massachusetts Constitution has required uniform assessment of all real properties subject to taxation. However, this provision has been rarely honored. Not only are different classes of properties within a single municipality, but properties within the same class are assessed differently as well. These disparities are illegal and precipitated a landmark court case

(Bettigole vs Assessors of Springfield) in 1961. The result of the case was a judicial mandate for statewide 100 percent assessment. Some years later, in response to growing pressure to enforce statewide 100 percent assessment, politicians began to lobby for property tax classification as a way to legalize existing inter-class disparities. In Massachusetts, classification required a state constitutional amendment which had to be passed by two sessions of the State Legislature and a majority of voters in a subsequent state wide referendum. In 1968 and 1969, although the State legislature passed the referendum, only 36% of the voters supported it and the classification amendment was not passed.

In 1974, a second case concerning 100 percent assessment (Town of Sudbury vs Commissioner of Corporations and Taxation) was brought before the Massachusetts Supreme Judicial Court. The outcome of the case generated pressure to reform assessment practices and renewed interest in property classification again. In 1975 and 1976, the State Legislature passed a new classification amendment and placed it on the ballot in 1978. The classification was successful given much publicity. In 1978, the citizens of the Commonwealth of Massachusetts adopted a constitutional amendment authorizing the General Court to classify real property into as many as four separate and distinct classes and thereafter to tax such classes differently. In 1979, the General Court adopted an act which implemented the amendment. Bloom (1979) argues that local citizens support property tax classification in Massachusetts because many local residents are not aware of the alternative measures and their implications. In addition, the continuous promotion of the advantages of property tax classification and the "good" analyses by media affect the citizens' perception of monetary self-interest and their votes.

The Classification does not raise additional dollars from the property tax, but serves to redistribute how much levy will be raised from each class. Preferential tax treatment for any class of property is not mandated, but the choice of distributing the levy burden among the various classes remains a local option. The Commissioner of Revenue supervises the implementation of property classification. After the Commissioner has determined that a city or town's assessed values represent full and fair cash values, the Assessor classifies all real property according to use. Local elected officials are then permitted to determine, within limits calculated by the Commissioner, what percentage of the tax burden is to be borne by each property class. The determination how to allocate the tax burden by class is made annually by the City Council, generally in November.

Massachusetts law provides for the implementation of the classification process through three phases: first, every city and town must value all taxable property at full and fair cash value; second, each city and town must classify every parcel of property according to use; third, each city and town which has revalued and classified may allocate its tax levy among the various classes of property. The first and second steps are mandatory. The third stage, determining whether to allocate the tax burden by class, is optional with each community.

According to the Massachusetts Tax Foundation, 102 communities used the classification law to shift property taxes from residential to business taxpayers in the financial year 1998. Although the classifying communities comprise fewer than a third of the Commonwealth's 351 cities and towns, they account for almost two thirds of the state's population and encompass most of the state's urban centers and developed suburbs, including all of the larger cities and many towns with significant amounts of business property. The shift to business property taxpayers in the classifying communities totaled over $600 million in 1998, almost one-third of the business property tax bill. In Boston, the largest classifying community, the rate of $38.45 paid by business taxpayers was more than three times the rate paid by residential taxpayers. In the 101 other classifying communities, the average tax rate for business of about $28 was almost double the rate for residences. In contrast, in the 239 communities that did not classify, businesses and residential taxpayers paid the same tax rates, averaging $14.60 per $1,000 of assessed value over all such communities. It is also interesting to note that communities that switched from universal tax regime to multiple tax regime do not switch back.

4.4 Theory

Suppose there Q number of jurisdiction; each has population n_i and houses l_i firms per capita. Each jurisdiction has a limit on the number of firms that they can build, k. The jurisdiction imposes a tax $t_2$on commercial properties and t_1 on residential properties. The jurisdiction adopts uniform tax rate system if the tax rates are equal $t_1 = t_2$, and property tax classification system if otherwise. Each resident earns income Y. The total number of firms and population for all the municipals is N and L respectively.

The residents' utility function is as follow:

$$U = (k - l)^{\alpha} S^{b} X^{1-b}, \tag{1}$$

where X is the amount of private goods consumed, S denotes the amount of public services demanded. ◘ and b are positive constants.

In addition, the income constraint for the resident is such that

$$Y = t_1 + X, \tag{2}$$

assuming that the price of private good is 1.

The government budget constraint is

$$S = t_1 + t_2 l_i, \tag{3}$$

Assumption 1:

This implies l cannot large- nearly equals to one- and $t_2 l'$ cannot be small- near to 0- at the same time.

$$l_i + t_2 l'_i < 0 \tag{4}$$

The assumption is required to ensure residents do not tax firms at an infinitely high tax rate. Without this assumption, the utility function with respect to tax is an upward sloping curve. The equality in (2) will ensure the utility function forms a concave function with respect to t_2.

Assumption 2:

We impose the condition that the number of firms that a municipal hold is no larger than the term on the right hand side of equation (3).

$$l_i \leq \frac{k}{1+\alpha} \tag{5}$$

In other words, we assume that the full capacity of firms a municipal can hold is not reached at any point in time. As seen later, the residents' tax function reaches an asymptote when $l_i = \frac{k}{1+\alpha}$. The residents solve the following optimization problem:

$$\max_{t_1,t_2}(k-l)^{\alpha}(t_1+t_2 l)^{b}\,(Y-t_1)^{1-b} \tag{6}$$

First order conditions give us the following

t_1: $$t_1 + t_2 l = b(Y + t_2 l) \tag{7}$$

t_2: $$-\frac{\alpha U}{k-l}l' + \frac{bU}{t_1+lt_2}l + \frac{bU}{t_1+lt_2}t_2 l' = 0 \qquad (8)$$

From the first order conditions in (3), we found that we need the following condition

$$l + t_2 l' < 0 \qquad (9)$$

Since $l' < 0$, the first derivative of Utility with respect to t_2 will be positive if assumption 1 is not met. This implies t_2 will go to infinity for residents to maximize their utility.

From (6), we simplify to get the following expression,

$$\frac{b(l+t_2 l')}{(t_1+lt_2)} = \frac{\alpha l'}{k-l} \qquad (10)$$

Substituting (5) into (8) gives us

$$\frac{(l+t_2 l')}{(Y+lt_2)} = \frac{\alpha l'}{k-l} \qquad (11)$$

Solving the above gives us

$$t_2 = \frac{\alpha Y}{(k-l)-\alpha l} - \frac{(k-l)l}{l'[(k-l)-\alpha l]} \qquad (12)$$

The firms are adverse to tax rates and will seek to locate in municipals that offer low tax rates. In addition, the firms are attracted by the population in the municipal as the residents provide labor and demand for finished goods. We specify the firms' decision to locate as follow:

$$\widetilde{l_i} = L\left[\frac{n_i e^{\beta t_{2i}+z_i}}{\sum_j n_j e^{\beta t_{2j}+z_j}}\right], \qquad (13)$$

where $l_i = \frac{\tilde{l_i}}{n_i}$.

Taking the first derivatives of $\tilde{l_i}$with respect to t_2 gives us

$$\tilde{l_i}' = \beta \tilde{l_i}\left(1 - \frac{\tilde{l_i}}{L}\right). \tag{14}$$

Hence

$$l_i' = \beta l_i\left(1 - \frac{n_i l_i}{L}\right) \tag{15}$$

The equilibrium solution using (7) and (8) cannot be readily observed and the relationship between number of firms and commercial property tax rates are uncertain.

We set the following parameters:

Parameters	Initial Values
b	0.2
α	0.3
k	5
β	-2
Y	1
L	1
n_i/N	0.003

N	1
Q (total number of municipals)	351

We assume that the population of each municipal is the same, *n*. Then

$$l_i = \frac{L}{n_i}\left[\frac{n_i e^{\beta t_{2i}+z_i}}{\sum_j n_j e^{\beta t_{2j}+z_j}}\right] = \frac{L}{n}\left(\frac{e^{\beta t_{2i}+z_i}}{\sum_j e^{\beta t_{2j}+z_j}}\right) \quad (16)$$

In the beginning, the tax rates for all N municipals are the same. Hence, the number of firms per capita in the municipal is $\frac{L}{351n}$.

