Neighborhood Externalities and Housing Price Dynamics

Veronica Guerrieri[*]
University of Chicago and NBER

Daniel Hartley[†]
Federal Reserve Bank of Cleveland

Erik Hurst[‡]
University of Chicago and NBER

September 7, 2009

Abstract

In this paper, we explore differential changes in house prices across neighborhoods within a city to better understand the nature of house price dynamics. Using a variety of different data sources, we find that two stable empirical relationships emerge. First, we find that initially low priced neighborhoods systematically experience higher appreciation rates during city wide housing booms than do initially higher priced neighborhoods. Second, we find that the extent of the convergence of prices across neighborhoods is greater the larger the price appreciation rate for the city as a whole. We then present a spatial equilibrium model of a city to explain the above facts. The key ingredient of our model is the presence of a positive neighborhood externality: agents like to live in areas where other agents live. The idea is that local amenities are better the more people live in the neighborhood. We show that if a positive house demand shock hits the city, house prices will appreciate more in low priced neighborhoods relative to high priced neighborhoods, as our first fact predicts. In addition, relative to a standard model with convex adjustment costs, the existence of a positive neighborhood externality generates a persistent increase in housing prices in the long run and may amplify the reaction of prices in the short run. We also show that our model is consistent with the second fact. In particular, we explore an extension of the model with gentrification. We show that cities at a more advanced stage of development feature more gentrification in reaction to similar demand shocks, hence experiencing bigger housing price booms and more neighborhood convergence. In the final part of the paper, we provide additional empirical support for the model's prediction.

[*]Contact: *vguerrie@chicagobooth.edu*
[†]Contact: *daniel.hartley@clev.frb.org*
[‡]Contact: *Erik.Hurst@chicagobooth.edu*

1 Introduction

According to the Case-Shiller index, real residential property prices appreciated by over 50 percent for the United States as a whole between 2000 and 2006. Between the first quarter of 2007 and the first quarter of 2009, real property prices have dropped by 33 percent nationwide. The national data, however, mask a large degree of heterogeneity at more localized levels. For example, real property prices increased by over 100 percent in Washington DC, Miami, and Los Angeles between 2000 and 2006 while property prices appreciated by less than 10 percent in Charlotte, Denver, and Detroit during the same time period. Despite the importance of changing housing prices in determining household consumption (Mian and Sufi (2009)) and borrower default (Foote, Gerardi, and Willen (2008)), relatively little work has been done trying to explain the nature of housing price dynamics.[1] In this paper, we explore differential changes in house prices *across neighborhoods within a city* to shed light on the nature of house price dynamics in reaction to housing demand shocks.

We start the paper by documenting a series of new stylized facts about the movement of house prices across neighborhoods within a city during city-wide housing price booms and busts. We are the first paper to systematically explore the movement of house prices across neighborhoods within a city for a large cross section of cities during many different time periods.[2] Using a variety of different data sources including the zip code level data that underlies the Case-Shiller indices, U.S. Census data, and deed data that we have compiled for certain cities, we find that two stable relationships emerge. First, we find that initially low priced neighborhoods systematically experience higher appreciation rates during city wide housing booms than do initially higher priced neighborhoods. For example, during $2000 - 2006$, initially low priced neighborhoods in New York (e.g., Harlem) appreciated at three times the rate of initially high priced neighborhoods in New York (e.g., Mid-town Manhattan). Second, we find that the extent of the convergence within the city (defined as the appreciation rate of lower priced neighborhoods relative to higher priced neighborhoods) is greater the larger the price appreciation rate for the city as a whole. For example, while low initially price neighborhoods appreciated at three times

[1] Some notable exemptions include Topel and Rosen (1988), Case and Shiller (1989), Glaeser, Gyourko, and Saks (2005a), Glaeser and Gyourko (2006), Gyourko (2009), Kiyotaki, Michaelides, and Nikolov (2008) and Saiz (2009). In the Topel and Rosen, Glaeser, Gyourko, and Saks, Glaeser and Gyourko, and Gyourko papers, the dynamics are achieved – in part – by focusing on constraints to adjusting housing supply. These constraints take the form of either standard convex adjustment costs which slow down the speed of building, regulations which prevent building, or physical land barriers (such as oceans or lakes) which also prevent building. In section 6, we discuss how our work relates to this and other relevant literatures.

[2] A few papers exist that examine zip code level price movements within a city for a given time period. For example, Case and Mayer (1996) examine price movements by neighborhood in Boston between 1982 and 1992 while Case and Marynchenko (2002) look at within city price movements for Boston, Chicago, and Los Angeles during the 1983 to 1993 period. We discuss these papers in greater depth in section 6.

the rate of higher priced neighborhoods in New York between 2000 and 2006, less expensive neighborhoods in Chicago appreciated at only twice the rate of more expensive neighborhoods and there was no difference in appreciation rates between high and low priced neighborhoods in Charlotte during that time period. Between 2000 and 2006, the New York metro area as a whole experienced a 70 percent real increases in housing prices while Chicago and Charlotte experienced real housing price increases of roughly 40 percent and 10 percent, respectively.

In this section, we also document that these relationships between city wide housing booms and the convergence of neighborhoods within a city are prominent features of the data in all time periods that we analyzed. In particular, we show that such results held during the recent property price boom as well as local property price booms that occurred during the 1980s and 1990s. Finally, we show some evidence that property prices across neighborhoods within a city diverge during city wide property price busts. In other words, low price neighborhoods depreciate more during periods of real housing declines.

In the next part of the paper, we present a spatial equilibrium model of property price movements across neighborhoods within a city to explain the above facts. There are two key features of our model. First, we analyze a linear city (in the spirit of Mills (1967) and Muth (1969)). Such an assumption is common in most spatial equilibrium models of a city.[3] Second, we model within cities positive neighborhood externalities.[4] In particular, we assume that there are increasing returns to scale in the production of neighborhood amenities (number and variety of restaurants, easier access to service industries such as dry cleaners, movie theaters, etc.) such that these amenities are more common as the number of people in the neighborhood increases. In an extended version of the model, we further refine the neighborhood externalities by making them positively associated with the number of "rich" people in the neighborhood and negatively related to the number of "poor" people. In doing this, we are capturing the fact that density - per se - does not yield better amenities but instead it is the density of a certain type of people that generates the positive consumption amenities. The importance of urban density in facilitating local consumption externalities has been recently emphasized in the work of Glaeser, Kolko, and Saiz (2001) and in Becker and Murphy (2003). Our innovation is to embed these consumption externalities into a model of neighborhood development within a city and show how such preferences affect the reaction of house prices to housing demand shocks both across various types of neighborhoods and at the aggregate level.