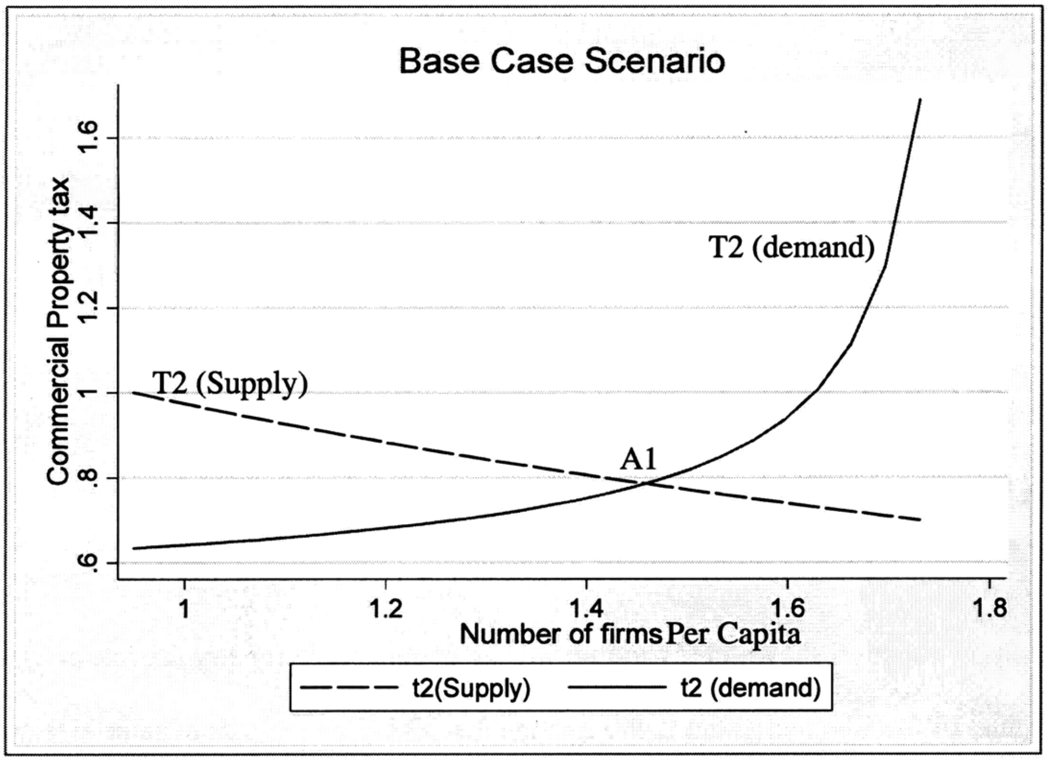
Using the supply function shown in (5) and setting the commercial property tax rate at 2, we solve for l, the number of firms in municipal i. We assume that all the municipals of same size and have the same tax rates initially. We let the tax rates increase over time and observe how the number of firms change. Similarly, using the number of firms derived from the supply function, we derive the tax rates using the demand function in (4).

We plot the residents' tax function and firms' location function in Figure 13. The equilibrium number of firms is 1.47 firms per capita and the equilibrium property tax rate is 0.76. As expected, we found that the tax rate derived from the demand function reaches an asymptote where

$$l = \frac{k}{1+\alpha} \quad . \quad (17)$$

When l is less than $\frac{k}{1+\alpha}$, the relationship is different.

Figure 13: Based on Initial Conditions

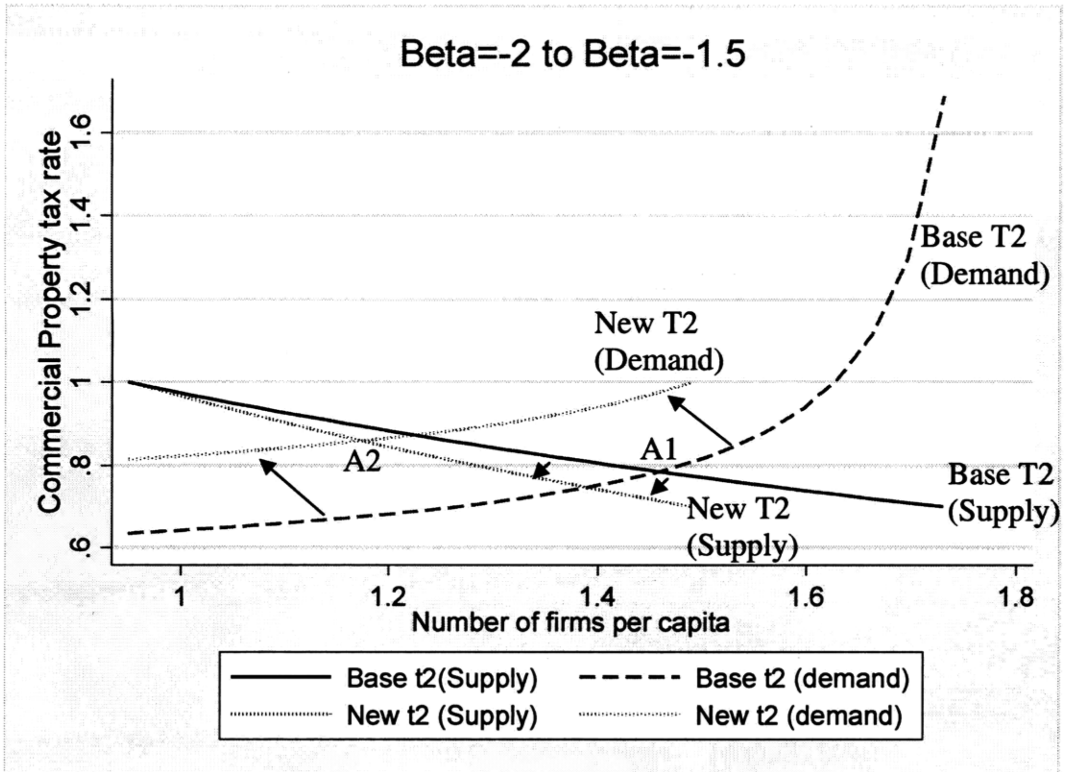


Note: A1 is the equilibrium point, where number of firms is 1.47 and property tax rate is 0.76. t2(supply) depicts the number of firms in the municipal given that the tax rates are known. t2 (demand) describes the property tax rate given the number of firms are known.

First Scenario: Beta=-1.5 vs Beta=-2.0

In our second scenario, we change the beta from -2.0 to -1.5. The changes in demand tax functions and supply tax functions are shown in chart 2. If firms are less responsive to changes in tax rate, the whole demand schedule will shift upwards. In addition, the gradient for the firms location function becomes less steep. The original equilibrium, which is at A1, shifts to A2. The new equilibrium commercial property tax is higher and the new equilibrium number of firms per capita is lower.

Figure 14

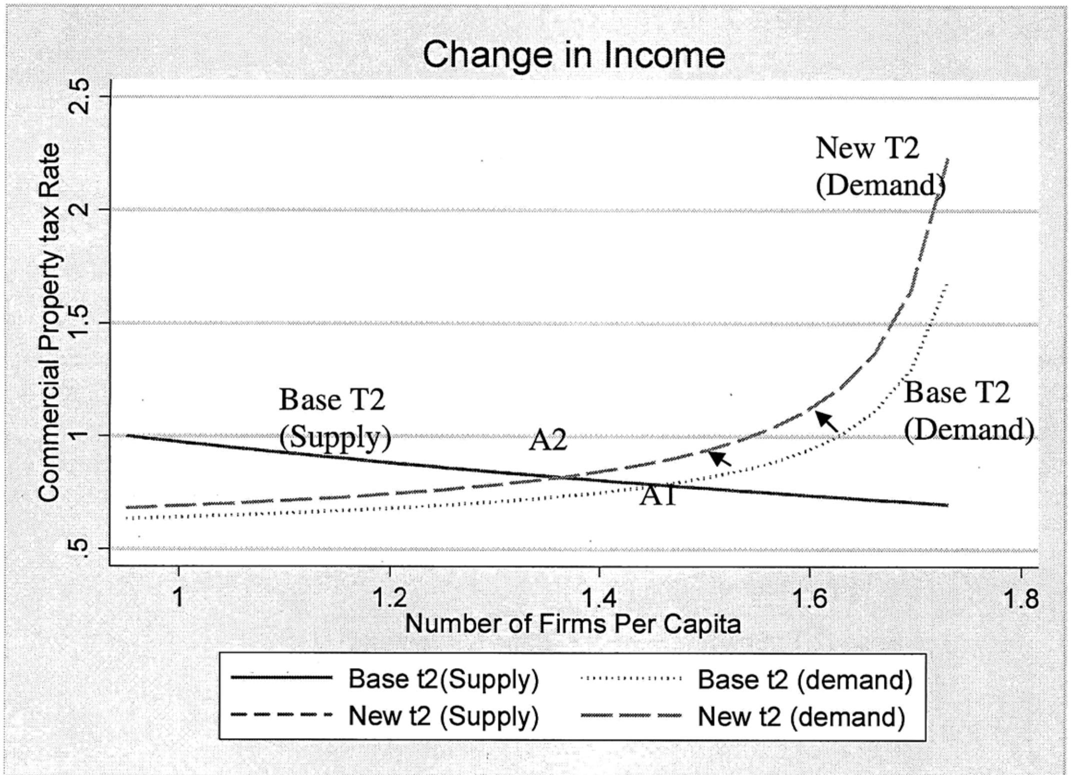


Note: The pre t2 are the tax rates with beta=-2, while t2 are tax rates derived with beta =-1.5.

Using our initial conditions, I found that the commercial property tax rates will fall as number of firms increase (See Figure 14). A2 is the new equilibrium point and A1 is the old equilibrium.

The next question is to find how the equilibrium tax rates and number of firms will change when income changes. We let income increases over time and observe how the equilibrium changes. As shown below in Figure 15, we find that the whole demand function shifts upwards. The equilibrium tax rate becomes higher and the equilibrium number of firms falls; the equilibrium position shifted from B1 to B2.

Figure 15: Changes in Income



Note: The pre t2 are the tax rates with income=1, while t2 are tax rates derived income=1.5. The chart shows that the demand function will shift upwards. The equilibrium position shifted from A1 to A2, signifying higher commercial property tax rates and fewer firms.

Given that high income municipals place a high value to environment, they have to rely less on firms for external revenue. Hence, they will increase commercial tax rates to "drive" the firms away. In addition, the increase in come means higher demand of public services. Residential tax rates will too have to increase as the residents consume more services.

Proposition 1:

Municipals with greater number of firms will impose a high commercial property tax rate, as long as assumptions (1) and (2) are met. From Chart 1, we see that commercial property tax rate is upwards sloping until assumption (2) is violated.

Proposition 2:

High income earning municipals that are adverse to firms will impose high commercial property tax rates, as they try to not to rely on firms to produce services. The number of firms will fall. This is depicted in Chart 3. The demand function increases, while the supply function is unaffected by changes in income.

Corollary 2a:

From Proposition 1, we know that municipals with high-income households will impose a higher commercial tax rate that that of municipals with low-income households. The increase in tax rates will discourage firms to locate in the municipal. We know that the residential property tax is provided by the following function,

$$t_1 + t_2 l(1-b) = bY \ .$$

Differentiating the above with respect to Y, we get

$$\frac{\partial t_1}{\partial Y} = b - \frac{\partial t_2}{\partial Y}(1-b)\left(l - t_2 \frac{\partial l}{\partial t_1}\right). \qquad (11)$$

Using the assumption in (7) as seen below and Proposition (2)

$$l + t_2 l' < 0, \qquad (7)$$

$\frac{\partial t_1}{\partial Y}$ must be more than 0. Hence, high income municipals will impose a high residential property tax rate, and the difference in rates for rich municipals is smaller than that for poor municipals.

4.5 Empirical test

I assembled the property tax rates, levy from the yearly publications of Massachusetts Taxpayers Association (MTA) and Department of Revenue of Massachusetts. I use the effective tax rates- rates based on Equalized Valuations by the department. Note that the actual tax rates are usually higher than the effective tax rates before 1981 because the valuations were usually lower than Equalized Valuations. Given that the Department of Revenue only provides information from 1980 onwards, I have to depend on the publications by MTA. Some of the early publications like the tax rates on 1974 was missing in the library and cannot be accounted. In addition, for some small municipals, information is not available and they are omitted from study. The series provided by Department of Revenue have missing information on the levies collected. I used the rates from MTA and the equalized valuations that stretched back to 1970 to derive the levies collected. I compared the levies collected in the MTA and those I derive and the differences are not large.

The wage and employment data is obtained from are obtained from the Department of Labor and Workforce of Massachusetts. It contains employment and wages in establishments in each of the 351 cities and towns in the Commonwealth. In other words, the wage does not represent the income of residents but what offered by firms in the municipals. In addition, the wage I took is the average wage for all establishments, including non-profit organizations and public sector. It is noted that no-profit organizations and public sector institutions do not pay property taxes. The results will be better reflect the interactions if we use private sector employment and wages, or the employment and wages for the detailed industries. Nevertheless, most of the information is subject to certain restrictions to protect the confidentiality of all data reported by individual employers. Summary level dad is confidential if there are less than three reporting units in total, or if with three or more units, one unit accounts for 80% or more of the total.

Using the published numbers for establishments is problematic- the number of establishments includes subsidiaries of firms and other related companies and companies sometimes use the establishment for tax rebates purposes. Hence, we use the employment base as a proxy for the number of firms.

4.1 Presence of firms on property tax rates

Our tests comprise of several fixed effect tests and Difference-In-Difference tests. In our first set of empirical tests, we wanted to know whether municipals with more firms will vote for higher commercial tax rates or not. This is our test for the first proposition. We first study the difference in tax rates using the panel of observations. In addition, we use a probit test to examine the characteristics of the municipals that selected the classified property tax system. It is interesting to note that the municipals that preferred the multi-tax regime never revert back to the uniform tax system.

For our panel tests, our dependent variable is the difference between commercial tax rate and residential tax rate. The dependent variables are the employment of the jurisdiction lagged one period and the difference in rates lagged on period. We ran a total of 6 tests. For the first three tests, we only use the post 81 dat. This is because the difference in rates is 0 before 1981, and this may generate bias in the results. In the latter three tests, we utilize the whole sample to test the relationship. Besides splitting the sample by time period, we also run different specification of the tests. We had tests with fixed effects for municipal and year, fixed effects for year only and no fixed effects. When the municipal fixed effects are included, the effect of employment per capita on property tax rates differences is reduced. This is because the differences in employment may be captured by the municipal fixed effects.

The results are presented in table 4.1. The results are very consistent across the different specifications we used. We found that the difference in tax rates are persistent- a high difference in property tax rates in previous period is likely to lead to an even higher difference in current period. The total employment per capita in the previous period has a positive effect on the current difference in rates. The results show support to our theory; municipals with large number of firms tend to impose a high commercial tax rate. Looking at the results for the post 81 data with year fixed effects only specification, we found that the difference in property tax rates will drop by 7.2% if the employment per capita drops by 100.

<u>Table 11: Panel Fixed Effects Test</u>

Dependent Variable: Difference in Property tax rates (Commercial tax rate – Residential tax rates)

	Test with year and municipal fixed effects using post 81data	Test with year Fixed Effect using post 81data	Test without fixed Effects using post 81data	Test with year and municipal fixed effects using whole sample	Test with year Fixed Effect using whole sample	Test without year fixed Effect using whole sample
	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Difference in Property tax rates lagged	0.7088138 (108.76)	0.95581 (274.41)	0.9516873 (274.11)	.8609834 (193.79)	0.9578126 (316.23)	0.9557697 (320.32)
Employment per capita Lag 1	0.361157 (2.05)	0 .713308 (10.60)	0.6717275 (9.93)	.5214304 (4.37)	0.6114604 (11.16)	0.6157899 (11.33)
Constant	-0.190418 (-0.70)	0.4843385 (6.04)	-0.0107785 (-0.42)	-.1247075 (-0.55)	- 0.0828233	- 0.0300367

					(-1.15)	(-1.57)
R-square	0.9247	0.9085	0.9069	0.9196	0.9135	0.9118

Alternatively, we ran a probit test. The dependent variable is the choice of whether the municipal has selected multiple-tax regime or not as in 1990. Included in the independent variables are the employment per capita in 1990 and the average household income of each municipal in 1990. The results are reported in Table 4.2.

Table 12: Dependent Variable: Choice of Tax Regime (1-dual tax regime,0- universal tax regime)

	Coefficient	Standard deviation	T stat	Pvalue	95% confidential level	
Median Household Income 1989	-0.000018	6.90e-06	-2.54	0.011	-.000031	-3.99e-06
Employment per capita 1990	3.002607	.5497244	5.46	0.000	1.925167	4.080047
Constant	-0.160275	.4128947	-0.39	0.698	-.9695344	.6489831

We found that municipals with high median household income are less likely to choose the dual tax regime, and municipals with high employment per capita are more likely to choose the dual tax regime. The results are consistent when we shift the cross-section to other years. The probit test provides much support to our results. First, municipals that have higher income residents place more importance (value) to the environment. Hence, they have to charge a high residential property tax rate as they want to rely less on the firms, while enjoying a high level of services. Low income municipals, alternatively, have to rely on firms to provide fiscal benefits. The only way for the residents to enjoy such fiscal benefits is to adopt the property tax classification system. Second, the coefficient for employment is consistent to the results we had in the panel tests. When municipals select the dual tax system, the municipals will charge the commercial properties at a rate higher than the residential property tax rate. Hence, with larger employment per capita, there will be more firms, attracting a higher commercial property tax rate.