[3]For an exception, see Robert E. Lucas and Rossi-Hansberg (2002).

[4]Our externality is a pure consumption externality, different from productive externalities that are usually emphasized in spatial equilibrium models.

In our baseline model, we focus on a linear city populated by a continuum of identical households where the neighborhood externalities are generated from living around other households. Moreover, we assume that the supply of houses is perfectly elastic. In equilibrium, the households will reside only on a fixed interval on the line describing the city, given that they strictly prefer to live close together. Toward the center of the interval, housing prices are the highest because the residential agglomeration effects are strongest. As households move away from the center city, the local consumption externalities get weaker. To compensate for the lower neighborhood externality, house prices fall.[5] At the boundaries of the city, per unit house prices are equal to the marginal cost of building an additional unit. In the center city, house prices are strictly above marginal cost.

The first insight of this simple model is that a shock to housing demand will push out the boundaries of the city. In doing so, the price in the center city will remain constant and prices toward the previous boundary of the city will grow rapidly as their neighbors get more developed. In this simple model, our first fact is generated. A housing demand shock will cause low price neighborhoods within the city to appreciate at a much higher rate than high price neighborhoods because the low price neighborhoods get more developed and, in doing so, offer higher amenities associated with the additional agglomeration. We show that the extent of the convergence of initially low priced neighborhoods is determined by both the size of the demand shock and the importance of the residential agglomeration effect in preferences.

We then extend our model slightly to include convex adjustment costs to building at the city wide level. We assume that the more the city as a whole adjusts supply during a given period, the higher the marginal cost of building within that period. Such adjustment costs are found in the models of housing put forth by Topel and Rosen (1988) and by Glaeser and Gyourko (2006). This simple addition to the baseline model yields a much richer set of predictions. First of all, different neighborhoods experience different price patterns as they adjust toward the long run level. The high price neighborhoods in the city center exhibit a standard overshooting pattern, as the demand pressure is stronger in the short run when the housing supply is slowly adjusting. The low price neighborhoods experience a hump-shaped pattern: initially they expand and the local externality makes them more desirable, but eventually new neighborhoods develop reducing the pressure of housing demand and driving prices down. The neighborhoods that are

[5] This is a similar prediction as the models put forth by Mills (1967) and Muth (1969) which focused on transportation costs. The farther away from the center city one lived, the larger their commuting costs and, as a result, the lower the price they were willing to pay to live far away from the city. In this paper, we abstract from commuting costs completely. Commuting costs would not be able to explain the features of the data that we describe in the first and third parts of our paper.

born later simply experience an increase in prices. We also compare our model to a similar model with adjustment costs and no neighborhood externalities. The main difference is that our model features a permanent increase in aggregate prices due to the development of low price neighborhoods. Moreover, for some parameter configurations, our model also delivers an amplified short run response of prices to demand shocks.

Finally, we extend our model by focusing on a city of bounded size and by including two types of households: rich and poor. In this model we assume that the rich have higher utility if they live around other rich and have lower utility if they live around other poor. In doing so, we get a further set of predictions about what types of neighborhoods evolve during positive shocks to housing demand. This gives us a link between demand shocks to housing, house price dynamics, and the gentrification of neighborhoods. As housing demand increases (perhaps due to falling interest rates) households will expand their desired consumption of housing. The rich do this by moving into the neighboring poor neighborhoods – in essence expanding the boundaries of the rich neighborhoods. Housing demand shocks result in the gentrification of neighborhoods directly next to the rich neighborhoods. As the rich move in and drive up property prices (given the agglomeration effects), the poor optimally choose to move out.

Given the city is bounded, the poor live in denser neighborhoods toward the edge of the city after each housing demand shock. We show that successive positive demand shocks to housing lead to larger housing price booms over time in a given city. The reason for this is that poor neighbors act as a barrier to the rich with respect to adjusting the amount of housing they can consume when faced with a positive demand shock. The denser the poor neighborhoods become at the fringe of the city, the larger the demand shock is necessary to get a richer person to move into the bordering poor neighborhood.

In the next part of the paper, we empirically test additional predictions of the model we developed. First, we test directly the spatial equilibrium predictions of our model. We show that it is the low priced neighborhoods that are directly next to the high priced neighborhoods that experienced the largest price increases during a city wide property price boom. Second, using detailed data from the IRS and from the U.S. Census, we show strong evidence that the neighborhoods that grew the fastest within a city during the housing price boom showed signs of gentrification. Finally, we use detailed data on building permits within Chicago to show that these gentrifying neighborhoods that we identify are also the ones that experienced the most new buildings.

In the last section of the paper, we discuss how our results touch on four literatures: 1)

models of spatial equilibrium, 2) papers focusing on house price dynamics, 3) papers focusing on neighborhood gentrification, and 4) papers that look at within city movements in housing prices. As we discuss in that section, our paper contributes substantially to all of these literatures.

In summary, our paper shows that the existence of neighborhood externalities has important implications for the nature of real estate prices dynamics across neighborhoods within a city and across cities. Also, the same city may experience different price dynamics at different stages of development in different points of time for a given sized housing demand shock. Moreover, we show empirical support for many of the models predictions – in particular the movement of housing prices across neighborhoods within a city when the city is faced with a positive increase in housing demand.

2 Data

To examine house price appreciation across neighborhoods within a city during a city-wide housing boom, we use a variety of different data sets. In particular, we use (1) zip code level housing price indices computed as part of the Case-Shiller index, (2) transaction level data on the universe (or near universe) of residential housing transactions for Chicago, New York, and Charlotte, and (3) micro data from the 1980 and 1990 U.S. Censuses. As we show in the following section, all three data sources yield similar results about how housing prices evolve within a city during a city-wide housing price boom.[6]

The bulk of our results use the Case-Shiller zip code-level price indices.[7] The Case-Shiller indices are calculated from data on repeat sales of single-family homes. The benefit of the Case-Shiller index is that it provides consistent constant-quality price indices for localized areas within a city or metropolitan area over long periods of time. Most of the zip code-level price indices go back in time through the late 1980s or the early 1990s. The data was provided to us at the quarterly frequency and our most recent data is for the fourth quarter of 2008. As a result, for each metro area, we have quarterly price indices on selected zip codes within selected metropolitan areas going back roughly 20 years.