4.2 Commercial Property tax rates on employment per capita

The second set of tests examines whether the firms' location decision function in our model is justifiable. Dye et al. (2001) argues that changes in the classification tax system have insignificant effect on economic development. With a larger number of cross-section observations and longer time series, we conduct the tests that are similar to the first set. The dependent variable is the lagged change in employment and the independent variables are the lagged changes in employment and lagged difference in tax rates. Note that the fixed effects of municipals removed a lot of explanatory power.

The panel results show that the change in employment per capita falls after the difference in property tax rates widened. Using level variables, we get the same results when we accounted for the year and municipal fixed effects. The results support a downward sloping firm location function that contributes to our propositions.

Table 13: Dependent Variable: Change in Employment Per Capita

	Fixed Effects (year) using post 1981 data	Fixed Effects (year and municipal) using post 1981 data	Pooled Regression (No fixed effects) using post 1981 data	Fixed Effects (year) using whole sample	Fixed Effects (year and municipal) using whole sample	Pooled Regression (No fixed effects) using whole sample
	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Change in Employment per capita lagged 1 period	-.0704006 (-7.32)	-.0929076 (-9.52)	-.0440589 (-4.68)	-.054094 -(5.78)	-.0786629 (-8.31)	-.0283833 (-3.04)
Difference in property tax rates lagged	-.0002331 (-2.63)	-.0006042 (-3.10)	-.0002005 (-2.24)	-.0002302 (-2.37)	-.001067 (-6.52)	-.0005162 (-5.31)
Constant	.0060902 (2.76)	-.0039234 (-0.49)	.0032615 (6.87)	0048096 (2.00)	-.0035089 (-0.44)	.0069929 (14.98)
R-square	0.0475	0.0784	0.0028	0.0744	0.1029	0.0032

Note: Values in parenthesis are t-values.

Table 14:Dependent Variable: Employment per Capita

	Fixed Effects (year)using post 1981 data	Fixed Effects (year and municipal)using using post 1981 data	Pooled Regression using using post 1981 data	Fixed Effects (year)using whole sample	Fixed Effects (year and municipal)using whole sample	Pooled Regression using whole sample
	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Employment per capita lagged 1 period	.9896879 (534.97)	.8465181 (168.65)	.9884476 (525.55)	.9944602 (540.22)	.8974636 (224.21)	.990705 (529.13)
Difference in property tax rates lagged	-.0000159 (-0.17)	-.0005534 (-2.97)	.0000441 (0.46)	-.0001096 (-1.07)	-.0006457 (-4.32)	-.0002709 (-2.63)
Constant	.0052566 (2.38)	.0389616 (4.99)	.0063364 (8.82)	-.0016737 (-0.69)	.0314663 (4.13)	.0088995 (13.43)

R-square	0.9732	0.9763	0.9720	0.9670	0.9695	0.9646

Note: Values in parenthesis are t-values

I further conduct Difference-in-Difference test, examining how the employment per capita changes after the law allowed the municipals to vote the tax system they want. Before the change in tax regime is allowed, "treated" municipals- those that select multiple tax regimes after 1981- chose has a growth rate of 7.2%. "Untreated" Municipals- municipals that do not opt for dual tax regimes after 1981- have a growth rate of 3.7%. After 1981, those municipals that selected the multiple tax regimes have a growth rate of 3.7%. Alternatively, municipals that continued the existing tax regime have a growth rate of 4.7%. The results provide strong evidence that the increase in commercial tax rates reduces the employment growth rate and hence, the growth rate in number of firms.

Table 15

Window period- from 1970-1978 and 1982-1991

Dependent variable: Absolute Change in employment per capita

	Coefficient	Standard deviation	T stat	Pvalue	95% confidential level	
Post Treatment	-.046132	.0155838	-2.96	0.003	-.0767292	-.0155348
Treatment	.0353707	.0110161	3.21	0.001	.0137417	.0569997
Post	0.0104364	.0084721	1.23	0.218	-.0061977	.0270705
Constant	.0377864	.0059846	6.31	0.000	.0260363	.0495365
Rsquare	0.0163					

4.6 Conclusion

Property tax classification system provides us an opportunity to study why municipals want to impose high property tax rates on commercial properties. The past studies approach the problem using Mieszkowski's approach, arguing that the tax rate differentials are passed to consumers from other jurisdictions or landowners. Even then, the theory could not explain why commercial properties are charged a higher tax rate than residential tax rate across the board. A second group of studies postulate that the fiscal benefits entice the population to impose a high tax rate on commercial properties. Nevertheless, that depends on the elasticity of firms' decision with respect to tax.

Instead of viewing the commercial property tax rate as a capital tax, we consider the commercial property tax as a payment to compensate for the negative externalities. We adopt this view because it is unlikely people stay in the same jurisdiction as they work. In addition, we model the firms' location decision using a logistic function. Using the above assumptions, we create a model to explain the phenomenon. We found that households that value the environment more will rely on themselves to provide better services. They will tax commercial firms more- not to increase their fiscal benefits but to enjoy better environment. Alternatively, for poor municipals who has less regard to the negative externalities the firm produces, they will try to attract firms to increase their fiscal savings. Hence, the difference in tax rates will fall. Our empirical tests show that demand for less reliance on firms is a strong explanatory variable on the differences in tax rates. In addition, we further show evidence that firms is adverse to high commercial property tax rates.

Chapter 5 Conclusion

The thesis aims to address three questions:

a) what is the theory that drives the co-movement between price and sales in housing market?

b) Is the housing market efficient and whether bargains are available or not?

c) Why do some municipals charge commercial properties at a rate higher than that for residential properties?

Given that the time series for data on housing and real estate markets are short, the use of panel data analysis techniques enable us to answer the above questions and develop new theories in the field. For the first question, we found a gap in the existing literature on co-movements of price and sales in housing market. Past studies fail to capture the inter-tenure transitions that represent 60% of the transactions. Further examination of the data shows that such inter-tenure transitions create a negative schedule between sales and price via inventory. The use of the panel data analysis further provides strong empirical support to our model. The finding has important policy implications. It shows that by increasing the affordability of rent-to-own households, the housing market can recover.

The second question has garnered lots of interest since 1980s, when housing market was booming. Under the user cost model, locations that at priced highly (lowly) will enjoy high (low) appreciation if the market is efficient. While many empirical studies have been done to check the efficiency of the market, my study is one of the first few that made use a panel data at zip code level within the Metropolitan Statistical Area. The database enabled us to leave out other possible reasons that could cause the same phenomenon. We found that the housing markets in Boston, San Diego, and Phoenix efficient, prompting us to conclude that "bargains" are unlikely be found. Alternatively, the results show Chicago housing market is inefficient, implying opportunities of arbitrage. Nevertheless, I believe the result in Chicago housing market could be resulted from the flight of firms from the center to the suburbs. Much research needs to be done to ascertain whether this is true.

The last question has drawn much attention across States as they look at different ways to increase funding for local services. One of the ways is to tax commercial property and residential properties at different rates. Studies have been fast to point out that such system will cause commercial property tax rate rise and residential tax rate fall. We took an extra step to find why certain jurisdictions will impose a high property tax rate for commercial properties and low tax rate for residential properties. Unlike past studies, we adopt Fischel (1975) framework- treating commercial property tax as compensation for the externalities the firms create. This is because few work in the same jurisdiction they live. I develop a model of firm location choice and residents' choice of tax rates. I found that rich neighborhoods tend to impose high commercial tax rates and high residential tax rates. This is because they wanted to deter the firms from locating in their neighborhood and reduce the reliance the revenue from them. In addition, the model suggests jurisdictions with large number of firms are likely to impose higher commercial property tax rates. From the panel database I assembled, the results strongly support the model. The empirical results and theory has important policy implications. First, the poor municipals need the firms to reduce their tax burden. Hence, their tax rates on commercial properties will be lower than those imposed by rich municipals. Second, DiMasi (1988) has shown that the multiple tax rate system is likely to generate less welfare. In this case, it seems that the rates in rich jurisdictions are unlikely to lose much welfare when they switch to property classification tax system because their rates do not differ much. Those staying in the poor jurisdictions, however, may suffer greater loss of welfare when they opt for the property tax classification system.

REFERENCES

Andrew, M. and Meen, G., 2003, "House price appreciation, transactions and structural change in the British housing market: A macroeconomic perspective", *Real Estate Economics*, 31, 99-116.

Baily, M.J., R.F. Muth, and H.O. Nourse, 1963, "A Regression Method for Real Estate Price Index Construction", *Journal of the American Statistical Association*, 58, 933-942.