There are four limitations to the Case-Shiller data. First, the Case-Shiller index only covers approximately 30 metropolitan areas.[8] Second, as noted above, for almost all cities the Case-

[6]For a complete discussion of the data sources - including a link to all of our online documentation - see the Data Appendix.

[7]The zip code indices are not publicly available. Fiserv, the information company overseeing the Case-Shiller index, provided them to us for the purpose of this research project. The data are the same as the data provided to other researchers studying very local movements in housing prices. See, for example, Mian and Sufi (2009).

[8]We were provided with zip code level data for at least 20 zip codes at the city level or at least 65 zip codes at the metro area level for the following cities/metropolitan areas: Atlanta, Baltimore, Boston, Charlotte, Chicago, Cincinnati, Columbus (OH), Dayton, Denver, Detroit, Hartford, Jacksonville, Las Vegas, Los Angeles, Memphis,

Shiller index does not start until late 1980s or early 1990s. This limits the analysis we can do for cities that experienced housing price booms in the early to mid 1980s. Third, the Case-Shiller index only covers single-family homes (as opposed to also including condos or multi-family buildings). To the extent that condos and multi-family buildings evolve differently during housing price booms and to the extent that different neighborhoods have different compositions of housing structures, the Case-Shiller data could yield a biased picture of within city house prices changes. Finally, given its focus on repeat sales of single-family homes, there is not enough data to compute reliable price indices for all zip codes within a metropolitan area. As a result, many zip codes - particularly those in center cities - have no reported Case-Shiller price indices. Appendix Figure A1 illustrates this point. In this Figure we show the zip codes within three cities: Chicago (panel A), New York City (panel B), and Charlotte (panel C). The darkened zip codes on the city maps are the ones for which a Case-Shiller index exists. Notice, that a Case-Shiller zip code index exists for less than 50 percent of the zip codes in Chicago, less than 10 percent of the zip codes in New York City, and essentially all of the zip codes in Charlotte.[9] Notice, for New York City proper, only some zip codes in Staten Island are included.

To overcome some of the shortcomings of the Case-Shiller zip code indices, we supplement our analysis using data from two additional sources. First, for three cities, we were able to get data for the universe (or near universe) of housing transactions. The three cities were Chicago, Charlotte, and New York. These data allow us to explore whether our results change using data that include condos and multi-family buildings (as well as single-family homes) and allows us to examine price behavior in all zip codes within-city. As we show below, the results in these more detailed datasets mirror the results from the Case-Shiller indices giving us confidence that these limitations of the Case-Shiller indices are not biasing our results.

For Chicago and Charlotte, we compiled the data on all residential real estate transactions ourselves. Using the Chicago Tribune website, we downloaded all residential real estate transaction data for the City of Chicago. The data we were able to download include all residential real estate transactions from 2000 through 2008 (inclusive). We believe this to be the universe of residential real estate transactions. We merged the data from the Chicago Tribune on Chicago real estate transactions with information from the Cook County Tax Assessor which included

Miami, Minneapolis, New York, Orlando, Philadelphia, Phoenix, Portland (OR), Sacramento, San Bernardino, San Diego, San Francisco, San Jose, Seattle, Tampa, and Washington DC.

[9] Coverage in the metro areas are much higher given that most zip codes outside the center city have a sufficiently large number of single-family home transactions to compute reliable price indices. The reason that zip codes in center cities tend to have lower coverage is that most homes sales are condos or multi-family buildings which are excluded from the Case-Shiller index.

information on the age of the structure, building type, etc.[10]. For Charlotte, the procedure was much easier. We downloaded all real estate transactions back to 1990 using the Mecklenburg County Real Estate Lookup System. These data include an extensive list of structural characteristics. Like the data from the Chicago Tribune, we believe the Charlotte Deed data to be the universe of all residential real estate transactions.

Given that we had some attributes of the structure in the Chicago data that we merged in from the Cook County Tax Assessor, we made a simple price index for each Chicago neighborhood.[11] The real estate transactions in the Chicago Tribune are mapped to Chicago neighborhoods (all within the city of Chicago). Chicago neighborhoods are slightly smaller than Chicago zip codes. For example, there are 77 Chicago neighborhoods and only 60 Chicago zip codes. In Appendix Figure A2, we show the Chicago neighborhood map. We have kept the Chicago Tribune price data at the level of Chicago neighborhood (instead of converting them to zip codes) so as to match our building permit data - which is at the neighborhood level - discussed in Section 5. To make the simple price index, we regressed all Chicago residential real estate transaction (log) price data on dummies for building type (multi-family, single-family, or condo), dummies for the age of the building (1-5 year old, 6-10, 11-20, 21-30, etc.), and dummies for neighborhood interacted with year. We evaluated the estimating equation using the mean structural characteristics for the entire city and added this component to coefficients for the neighborhood-year variables to form the neighborhood price indices over time. We used the same procedure to construct a hedonic index for Charlotte zip codes over time. The only difference was that we had a much richer set of building characteristics for each property in Charlotte.

For New York City, we use the Furman Center repeat sales index which covers all of NYC. The Furman data uses NYC community districts as its level of aggregation.[12] There are 59 community districts in New York City which represent clusters of several neighborhoods. The Furman data for New York City extends back to 1974. The benefit of the Furman data is that it gives us extensive coverage of New York neighborhoods over a long time period and covers all residential real estate transactions in New York City (not just single-family homes). Additionally, this is a repeat sales index measuring the change in price for constant quality housing units. A map of the NYC community districts can be found in the second panel of

[10] In the Data Appendix, we discuss our methodology in much greater detail

[11] For Chicago, the only structural characteristics that are available for all property types are the age of the building and the property type. A much richer set of structural characteristics are available for all Charlotte properties. See the data appendix for further details.

[12] See `http://furmancenter.org/`.

Appendix Figure A2.

To help examine neighborhood trends within a broad set of cities during the 1980s, we augment our analysis using data from the 1980 and 1990 U.S. census. We restricted our analysis of the U.S. Census data to census tracts within a metro area. So, when we use the census data, our unit of analysis - a census tract - is much smaller than zip code or community area. We only focus on census tracts whose boundaries remained constant between 1980 and 1990. For each census tract within the metro area, we computed the growth rate in median home prices. When computing the growth rate in median house prices by census tract, we did not hedonicly adjust the series for changing housing characteristics. As a result, using this data, we can explore changes in house prices within a neighborhood that are both due to the fact that the price of a constant-quality unit of housing may be changing and due to fact the quality of housing in the neighborhood may be changing. Again, the Data Appendix gives a complete description of all of our data sources and our sampling restrictions.

3 Within City House Price Movements During Housing Booms

In this section, we explore the nature of housing price movements within a city or metro area during local housing price booms. We begin by exploring data from the 2000s. In separate sub-sections, we explore similar patterns within cities that experienced housing price booms in the 1990s and then in the 1980s. The key results documented in this section are very robust to both the use of alternate housing price series and to alternate time periods.

3.1 Within City Movements in Housing Prices: The 2000s

We begin our analysis by estimating the following relationship:

$$\overset{g}{HP}{}^{i,j}_{t,t+k} = \alpha^{j}_{t,t+k} + \beta^{j}_{t,t+k} \ln(HP^{i,j}_{t}) + \epsilon^{i,j}_{t,t+k} \tag{1}$$

where $HP^{i,j}_{t}$ is the level of housing prices in neighborhood i, within city (metro area) j, in year t and $\overset{g}{HP}{}^{i,j}_{t,t+k}$ is the growth in housing prices in neighborhood i, within city (metro area) j, between years t and $t+k$. We estimate these relationships separately within each city (metro area). $\beta^{j}_{t,t+k}$ is an estimate of the relationship between the initial level of house prices (in logs) and the subsequent house price appreciation within city (metro area) j between t and $t+k$. We will examine neighborhood price movements within both cities and broader metro areas. As a result, sometimes j will index a city and sometimes it will index the broader metro area that contains that city.

Figure 1 shows the estimates of (1) where j = Chicago (Panel A), New York City (Panel B), and Charlotte (Panel C) over the 2000-2006 time period.[13] Along with the actual estimated regression line, we plot the underlying neighborhood data. The house price growth data (on the y-axis) is the growth rate in the neighborhood level house prices indices (which were discussed in the previous section). The initial level of house prices (on the x-axis) is the median house price in the neighborhood as reported by 2000 census.[14] We chose these three metro areas to start our analysis with for two reasons. First, as noted in the previous section, we have house price measures from multiple sources for these three cities. Second, the cities provide a nice contrast with each other given that the New York metro area as a whole experienced a substantial housing price boom between 2000 and 2006 (of 75 percent), while the Chicago metro area experienced a medium sized housing price boom during that period (of 40 percent) and the Charlotte metro area experienced a very small housing price boom during that period (of 8 percent).[15]

For each of the three cities in Figure 1, we provide three panels that shed light on the relationship between house price growth across neighborhoods and the initial level of prices in the neighborhood. In the left most panel, we use our alternate data sets for each city. For Chicago, we use the Chicago Tribune data. For New York, we use the Furman data for Manhattan. And for Charlotte, we use the universe of deed data. In the middle panel, we use the zip code level data from Case-Shiller. As noted above, for New York City, the Case-Shiller data only covers some zip codes in Staten Island (see Appendix Figure A1). The left and middle panels focus on only neighborhoods within the city of Chicago, New York or Charlotte. In the right most panel, we expand our analysis to include all the zip code data from Case-Shiller for the entire metro area.

Focusing on the Chicago panel of Figure 1 (Panel A), we see that both the Chicago Tribune data and the Case-Shiller data show a large amount of convergence of zip codes across neighborhoods within Chicago during the 2000-2006 period. Focusing on the Tribune data (which, as discussed in the previous section, only control for limited property characteristics in the hedonic relationship), we find that prices of transacted properties in the initially low priced

[13] All prices in the paper are measured in year 2000 dollars.

[14] Figures with zip code-level observations use 2000 Census summary file (SF3) tabulations of median value of owner-occupied housing units by zip code, which are available from `http://factfinder.census.gov/`. The left-most figure of Panel B also uses 2000 Census tabulations of median value of owner-occupied housing units, but these are specially tabulated for NYC community districts and are are available from NYC's GIS department at `http://gis.nyc.gov`. Finally, the x-axis for the left-most figure of panel A uses the mean owner-occupied housing value calculated from census 2000 tabulations by census tract and described in the data appendix.

[15] To compute the change in the metro area price indices as a whole, we use the Case-Shiller metro area price indices.

Chicago neighborhoods grew at a rate that was 2 to 3 times higher than the prices of transacted properties in the initially high priced Chicago neighborhoods. For example, the average transacted property grew at about 25 percent or less in Chicago's high priced neighborhoods of Lincoln Park (community 7), Lakeview (community 6) and The Loop (community 32). Conversely, some of the neighborhoods that had initially low price levels experienced an increase of 75 percent or more for the average value of transacted property prices.

The left hand panel of Figure 1A also includes estimates of (1). The estimated β from (1) for Chicago using the Chicago Tribune data during the 2000-2006 period is -0.33 with a robust standard error of 0.05. In other words, a 100 percent increase in the initial level of housing prices reduces the growth rate in housing prices by 33 percentage points. The simple R-squared of the scatter plot is 0.29.

The data in the left-hand panel of Figure 1A includes the universe of transacted properties within the city of Chicago during the 2000-2006 period. The drawback of this data is that it is not based on repeat sales transactions. To explore the movement of prices using a better neighborhood repeat sales price index, we use the Case-Shiller zip code data for Chicago. As noted above, the Case-Shiller data only cover about 45 percent of Chicago's 60 or so zip codes. Moreover, the Case-Shiller data only focus on the price movements of single-family homes. Despite the differences in coverage, the results in the middle panel of Figure 1A (using the Case-Shiller data) are very similar to the results in the left panel of Figure 1A (using the Chicago Tribune data). A one-hundred percent increase in initial housing prices reduces the growth rate in housing prices by 33 percentage points (with a standard error of 4 percentage points). The simple R-squared of the scatter plot is 0.78. The right hand panel of Figure 1A shows the results using the Case-Shiller zip code data for the entire Chicago metro area (as opposed to just the city of Chicago). Again, a similar pattern emerges. High price neighborhoods experienced lower appreciation rates than lower priced neighborhoods (β = -0.23 with a standard error or 0.04).