Berkovec, J.A. and Goodman, J.L. Jr., 1996, "Turnover as a Measure of Demand for Existing Homes", *Real Estate Economics*, 24 (4), 421-440.

Bloom, H.S. (1979). "Public Choice and Private Interest: Explaining the Vote for Property Tax Classification in Massachusetts," National Tax Journal, 32(4), 527-534.

Brunnermeier, Markus and Christian Julliard. 2008. "Money Illusion and Housing Frenzies." Review of Financial Studies, 21(1):135-180

Bucovetsky, S. 1991. "Assymetric tax competition," Journal of Urban Economics30, 167-181.

Cannon, S., N. Miller, G. Pandher "Risk and Return in the US Housing Market: A Cross-Section Asset-Pricing Approach", Real Estate Economics, 34, 4 (2006) 519-552.

Capozza, D., Hendershott, P., and Charlotte Mack, 2004, "An Anatomy of Price Dynamics in Illiquid Markets: Analysis and Evidence from Local Housing Markets," *Real Estate Economics*, 32:1, 1-21.

Capozza, Dennis, Hendershott, Patric, and Charlotte Mack. "An Anatomy of Price Dynamics in Illiquid Markets: Analysis and Evidence from Local Housing Markets," Real Estate Economics, 32 (2004) 1-21.

Capozza, Dennis, Hendershott, Patric, Mayer,C. and Charlotte Mack. (2002), Determinants of Real House Price Dynamics. NBER Working Paper No. W9262.

Case, B., Pollakowski, H., and Susan M. Wachter, 1991. "On Choosing among Housing Price Index Methodologies," *AREUEA Journal*, 19, 286-307.

Case, Karl and Robert Shiller. (1989). "The Efficiency of Market for Single-Family Homes," American Economic Review, Vol79, no.1, 125-137.

Case, Karl E. and Robert J. Shiller. 1988. "The Behavior of Home Prices in Boom and Post-Boom Markets", New England Economic Review, November, pp. 29-46.

Chan, S., 2001, "Spatial Lock-in: Do Falling House Prices Constrain Residential

Mobility?", *Journal of Urban Economics*, 49, 567-587.

Charney, A. (1980) Intraurban Manufacturing Location Decisions and Local Tax Differentials, Journal of Urban Economics, 14, 184–205.

Connecticut Center for Economic Analysis. (1999). The Economic Effects of Revaluation and Tax Policy on The City of Hart Ford

DiMasi, J. "Property tax classification and Welfare in Urban Areas: A General Equilibrium Computational Approach" Journal of Urban Economics, 23, 131-145.

DiPasquale, D. and Wheaton, W.C. Urban Economics and Real Estate Markets, Prentice Hall,New Jersey (1996).

Dye F. R., McGuire, T.J. and Merriman D.F. (1999). "The Impact of Property Tax and Property Tax Classification on Business in Chicago Metropolitan Area," Lincoln Institute of Land Policy Working Paper.

Engelhardt, G. V., 2003, "Nominal loss aversion, housing equity constraints, and household mobility: evidence from the United States." *Journal of Urban Economics*, 53 (1), 171-195.

Erickson, A. R. and Wollover, David. R. 1987. "Local tax burdens and supply of business sites in suburban municipalities" Journal of Regional Science, 27(1), 25-37.

Fischel, William. A. 1975. "Fischel and Environmental Considerations in the Locations of Firms in Suburban Communities," in Fiscal Zoning and Land Use Controls. Eds: Edwin S. Mills and Wallace E.Oates. Lexington, MA. Heath-Lexington Books, 1975, pp.119-173.

Fox, W. (1981) Fiscal Differentials and Industrial Location: Some Empirical Evidence, Urban Studies, 18, 105–11.

Fox, W. 1978. "Local Taxes and Industrial Location." Public Finance Quarterly. 6(Jan). 93-114.

Genesove, D. and C. Mayer, 2001, "Loss aversion and seller behavior: Evidence from the housing market." *Quarterly Journal of Economics*, 116 (4), 1233-1260.

Greene, W.H., 2003, *Econometric Analysis,* (5th Edition), Prentice Hall, New Jersey.

Grossman, Sandford, and Guy Laroque. "Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods", Econometrica, 58 (1990) 25-51.

Gyourko, J., Mayer, C. and Sinai,T. (2006) "Superstar Cities" National Bureau of Economic Research Working Paper No. 12355. June 2006.

Harding, J., Rosenthal, S., and C.F.Sirmans, 2007, "Depreciation of Housing Capital, maintenance, and house price inflation…", *Journal of Urban Economics*, 61,2, 567-587.

Hendershott, Patric and Joel Slemrod. 1983. "Taxes and the User Cost of Capital for Owner-Occupied Housing." Journal of the American Real Estate and Urban Economics Association. Winter, 10:4, pp. 375–93.

Himmelberg, C., C. Mayer, and T. Sinai (2005). Assessing high house prices: Bubbles, fundamentals and misperceptions. Journal of Economic Perspectives, 19(4), 67-92.

Holtz-Eakin, D., Newey, W., and Rosen, S.H, 1988, "Estimating Vector Autoregressions with Panel Data", *Econometrica*, Vol. 56, No. 6, 1371-1395.

Kindleberger, Charles. 1987. "Bubbles." In The New Palgrave: A Dictionary of Economics, edited by John Eatwell, Murray Milgate, and Peter Newman. New York: Stockton Press.

Kyung So Im, M.H. Pesaran, Y. Shin, 2002, "Testing for Unit Roots in Heterogeneous Panels", Cambridge University, Department of Economics

Lamont, O. and Stein, J., 1999, "Leverage and House Price Dynamics in U.S. Cities." *RAND Journal of Economics*, 30, 498-514.

Leung, C.K.Y., Lau, G.C.K. and Leong, Y.C.F., 2002, "Testing Alternative Theories of the Property Price-Trading Volume Correlation." *The Journal of Real Estate Research*, 23 (3), 253-263.

Levin, Andrew and Chien-Fu Lin, 1993, "Unit Root Tests in Panel Data: new results." *Discussion Paper No. 93-56*, Department of Economics, University of California at San Diego.

Massachusetts Tax Payers Foundation, 1998. Unequal Burdens: Property Tax Classification in Massachusetts. (Massachusetts Taxpayers Foundation, Boston).

Mcdonald, J.F. (1993). "Property tax classification and Welfare in Urban Areas: A General Equilibrium Computational Approach" Journal of Urban Economics, 23, 131-145.

McGuire, T. (1985) Are Local Property Taxes Important in the Intrametropolitan Location Decisions of Firms? An Empirical Analysis of the Minneapolis-St. Paul Metropolitan Area, Journal of Urban Economics, 18, 226–34.

Meese, Richard, and nancy Wallace. 1993. Testing the Present Value Relation for Housing Prices: Should I Leave My House is San Francisco? Journal of Urban Economics 35(3), 245-66.

Nickell, S., 1981, "Biases in Dynamic Models with Fixed Effects," *Econometrica*, Vol. 49, No. 6. pp. 1417-1426.

Nierwerburgh, S.V and Weill, P.O. Why has house Price Dispersion Gone UP? Working Paper.

Oakland, William H. and William A. Testa. 1996. "State-Local Business Taxation and the Benefits Principle." Economic Perspectives 20.1: 2-19.

Ortalo-Magne, Francois and Sven Rady.. 2001. "Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints," Discussion Paper 470, CESIfo.

Ortalo-Magne,F. and S. Rady (1999). Boom in, bust out: Young households and the housing price cycle. European Economic Review 43(4), 755-766.

Per Lundborg, and Per Skedinger, 1999, "Transaction Taxes in a Search Model of the Housing Market", *Journal of Urban Economics,* 45, 2, 385-399.

Pissarides, Christopher, 2000, *Equilibrium Unemployment Theory*, 2nd edition, MIT Press, Cambridge, Mass.

Poterba, J.M. (1984). Tax Subsidies to owner occupied housing: An asset market approach. The Quarterly Journal of Economics 99(4), 729-52.

Smith, Margaret Hwang and Gary Smith. "Bubble, Bubble, Where's the Housing Bubble?", Brooking Papers.

Sonstelie,J. "The incidence of a classified property tax." Journal of Public Economics, 12, 1979, 75-85.

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

Stein, Jeremy. 1995. "Prices and Trading Volume in the Housing Market: A Model with Downpayment Effects," Quarterly Journal of Economics 110, May, pp. 379-406.

W.C. Wheaton, The incidence of inter-jurisdictional differences in commercial property taxes. National Tax Journal. **37** (1984), pp. 515–528.