In Figure 1B, similar results are shown for New York. Using the Furman data for community districts in Manhattan (left-hand panel) or using Case-Shiller zip codes from Staten Island (middle panel), we find that a 100 percent increase in initial housing prices reduced the subsequent growth rate in housing prices between 2000 and 2006 by 39 percentage points or 32 percentage points, respectively. Both estimated slope coefficients are significant at the 1 percent level. For example, the Harlem area of Manhattan appreciated at twice the rate of midtown Manhattan. As seen in the right hand panel of Figure 1B, the results also hold broadly for the New York metro area as whole (β = -0.35 with a standard error of 0.02).

In Figure 1C, we see similar graphical representations for Charlotte. The results, however, for Charlotte are very different. No matter what level of aggregation and no matter what measure of housing prices, there is either no systematic relationship between the initial level of housing prices and the subsequent growth in housing prices or evidence in the opposite direction. Across the three figures, β equals 0.14, 0.07 and 0.08 - only the first coefficient is significant at standard levels.

Similar patterns can be found in Figure 2. In the top panel of Figure 2, we show the convergence patterns for other metropolitan areas (using the Case-Shiller data) that experienced large metropolitan wide housing price increases. These metro areas include Boston, Los Angeles, San Francisco, and Washington DC. These pictures are analogous to the right hand panels of Figure 1. In every one of these metro areas, property prices as a whole increased by at least 50 percent throughout the metro area. Also, in every one of these metro areas, the neighborhoods with initial low levels of housing prices increased by a substantially higher amount than higher price neighborhoods. For example, the estimated β's for Boston, L.A., San Francisco, and Washington, D.C., were, respectively, -0.22, -0.40, -0.44 and -0.49 (all significant at the 1 percent level).

In the bottom panel of Figure 2, we present similar plots for metro areas that did not experience large property price booms. These metro areas include: Atlanta, Cincinnati, Denver, and Seattle. In none of these metro areas do we find that subsequent local neighborhood price appreciation during the 2000-2006 period was correlated with the initial level of property prices in that neighborhood. The estimated β's for these metro areas, respectively, were -0.02, 0.01, 0.04, and -0.02.

Figure 3 formalizes the relationship between the level of housing price growth within the metro area as a whole and the amount of cross-neighborhood convergence that takes place within the city. In particular, we estimate the following:

$$\beta^j_{t,t+k} = \gamma_0 + \gamma_1 \overset{g}{HP}{}^{j}_{t,t+k} + \eta^j_{t,t+k} \tag{2}$$

where the β^j's are estimated for each metro area as described in (1) and $\overset{g}{HP}{}^{j}_{t,t+k}$ is house price appreciation in the entire metro area j between t and $t+k$. When estimating (2), we restrict our analysis to the 30 metro areas where we have zip code level price indices from Case-Shiller.

Figure 3A shows that the larger the price increase in the metropolitan area, the more convergence takes place within the city (the estimated metro area β is more negative). This Figure restricts the analysis to the 2000 to 2006 period. In other words, the β's are the same as ones estimated in Figures 1 and 2 - but for all the metro areas in the Case-Shiller data with

an adequate number of zip codes.[16] The figure shows that places that experienced large price increases also had the price appreciation rates of initially lower price neighborhoods that were much higher than the price appreciation of initially higher priced neighborhoods. The simple correlation from the data underlying Figure 3 is 0.50. The negative relationship across metro areas between the size of house price growth in the metro area and the rate of convergence (the β's) is highly statistically significant (γ_1 = -0.37 with a bootstrapped standard error = 0.03).[17]

Are the results symmetric during periods of housing price declines? For the recent episode this is hard to answer given that prices are still falling (and our data only goes through the fourth quarter of 2008). However, Figure 3B shows some evidence that the results are weakly symmetric. During this time period, every metro area in our sample experienced some decline in housing prices between 2007 and 2008. Figure 3B shows the estimates from (2) except the time period is for house price changes 2007 to 2008. On average, places with bigger price declines for the metro area experienced a greater divergence in housing prices between initially high and low priced neighborhoods between 2007 and 2008 (i.e., β was more positive). Specifically, the estimated relationship between the change in house prices and the rate of converge (β) was -0.19 with a bootstrapped standard error of (0.04). Despite the economic and statistical significance of the relationship, the R-squared from the regression in Figure 3B is low (0.09). This may be due to the fact that we have a very short period of time where housing prices have fallen in our data. In the subsequent sub-section, we will explore this question in greater depth using the housing price busts that occurred in many metropolitan areas during the 1990s.

3.2 Within City Movements in Housing Prices: The 1990s

Are the patterns documented above robust to other time periods? The answer is yes. In this sub-section, we focus on results from the 1990-1997 period.[18] The early 1990s is another good laboratory to study the movement of housing prices across neighborhoods within a metro area given that some metro areas experienced rapidly accelerating housing prices while others experienced non-trivial housing price declines.

Figure 4 shows the results of estimating (1) for the Denver metro area (Panel A) and Portland, Oregon metro area (Panel B) for the 1990 - 1997 time period. Denver and Portland both experienced very large metro area real housing price booms during the 1990s (on the order

[16] Our zip code sample includes any city with at least 20 city zip codes or 65 metro area zip codes present in the Case-Shiller data.

[17] To compute bootstrapped standard errors we re-sample with replacement stratified by city/metro area and then compute the second stage coefficient. We repeat this operation 500 times.

[18] We stop at 1997 because that is the beginning of subsequent housing booms that continued through 2006 in many metro areas. We wanted to focus on housing price dynamics unrelated to the current period of rapidly appreciating housing prices.

of 30 - 40 percent). In both of these cities (the left panels) or metro areas more broadly (the right panel), there was a sharp convergence of property prices during the housing price boom. As with the property housing price booms during the 2000s, the appreciation of housing prices in initially lower housing price neighborhoods was substantially higher than the appreciation of housing prices in initially higher housing price neighborhoods. The estimated convergence metro area β's were -0.49 and -0.35, respectively, for Denver and Portland.

Figure 5 reports the estimates for (1) for areas that experienced housing price declines during the 1990-1997 period. The areas examined are: New York City (city only), New York metro area, San Francisco metro area and Boston metro area. The first uses data from the Furman Center. The latter three use the Case-Shiller data. Within New York City (left panel), there is no evidence that housing prices started to diverge (i.e, β was not positive) as housing prices were falling. However, within the New York metro area broadly, within the San Francisco metro area, and within the Boston metro area, there again is week evidence that housing prices started to diverge during the housing price bust (β's were positive). The estimated β's were 0.07, 0.09, and 0.22, respectively, for the New York, San Francisco, and Boston metro areas. All estimates were significant at the 1 percent level. The results seem to indicate that in metro areas - broadly defined - lower priced neighborhoods fell by a greater amount during the bust period than did the higher priced neighborhoods.

Figure 6 is the analog to Figure 3 in that it shows the results from specification (2) for the 1990-1997 time period using the 30 metro areas for which we have Case-Shiller data. The results are striking. Just as in the 2000-2006 period, the amount of convergence (divergence) is positively related to the size of the housing price increases (declines). Areas such as Denver and Portland had big price increases and sizable convergence while areas such as Chicago, Atlanta, and Minneapolis had little price appreciation or depreciation and little convergence or divergence. Areas like San Bernardino, L.A., and San Diego had sizable price declines and some divergence. The R-squared of the scatter plot of change in house prices against the estimated β was 0.61 with the estimated γ being -0.48 (with a bootstrapped standard error of 0.03).

3.3 Within City Movements in Housing Prices: The 1980s

For our last set of results, we examine property price movements across neighborhoods within various metro areas during the 1980s. Given that our Case-Shiller data does not exist for many metropolitan areas prior to 1987, we will primarily use data from the U.S. Census to examine the robustness of our fact during the 1980s.

We start by focusing on New York and Boston - two cities that experienced substantial

housing price booms during the 1980s. Average real housing prices increased in these two cities between 1984 and 1989 by well over fifty percent. In the left panel of Figure 7, we show the estimates from (1) for New York using the Furman data for the period 1984-1989. In the middle panel, we show the estimates from (1) for Boston using the Case-Shiller data for the period 1984-1989. The Boston metro area was one of the few metro areas where Case-Shiller data existed throughout the entire 1980s. In the last panel, we show the data from the U.S. Census - by census tract - for the city of Boston during the 1980 - 1990 period. What we plot in the right-most panel is just the raw change in median house price by census tract between the 1980 and 1990 period. Like the results for the 1990s and 2000s, New York and Boston saw sharp convergence in housing prices across neighborhoods during their 1980s property price booms. The estimated β's from the three panels are -0.38, -0.13, and -1.43, respectively. All convergence estimates are significant at the 1 percent level. Given that the Census data is not holding the quality of the housing stock constant in the price estimates, the magnitude of the housing price increases are much higher. That is why the estimated β from the Census data is so much higher than the estimated β's from the other two specifications.

In Figure 8, we estimate (2) using the Census data for the 1980-1990 period. Our unit of analysis is any metro area that had at least 50 distinct census tracts in the first stage regression where we estimated the β's.[19] There were 41 such metro areas that met our definition. As with the results in Figures 3 and 6, those metro areas that experienced large property price increases between 1980 and 1990 experience substantial convergence in house prices across neighborhoods within the MSA during that same time period.

3.4 Summary of Fact

In this section we have shown a very robust statistical relationship between area wide changes in house prices and the amount of convergence or divergence that takes place across neighborhoods in that area. Specifically, when a local area (either city or metro area) experiences a large property price increase, there is substantial convergence in house prices across neighborhoods. During the housing price boom, it is the initially lower house price neighborhoods that experience substantially higher rates of appreciation compared to the higher priced neighborhoods. The city-wide boom, the more convergence takes place. This fact is found with cities/metro areas in the 1980s, 1990s, and 2000s and is robust to using many different types of price indices

[19] When ensuring that there were 50 distinct census tracts in the metro area, we restricted our analysis to only those census tracts whose boundaries remained intact between 1980 and 1990 and where the Census actually reported a median home price for the district in both 1980 and 1990. See the data appendix for a complete discussion of our sample selection criteria.

to measure changes in prices within neighborhoods. Lastly, there is some evidence that the results are symmetric. When a local area experiences a period of property price declines, there is a divergence in house prices across neighborhoods within the area.

4 Model

In order to rationalize the facts that we have just documented, in this Section, we develop a spacial equilibrium model of housing prices across neighborhoods within a city. The key ingredient of the model is a positive neighborhood externality: people like to live next to more people. In the spirit of Glaeser, Kolko, and Saiz (2001), we assume that there are increasing returns to scale in the production of amenities, so that the more people are around, the greater is the number of restaurants, museums, coffee shops, dry cleaners, etc. First, we present a baseline version the model, that is able to capture our main mechanism. Then, we present two extended versions: a version of the model enriched with adjustment costs and a version where we refine the nature of the externality in order to generate gentrification.

4.1 Baseline model

4.1.1 Set up

Time is discrete and runs forever. We consider a city populated by a continuum of infinitely-lived households of measure N. The city is represented by the real line and each point on the line $i \in (-\infty, +\infty)$ is a different location. Agents are fully mobile and can choose to live in any location i. Denote by $n_t(i)$ the measure of agents who live in location i at time t and by $h_t(i)$ the size of the house they choose. In each location, there is a maximum space that can be occupied by houses normalized to 1, that is,

$$n_t(i)\, h_t(i) \leq 1 \text{ for all } i, t.$$

Moreover, market clearing requires

$$\int_{-\infty}^{+\infty} n_t(i)\, dni = N. \tag{3}$$

There is a positive location externality, as agents like to live in areas where other agents live. Each location i has an associated neighborhood, given by the interval centered at i of radius γ. Let $H_t(i)$ denote the total space occupied by houses in the neighborhood around location i, that is,

$$H_t(i) = \int_{i-\gamma}^{i+\gamma} h_t(j)\, n_t(j)\, dj. \tag{4}$$

Households have separable utility for non-durable consumption c and housing services h. The location externality is captured by the fact that agents enjoy more to live in locations with higher $H_t(i)$. The utility of an agent located in location i at time t is given by $u(c) + v(h, H_t(i))$, where $u(.)$ and $v(.)$ are both weakly concave functions. For simplicity, we assume that u is linear and that v takes the following functional form: $v(h, H) = h^{\alpha} H^{\beta}$. Also, each period agents receive an exogenous endowment of consumption goods equal to y.

On the supply side, there is a representative constructor who can build housing at marginal cost C in any location. The (per square foot) price of a house in location i at time t is equal to $p_t(i)$. Hence there is going to be construction in location i as long as $p_t(i) \geq C$. If there is a location where there is empty space and no construction, the price cannot be higher than C.