Wasylenko, M. 1980." Evidence of Fiscal Differentials and Intrametropolitan Firm Relocation," Land Economics, 556 (13), 339-349.

Wheaton, W.C., 1990, "Vacancy, Search, and Prices in a Housing Market Matching Model," *Journal of Political Economy*, 98, 1270-1292

Wheaton, W.C., and Gleb Nechayev 2008, "The 1998-2005 Housing 'Bubble' and the Current 'Correction': What's different this time?" *Journal of Real Estate Research*, 30,1, 1-26.

Wilson, J.D., 1986, A theory of interregional tax competition, Journal of Urban

Economics, 19, 296-315.

Zodrow, G.R. and P. Mieszkowski, 1986, “Pigou, Tibout, property taxation, and the underprovision of local public goods”, Journal of Urban Economics, 19, 356-370.

APPENDIX 2.1: Sales, Price Panel Statistics

Market Code	Market	Average GRRHPI (%)	Average GREMP (%)	Average SFSALES RATE	Average GRSALES RATE (%)
1	Allentown*	2.03	1.10	4.55	4.25
2	Akron	1.41	1.28	4.79	4.96
3	Albuquerque	0.59	2.79	5.86	7.82
4	Atlanta	1.22	3.18	4.31	5.47
5	Austin	0.65	4.23	4.36	4.86
6	Bakersfield*	0.68	1.91	5.40	3.53
7	Baltimore*	2.54	1.38	3.55	4.27
8	Baton Rouge	-0.73	1.77	3.73	5.26
9	Beaumont	-1.03	0.20	2.75	4.76
10	Bellingham*	2.81	3.68	3.71	8.74
11	Birmingham	1.28	1.61	4.02	5.53
12	Boulder	2.43	2.54	5.23	3.45
13	Boise City	0.76	3.93	5.23	6.88
14	Boston MA*	5.02	0.95	2.68	4.12
15	Buffalo	1.18	0.71	3.79	2.71
16	Canton	1.02	0.79	4.20	4.07
17	Chicago IL	2.54	1.29	4.02	6.38
18	Charleston	1.22	2.74	3.34	6.89
19	Charlotte	1.10	3.02	3.68	5.56
20	Cincinnati	1.09	1.91	4.87	4.49
21	Cleveland	1.37	0.77	3.90	4.79
22	Columbus	1.19	2.15	5.66	4.61
23	Corpus Christi	-1.15	0.71	3.42	3.88
24	Columbia	0.80	2.24	3.22	5.99
25	Colorado Springs	1.20	3.37	5.38	5.50
26	Dallas-Fort Worth-Arlington	-0.70	2.49	4.26	4.64
27	Dayton OH	1.18	0.99	4.21	4.40
28	Daytona Beach	1.86	3.06	4.77	5.59
29	Denver CO	1.61	1.96	4.07	5.81
30	Des Moines	1.18	2.23	6.11	5.64
31	Detroit MI	2.45	1.42	4.16	3.76
32	Flint	1.70	0.06	4.14	3.35
33	Fort Collins	2.32	3.63	5.82	6.72
34	Fresno CA*	1.35	2.04	4.69	6.08
35	Fort Wayne	0.06	1.76	4.16	7.73
36	Grand Rapids MI	1.59	2.49	5.21	1.09

37	Greensboro NC	0.96	1.92	2.95	7.22
38	Harrisburg PA	0.56	1.69	4.24	3.45
39	Honolulu	3.05	1.28	2.99	12.66
40	Houston	-1.27	1.38	3.95	4.53
41	Indianapolis IN	0.82	2.58	4.37	6.17
42	Jacksonville	1.42	2.96	4.60	7.23
43	Kansas City	0.70	1.66	5.35	5.17
44	Lansing	1.38	1.24	4.45	1.37
45	Lexington	0.67	2.43	6.23	3.25
46	Los Angeles CA*	3.51	0.99	2.26	5.40
47	Louisville	1.48	1.87	4.65	4.53
48	Little Rock	0.21	2.22	4.64	4.63
49	Las Vegas	1.07	6.11	5.11	8.14
50	Memphis	0.46	2.51	4.63	5.75
51	Miami FL	1.98	2.93	3.21	6.94
52	Milwaukee	1.90	1.24	2.42	5.16
53	Minneapolis	2.16	2.20	4.39	4.35
54	Modesto*	2.81	2.76	5.54	7.04
55	Napa*	4.63	3.27	4.35	5.32
56	Nashville	1.31	2.78	4.44	6.38
57	New York*	4.61	0.72	2.34	1.96
58	New Orleans	0.06	0.52	2.94	4.80
59	Ogden	0.67	3.25	4.22	6.08
60	Oklahoma City	-1.21	0.95	5.17	3.66
61	Omaha	0.65	2.03	4.99	4.35
62	Orlando	0.88	5.21	5.30	6.33
63	Ventura*	3.95	2.61	4.19	5.83
64	Peoria	0.38	1.16	4.31	6.93
65	Philadelphia PA*	2.78	1.18	3.52	2.57
66	Phoenix	1.05	4.41	4.27	7.49
67	Pittsburgh	1.18	0.69	2.86	2.75
68	Portland*	2.52	2.61	4.17	7.05
69	Providence*	4.82	0.96	2.83	4.71
70	Port St. Lucie	1.63	3.59	5.60	7.18
71	Raleigh NC	1.15	3.91	4.06	5.42
72	Reno	1.55	2.94	3.94	8.60
73	Richmond	1.31	2.04	4.71	3.60
74	Riverside*	2.46	4.55	6.29	5.80
75	Rochester	0.61	0.80	5.16	1.01
76	Santa Rosa*	4.19	3.06	4.90	2.80
77	Sacramento*	3.02	3.32	5.51	4.94
78	San Francisco CA*	4.23	1.09	2.61	4.73

79	Salinas*	4.81	1.55	3.95	5.47
80	San Antonio	-1.03	2.45	3.70	5.52
81	Sarasota	2.29	4.25	4.69	7.30
82	Santa Barbara*	4.29	1.42	3.16	4.27
83	Santa Cruz*	4.34	2.60	3.19	3.24
84	San Diego*	4.13	2.96	3.62	5.45
85	Seattle*	2.97	2.65	2.95	8.10
86	San Jose*	4.34	1.20	2.85	4.55
87	Salt Lake City	1.39	3.12	3.45	5.72
88	St. Louis	1.48	1.40	4.55	4.82
89	San Luis Obispo*	4.18	3.32	5.49	4.27
90	Spokane*	1.52	2.28	2.81	9.04
91	Stamford*	3.64	0.60	3.14	4.80
92	Stockton*	2.91	2.42	5.59	5.99
93	Tampa	1.45	3.48	3.64	5.61
94	Toledo	0.65	1.18	4.18	5.18
95	Tucson	1.50	2.96	3.32	8.03
96	Tulsa	-0.96	1.00	4.66	4.33
97	Vallejo CA*	3.48	2.87	5.24	5.41
98	Washington DC*	3.01	2.54	4.47	3.26
99	Wichita	-0.47	1.43	5.01	4.39
100	Winston	0.73	1.98	2.92	5.51
101	Worcester*	4.40	1.13	4.18	5.77

Notes: Table provides the average real price appreciation over the 25 years, average job growth rate, average sales rate, and growth in sales rate.
* Denotes "Costal city" in robustness tests.

APPENDIX 2.2

Let $\Delta p_T = [\Delta P_{1T},....,\Delta P_{NT}]$ 'and $\Delta s_T = [\Delta S_{1T},....,\Delta S_{NT}]$ ', where N is the number of markets. Let $W_T = [e, \Delta p_{T-1}, \Delta s_{T-1}, \Delta X_{i,T}]$ be the vector of right hand side variables, where e is a vector of ones. Let $V_T = [\varepsilon_{1T},...,\varepsilon_{NT}]$ be the N x 1 vector of transformed disturbance terms. Let $B = [\alpha_0, \alpha_1, \alpha_2, \beta_1, \delta_1]$ ' be the vector of coefficients for the equation.

Therefore,

$$\Delta p_T = W_T B + V_T \qquad (1)$$

Combining all the observations for each time period into a stack of equations, we have,

$$\Delta p = WB + V. \qquad (2)$$

The matrix of variables that qualify for instrumental variables in period T will be

$$Z_T = [e, \Delta p_{T-2}, \Delta s_{T-2}, \Delta X_{i,T}], \qquad (3)$$

which changes with T.

To estimate B, we premultiply (2) by Z' to obtain

$$Z'\Delta p = Z'WB + Z'V. \qquad (4)$$

We then form a consistent instrumental variables estimator by applying GLS to equation (4), where the covariance matrix $\Omega = E\{Z'VV'Z\}$. Ω is not known and has to be estimated. We estimate (4) for each time period and form the vector of residuals for each period and form a consistent estimator, $\tilde{\Omega}$, for Ω. $\tilde{B}$, the GLS estimator of the parameter vetor, is hence:

$$\tilde{B} = [W'Z(\tilde{\Omega})^{-1}Z'W]^{-1}W'Z(\tilde{\Omega})^{-1}Z'\Delta p. \qquad (5)$$

The same procedure applies to the equation wherein Sales (S) are on the LHS.

APPENDIX 2.3: AHS Data (House Price Data from Census)

year	1985	1987	1989	1991	1993	1995	1997	1999	2001	2003	2005	2007
R change	1061	300	414	266	616	-312	116	-65	-53	-181	665	952
O change	481	1071	1055	762	710	1603	1102	1459	1343	776	978	-221
New Del.	1072.5	1122.8	1026.3	837.6	1039.4	1065.5	1116.4	1270.4	1241.8	1386.3	1635.9	1216.5
New O	537	558	482	473	557	579	409	430	564	501	641	688
New R	2792	2877	2751	2881	2725	2959	2377	2387	2445	2403	2507	2686
RR	7325	7438	7563	7485	7184	7714	7494	6934	6497	6889	7291	7152
OR	1491	1448	1654	1129	1143	1186	1413	1309	1330	1233	1273	1426
OO	1998	2049	1918	1697	1769	1983	2074	2478	2249	2381	2913	2391
RO	2074	2256	2110	1980	2177	2337	2208	2378	2468	2305	2607	2032
O Exits	639	295	-117	562	881	127	102	40	359	797	997	1515
R Exits	1148	1769	1881	1764	1075	2120	1466	1383	1360	1512	508	1128
Lists	5200.5	4914.8	4481.3	4225.6	4832.4	4361.5	4705.4	5097.4	5179.8	5797.3	6818.9	6548.5
Sales	4609	4863	4510	4150	4503	4899	4691	5286	5281	5187	6161	5111
Price	244.1647	264.3485	270.2479	256.9754	253.6429	253.331	258.1007	274.0435	295.7802	322.0829	371.4579	387.985
Real Median House Price	145488.1	156215.7	158258.4	156384.1	156575.9	159199.3	166624.9	175750.3	183404.1	203062.5	232526.1	217900

Appendix 3.1 Hedonics Equation

A) Boston Hedonic Price Regression (1998)

Dependent Variable: House price

lnsale	Coef.	t	P>|t|	sig coeff
Built 1960-1980	-0.05026	-7.11	0	-0.05026
Built 1940-1960	-0.13523	-16.09	0	-0.13523
Built 1900-1940	-0.1662	-18.46	0	-0.1662
Built pre-1900	-0.19206	-15.89	0	-0.19206
1 bedroom	-0.12093	-3.78	0	-0.12093
2 bedrooms	-0.06042	-4.29	0	-0.06042
3 bedrooms	-0.00808	-0.66	0.507	0
4 bedrooms	0.005488	0.46	0.643	0
1 bath	-0.20654	-16.47	0	-0.20654
1.5 bath	-0.14318	-12.71	0	-0.14318
2 bath	-0.16332	-13.98	0	-0.16332
2.5 bath	-0.06041	-6.19	0	-0.06041
interior square feet	0.000298	21.11	0	0.000298
sq. feet squared	-1.32E-08	-5.58	0	-1.3E-08
Lot size	0.14155	13.27	0	0.14155
Lot size sq	-0.01635	-8.98	0	-0.01635
regsco	0.015277	4.09	0	0.015277
density	0.041415	8.51	0	0.041415
Ln median income	0.357986	2.16	0.031	0.357986
_cons	-2.95215	-2.95	0.003	-2.95215

R-Square= 0.7168

B) Phoenix Hedonic Price Regression(1998)

lnsale	Coef.	P>|t|	sig coeff
Built 1980-1990	-0.0886	0	-0.0886
Built 1970-1980	-0.14813	0	-0.14813
Built 1960-1970	-0.20073	0	-0.20073
Built before 1960	-0.2527	0	-0.2527
1 bathroom	-0.11637	0	-0.11637
1.5 bathroom	-0.12147	0	-0.12147
2 bath rooms	-0.05861	0	-0.05861
2.5 bathrooms	-0.02484	0	-0.02484
3bathrooms	-0.00131	0.814	0
5-room	-0.00331	0.709	0
6-room	-0.02291	0.037	-0.02291
7-room	-0.0511	0	-0.0511
Room8-	-0.08753	0	-0.08753
Interior Sq Feet	0.0006	0	0.0006
Interior Sq feet Square	-4.42E-08	0	-4.4E-08
Lot Size	0.260908	0	0.260908
Lot Size Sq	-0.02199	0.001	-0.02199
Pool	0.052737	0	0.052737
Garage	0.057881	0	0.057881
Density	-0.00434	0.511	0
Ln Median Income	0.011572	0.901	0
Constant	11.18304	0	11.18304

R-Square= 0.776

C) San Diego Price Regression (1998)

lnsale	Coef.	P>|t|	sig coef
built80	0.009864	0.45	0
built70	-0.09978	0	-0.09978
built60	-0.09502	0	-0.09502
builtp60	-0.0653	0.001	-0.0653
bed1	0.076764	0.177	0
bed2	0.033985	0.085	0.033985
bed3	0.052283	0.001	0.052283
bed4	0.032621	0.02	0.032621
bath1	-0.0742	0.015	-0.0742
bath15	-0.06265	0.059	-0.06265
bath2	-0.01135	0.667	0
bath25	-0.03225	0.188	0
bath3	-0.02204	0.344	0
sqft	0.000467	0	0.000467
sqftsq	-2.58E-08	0	-2.6E-08
lotszsq	-1.46E-06	0.847	0
lotsize	0.001165	0.734	0
garage	0.135294	0	0.135294
pool	0.076756	0	0.076756
density	0.00358	0.284	0
lnmedinc	0.85806	0	0.85806
_cons	2.262603	0.001	2.262603

R-Square= 0.6994

D) Chicago Hedonic Regression (1998)

lnsale	Coef.	P>|t|	sig coeff
built after 1980	0.195267	0	0.195267
built 1970-1980	0.124359	0	0.124359
Built 1970-60	0.078294	0	0.078294
Built 1960-1970	0.033203	0	0.033203
1 Bathroom	-0.27247	0	-0.27247
1.5 Bathrooms	-0.21495	0	-0.21495
2 Bathrooms	-0.19741	0	-0.19741
2.5 Bathrooms	-0.14139	0	-0.14139
3 Bathrooms	-0.11663	0.004	-0.11663
Interior Sq feet	0.00036	0	0.00036
Interior Sq feet Square	-2.60E-08	0	-2.6E-08
Lot Size	0.381877	0	0.381877
Lot Size Sq	-0.0505	0	-0.0505
density	0.025079	0	0.025079
nonwhite	-0.96735	0	-0.96735
_cons	11.77202	0	11.77202

R-square = 0.6372

Appendix 3.2

Estimation of Rhos

Chicago			Boston		
zip	rho	std deviation	zip	rho	Std deviation
60004	0.048611	-0.2268	1460	0.146566	-0.3752
60007	-0.08801	-0.3377	1532	-0.14064	-0.712
60008	-0.05092	-0.1896	1581	-5.67E-08	-0.383
60016	0.092494	-0.1198	1701	0.108319	-0.5193
60025	0.032324	-0.35	1702	0.254682	-0.4124
60056	0.032324	0.0089	1719	0.3525	-0.2744
60062	0.131688	0.0619	1720	0.358005	-0.349
60067	-0.06284	-0.3829	1721	0.184879	-0.5264
60068	0.169501	0.687	1730	0.490579	-0.3903
60076	0.415859	-0.02	1741	0.06545	-0.4564
60090	-0.15164	-0.1208	1742	0.689924	0.0439
60091	0.33202	-0.2525	1746	0.293363	-0.486
60103	0.158114	-0.3659	1747	0.363329	-0.0325
60107	0.177947	-0.2378	1748	0.211327	-0.2582
60193	-0.137	-0.4917	1752	0.139121	-0.4379
60194	0.093376	-0.2411	1757	0.355482	0.0567
60195	0.088855	-0.302	1760	0.37052	-0.4152
60201	0.432982	-0.3046	1770	0.296614	-0.3547
60302	0.265567	-0.2612	1772	0.01312	-0.6093
60402	0.224017	0.1174	1773	0.483241	-0.572
60430	0.524584	-0.1101	1776	0.36285	-0.1775
60438	0.407394	-0.3561	1778	0.42361	-0.3491
60445	0.664648	0.2258	1801	-0.02617	-0.2153
60452	0.36051	-0.3487	1803	0.479275	-0.182
60459	0.527129	0.2227	1810	0.088643	-0.5361
60462	0.303743	-0.3147	1824	0.081071	-0.3905
60477	0.478473	-0.1746	1843	0.456403	-0.4672
60521	0.365061	0.0937	1844	0.290716	-0.3762
60525	-0.11994	-0.388	1845	0.094984	-0.2466
60618	0.394281	-0.0711	1850	0.492241	-0.2137
60620	-0.00288	-0.1334	1851	0.542552	-0.2919
60625	0.346792	-0.2329	1852	0.543744	-0.3503
60629	0.6299	0.2592	1854	0.535427	-0.2764
60630	0.141456	-0.5705	1862	0.200481	-0.2924
60631	0.272147	0.3063	1864	0.066686	-0.4374
60632	0.297605	-0.131	1867	0.122867	-0.3155