Finally, there is a continuum of competitive intermediaries who own the houses and rent them to the households. The intermediaries are only introduced for ease of exposition and nothing would change if we allowed the households to own. The (per square foot) rent for a house in location i at time t is denoted by $R_t(i)$. As long as the rent in location i at time t is positive, the intermediaries find it optimal to rent all the houses in that location. Also, for simplicity, assume that houses do not depreciate. Competition among intermediaries requires that for each location i the following arbitrage equation holds:

$$p_t(i) = R_t(i) + \frac{1}{1+r} p_{t+1}(i). \tag{5a}$$

4.1.2 Equilibrium

An equilibrium is a sequence of rent and price functions $\{R_t(i), p_t(i)\}_{i \in R}$ and of allocations $\{h_t(i), n_t(i)\}_{i \in R}$ such that 1) households maximize their utility, 2) the constructor maximizes his profits, 3) the intermediaries maximize their profits, and 4) markets clear.

At each given time, households decide their non-durable consumption, in what location to live, and the size of their house. Because of full mobility, the household's maximization problem reduces to a series of static problems of the form

$$\max_{c,h,i \in \mathcal{I}_t} c + \phi h^{\alpha} H_t(i)^{\beta}$$

s.t.

$$c + hR_t(i) \leq y,$$

where $H_t(.)$ is taken as given and $\mathcal{I}_t$ denotes the set of locations where there is housing at time t. Conditional on choosing location i, the optimal house size is then given by

$$h_t(i) = (\phi\alpha)^{\frac{1}{1-\alpha}} R_t(i)^{\frac{1}{\alpha-1}} H_t(i)^{\frac{\beta}{1-\alpha}}. \tag{6}$$

Households prefer to rent bigger houses when the rent is lower and the neighborhood is more developed.

Given that households are fully mobile, it must be that at each point in time, the equilibrium rents in different locations make them indifferent. In particular, rents must be higher in locations with more developed neighborhoods. Substituting (6) in the household objective function shows that the net surplus of living in location i is proportional to $R_t(i)^{\frac{\alpha}{1-\alpha}} H_t(i)^{\frac{\beta}{1-\alpha}}$. For households to be indifferent between all locations in $\mathcal{I}_t$ this quantity must be constant. Therefore, the household's optimal location choice implies that the rent function satisfies

$$R_t(i) = K H_t(i)^{\frac{\beta}{\alpha}} \text{ for all } i \in \mathcal{I}_t$$

for some constant K.

As a normalization, let us choose point 0 as the center of the city. Then, in equilibrium the set of developed locations $\mathcal{I}_t$ will be an interval $[-I_t, I_t]$ for some $I_t > 0$. Since the model is symmetric, from now on we can focus on the positive portion of the real line. Given households' preferences, it must be that $h_t(i)\, n_t(i)$ is equal to 1 for all locations in the interval $[0, I_t]$. Then we can solve for $H_t(i)$ as follows

$$H_t(i) = \begin{cases} 2\gamma & \text{for } i \leq I_t - \gamma \\ \gamma + I_t - i & \text{for } i \in (I_t - \gamma, I_t] \\ \max\{0, \gamma - i\} & \text{for } i > I_t \end{cases}. \tag{7}$$

Hence, the equilibrium function $R_t(i)$ must satisfy

$$R_t(i) = \begin{cases} R_t(I_t)\, 2^{\frac{\beta}{\alpha}} & \text{for } i < I_t - \gamma \\ R_t(I_t) \left(1 + \frac{I_t - i}{\gamma}\right)^{\frac{\beta}{\alpha}} & \text{for } i \in (I_t - \gamma, I_t] \end{cases}. \tag{8}$$

Next, combining the market clearing condition (3) with the optimality condition (6) together with the expressions for (7) and (8), one obtains the following equilibrium condition for the rent in the marginal location I_t:

$$R_t(I_t) = \rho(I_t) \equiv \gamma^{\beta} \phi \alpha \left(\frac{N}{2}\right)^{1-\alpha} \left[2^{\frac{\beta}{\alpha}} (I_t - \gamma) + \frac{\alpha\gamma}{\alpha+\beta}\left(2^{\frac{\alpha+\beta}{\alpha}} - 1\right)\right]^{\alpha-1}. \tag{9}$$

Also, the optimization problem of the constructor implies that he constructs in all locations with $p_t(I_t) \geq C$. This implies that the price in the marginal location must be equal to the marginal cost, that is, $p_t(I_t) = C$, given that if it was bigger, by continuity there would be construction in all the locations next to I_t and I_t could not be city boundary.

In equilibrium prices are constant and hence arbitrage conditions (5a) require that for each location i

$$p^*(i) = \frac{1+r}{r} R(i). \tag{10}$$

This immediately implies that rents in the equilibrium marginal location I^* are equal to

$$R^*(I^*) = \frac{r}{1+r}C, \tag{11}$$

and substituting this condition into (8) allows us to specify the rent function as a function only of the size of the city I^*. We can then combine conditions (11) and (9) in order to solve for I^* and obtain

$$I^* = \gamma + \frac{1}{2^{\frac{\beta}{\alpha}}}\left[\frac{N}{2}\left(\frac{r}{1+r}C\frac{1}{\phi\alpha\gamma^{\beta}}\right)^{\frac{1}{\alpha-1}} - \frac{\alpha\gamma}{\alpha+\beta}\left(2^{\frac{\alpha+\beta}{\alpha}}-1\right)\right]. \tag{12}$$

As intuition suggests, the city is bigger when there are more households N and when the marginal cost of construction C or the interest rate r are lower.

To complete the equilibrium characterization, we can use (8) and (10) together with (9) and (11), to solve for the equilibrium, price schedule

$$p^*(i) = \begin{cases} C2^{\frac{\beta}{\alpha}} & \text{for } i < I^* - \gamma \\ C\left(1+\frac{I^*-i}{\gamma}\right)^{\frac{\beta}{\alpha}} & \text{for } i \in (I^*-\gamma, I^*] \end{cases}. \tag{13}$$

In equilibrium, the households are all concentrated in the center of the city because they like to live next to each other. Moreover, equilibrium prices reflect the fact that locations that are further away from the city center and closer to empty space are less appealing. In particular, prices are the highest in the city center where the neighborhoods are fully developed and there is no empty land. As we move away from the center, prices start declining because the neighborhoods start to be less and less developed. At the boundary of the city, prices are equal to the marginal cost of construction given that house supply is perfectly elastic and there is empty land in the location right next to the marginal one.