Chicago			Boston		
zip	rho	std deviation	zip	rho	Std deviation
60634	0.043618	-0.3188	1876	0.02212	-0.3016
60638	0.407027	-0.0829	1880	-3.71E-06	-0.431
60639	0.890088	0.0525	1886	0.05403	-0.2758
60640	0.51791	0.0858	1887	0.220483	-0.072
60641	0.87808	0.1713	1890	-0.21493	-0.4835
60645	0.492665	0.0239	1904	0.435805	-0.3608
60646	0.220697	-0.4814	1905	0.430105	-0.4197
60647	0.654336	-0.0185	1906	0.093499	-0.4434
60649	0.111736	-0.0111	1907	-0.02971	-0.4686
60652	0.559695	0.1167	1915	-0.20751	-0.4624
60655	0.357254	-0.0803	1921	-0.26375	-0.4472
60656	-0.263	-0.5569	1923	-0.22053	-0.3879
60659	0.44887	-0.0643	1940	-0.07179	-0.5552
			1945	-0.36572	-0.5748
			1960	-0.05747	-0.5169
			1970	0.175351	-0.3639
			1983	-0.12863	-0.5332
			2019	0.683471	0.2886
			2025	0.117491	-0.2387
			2026	-0.27677	-0.8183
			2030	-0.15028	-0.265
			2035	0.448787	-0.3262
			2038	0.510869	-0.1835
			2045	-0.08283	-0.5645
			2053	0.322168	-0.4034
			2054	0.510805	-0.251
			2056	0.318303	-0.419
			2062	-0.41902	-0.5932
			2067	0.285696	-0.3552
			2072	0.393808	-0.1077
			2081	0.079075	-0.5905
			2093	0.630422	0.0699
			2124	0.548664	-0.2528
			2126	0.383723	-0.3164
			2131	0.239554	-0.4592
			2132	-0.35573	-0.8064
			2136	0.259943	-0.3317
			2140	0.118145	-0.5349

Boston		
zip	rho	Std deviation
1876	0.02212	-0.3016
2149	0.367784	-0.5034
2155	0.123323	-0.4882
2169	0.62796	-0.1304
2170	0.378005	-0.3476
2171	0.04049	-0.7303
2176	-0.22261	-0.5753
2180	-0.24702	-0.5765
2184	0.624568	0.0106
2186	0.280834	-0.5063
2343	0.454607	-0.368
2368	0.451192	-0.338
2420	0.43438	-0.4771
2421	0.500384	-0.4405
2445	-0.32986	-0.3987
2451	0.354868	-0.4878
2452	-0.00656	-0.761
2453	0.456466	-0.3778
2458	0.257591	-0.5663
2459	0.40158	-0.4256
2460	0.435922	-0.5738
2461	0.689296	-0.1327
2465	0.505042	-0.6212
2472	0.273476	0.377
2476	0.26875	-0.2661
2478	0.075171	-0.482
2481	0.554466	-0.2635
2482	0.403213	-0.3269
2492	0.644959	-0.1906
2493	0.500609	-0.3539
2494	0.598435	-0.2175

Phoenix			San Diego		
zip	rho	std deviation	zip	rho	Std Deviation
85003	0.39254	-0.6493	92065	0.215272	-0.3582
85006	0.268164	-0.7713	92131	0.104271	-0.1402
85007	0.066254	-0.6559	92129	-0.05869	-0.5524
85008	0.274727	-0.1764	92130	0.178116	-0.3139
85009	0.141134	-0.2891	92056	0.250375	-0.2143
85013	0.015599	-0.6646	92127	-0.08271	-0.5437
85014	-0.4197	-0.7427	92082	0.040015	-0.2781
85015	0.084183	-0.608	92064	0.021091	-0.3073
85016	0.281585	-0.5888	92019	0.08729	-0.2904
85017	0.09509	-0.4141	92128	0.394091	-0.2922
85018	-0.09609	-0.5995	92126	0.095368	-0.5086
85019	0.145434	-0.2707	92026	0.22069	-0.2433
85020	0.316775	-0.1634	92009	0.318531	-3.58E-01
85021	0.547019	-1.60E-03	92139	0.110018	-0.326
85022	0.643009	0.0999	92071	0.312465	-0.0807
85023	0.042151	-0.2227	92027	-0.03829	-0.3907
85024	0.404535	-0.0113	92114	0.151272	-0.3881
85027	0.491612	0.5986	92040	0.245662	-0.3119
85028	0.141711	-0.3827	92028	-0.13326	-0.3293
85029	0.301491	-0.1011	92083	0.086667	-0.2413
85031	0.143182	-0.2521	92084	0.173276	0.0356
85032	0.479644	0.0495	92154	0.255318	-0.0914
85033	0.368969	-0.102	92124	0.027316	-0.5508
85035	0.174849	-0.3495	92119	-0.05619	-0.434
85037	0.128464	-0.1903	92069	0.005387	-0.4243
85040	0.151643	-0.2779	92120	-0.03749	-0.4243
85041	0.034916	-0.361	92123	0.428613	-0.2153
85043	0.028465	-0.2166	92014	0.337878	-0.4111
85044	0.350704	-0.2431	92024	0.169606	-0.2844
85048	0.370531	0.1757	92025	-0.28027	-0.6085
85050	0.307235	-0.1289	92021	0.209972	-0.4419
85051	0.34648	-0.0209	92008	-0.19169	-0.5263
85053	0.180449	0.0347	92007	-0.07142	-0.5276
85201	0.006612	-0.3206	92117	-0.23493	-0.7751
85202	-0.10424	-0.504	92020	-0.15139	-0.5764
85203	-0.13484	-0.4828	92111	0.038558	-0.5445
85204	-0.31658	-0.586	92122	-0.17415	-0.5944
85205	-0.09911	-0.818	92075	0.02309	-0.512
85206	-0.1438	-0.671	92116	-0.07402	-0.6058

Phoenix			San Diego		
zip	rho	std deviation	zip	rho	Std Deviation
85207	0.67333	-0.134	92104	0.029577	-0.5645
85208	0.194965	-0.406	92107	0.197752	-0.3628
85210	0.14067	-0.1402	92103	-0.10906	-0.5376
85213	0.230783	-0.5629	Phoenix		
85215	0.3071	-0.1548	zip	rho	std deviation
85220	0.113442	-0.1736	85353	0.208344	0.1216
85224	-0.0301	-0.405	85363	0.754341	0.0357
85225	-0.05273	-0.5287	85373	0.713871	-0.2103
85226	0.119584	-0.4815	85374	0.238065	0.3001
85233	0.627211	0.1785	85375	0.89958	0.7127
85234	0.383948	-0.4632	85382	0.890297	0.6528
85242	0.096535	-0.3492			
85248	0.476103	-0.0657			
85249	0.094162	-0.3157			
85250	0.302333	-0.3067			
85251	0.447871	-0.1847			
85254	0.376367	-0.091			
85255	0.100586	-0.3553			
85257	0.509044	-0.3276			
85258	0.257615	-0.1271			
85259	0.032463	-0.1467			
85260	0.232409	-0.2502			
85268	0.115723	-0.2			
85281	0.780484	0.0009			
85282	0.364922	-0.4778			
85283	-0.00892	-0.3143			
85284	-0.30106	-0.1369			
85296	0.159502	-0.4182			
85301	0.2301	-0.2181			
85302	0.185275	-0.3729			
85303	0.197384	-0.2725			
85304	0.476978	-0.3496			
85305	-0.0788	-0.1495			
85306	0.709495	-0.0371			
85308	0.694061	-0.3243			
85310	0.808134	0.6545			
85331	0.401163	-0.0069			
85345	0.442559	-0.5326			
85351	0.769993	0.584			