4.1.3 Demand shock

We are interested in analyzing how prices in different locations react to shocks to the demand for housing. In particular, our baseline model is consistent with the first stylized fact we have documented: in reaction to a demand shock, prices in locations with higher initial price level react less than prices in locations where houses are cheaper to start with. The mechanism is driven by the externality at the core of our model. As a demand shock hits the economy, the city starts expanding and the neighborhoods that are less developed start growing. Due to our externality, the house prices in these neighborhood are driven up. This generates the counterintuitive effect that prices react more in neighborhoods where the house supply is more elastic.

For concreteness, let us analyze the economy's reaction to an interest rate shock. Let us start from an economy at an equilibrium with $r = r^H$. Imagine that at time $t = 0$ the economy is hit by an unexpected shock that permanently reduces the interest rate to $r^L < r^H$.

At the initial equilibrium, the city size I^H and price schedule $p^H(i)$ are given by conditions (12) and (13) with $r = r^H$. When the economy is hit by the shock, it immediately reaches a new equilibrium, with I^L and $p^L(i)$ given by the same conditions with $r = r^L$. From these conditions, it is easy to see that: (1) the city expands, $I^L > I^H$, (2) prices remain constant in the central locations $i \in \left[-I^H + \gamma, I^H - \gamma\right]$, (3) prices strictly increase in the more peripheral locations $i \in \left(-I^H, -I^H + \gamma\right)$ and $i \in \left(I^H - \gamma, I^H\right)$. Notice that initial prices are not well defined for the locations $i \in [-I^L, -I^H) \cup (I^H, I^L]$, given that these locations were not developed before the shock. At the original equilibrium, the shadow price in these locations can be any price less than or equal to C. To be conservative, when we calculate the price change after the shock, we assume that in those location the shadow price is exactly equal to C.

Figure 9 shows the reaction to a positive demand shock (a decrease in the interest rate) of prices in the different locations. Given that the city is symmetric, the picture shows only the positive portion of the real line.

Figure 10 shows the price growth rate as a function of the initial price level in different locations together with the OLS regression that corresponds to the regressions run in the data. Our model delivers a negative slope of the regression line as in our first stylized fact.

Next we want to use our simple model to explore our second stylized fact, that is, that, across cities, the slope of this regression line is steeper the higher is the average growth rate of prices. We consider two different stories which can rationalize this relationship within our model. First, it could be that different cities are hit by demand shocks of different size. Second, it could be that different cities are hit by a common demand shock, but differ in preferences and/or technology. In both cases, our model is able to generate a positive relationship between the degree of gentrification and the growth rate of average house prices. The following to exercise show this.

Figure 11 shows the reaction of house prices in different neighborhoods after demand shocks of different sizes. In particular, the figure shows the house price growth rate in different locations after a decrease of the interest rate from 5% to 4.5% (blue dots) and from 5% to 4% (red dots). The figure also shows the corresponding OLS regression lines (solid lines). It is evident that after a larger demand shock average prices increase more and, at the same time, the regression coefficient is bigger, which delivers our second empirical fact.

Figure 12 shows the reaction of house prices to the same demand shock in cities with different preferences. In particular, the figure shows the house price growth rate in different locations after a decrease of the interest rate from 5% to 4% in a smaller city with lower ϕ (blue dots) and in a more developed one with higher ϕ (red dots). The city with higher ϕ is more developed, for any given interest rate, because households enjoy more living there. The figure also shows the associated OLS regression lines (solid lines), showing that the regression line is steeper in the city that is more developed to start with, which is also the city where prices grow more on average, delivering again our second empirical fact.

Both these stories may be relevant for different cross sections in different time periods. However, we are particularly interested in exploring the second story, to understand, for example, why Charlotte did not react as Chicago or Boston did to the recent aggregate demand shock. In particular, in Section 4.3 we push this line and enrich the formalization of the externality so that cities that are more developed tend to gentrify more and at the same time to experience higher price growth.

4.2 Model with Adjustment Costs

In this section, we extend the previous model, by introducing standard adjustment costs on the supply side. This extension is interesting for two reasons. First, adjustment costs make the dynamics of the model richer. Second, we can compare our model to a standard adjustment cost model and show that the aggregate implications are substantially different.

4.2.1 Equilibrium

The set up of the model is exactly the same as in Section 4.1.1, except that the representative constructor now faces construction costs that increases with the amount of construction at each time t. By market clearing, the amount of construction at time t needs to be equal to the increase in the size of the city, that is, $I_t - I_{t-1}$. Then, the construction cost function is a convex function of $I_t - I_{t-1}$.

The analysis in Section 4.1.2 still goes through, except that now the price in the marginal location I_t is given by

$$p_t(I_t) = c'(I_t - I_{t-1}). \tag{14}$$

In our numerical examples, we set the cost function to be $c(x) = C_1 x + C_2 x^{\phi}/\phi$. Moreover, the arbitrage equation (5a) together with the expression for the rents (8) for any location $i \leq I_t$

yields

$$p_t(i) = \sum_{j=0}^{\infty} \left(\frac{1}{1+r}\right)^j \rho(I_{t+j}) \min\left\{\left(1 + \frac{I_{t+j} - i}{\gamma}\right)^{\frac{\beta}{\alpha}}, 2^{\frac{\beta}{\alpha}}\right\}, \tag{15}$$

where $\rho(I_t)$ is given by equation (9).

Letting $\tilde{p}_t$ denote the price at the boundary location at time t, $p_t(I_t)$, equation (14) and equation (15) evaluated at $i = I_t$ give us a dynamic system that characterizes the equilibrium sequence $\{I_t, \tilde{p}_t\}_{t=0}^{\infty}$. Once we have solved for the sequence $\{I_t, \tilde{p}_t\}_{t=0}^{\infty}$, we can solve for the sequence of rents for any location i, $\{R(i)\}_{t=0}^{\infty}$, using expression (8). Finally, we can solve for the sequence of prices for any location i, $\{p_t(i)\}_{t=0}^{\infty}$, using the arbitrage conditions (5a).

The steady state equilibrium of the model with adjustment costs looks identical to the equilibrium in the baseline model, if we assume $c'(0) = C$. However, now the equilibrium dynamics are interesting and if a shock hits, the economy does not immediately switch to the new steady state, but it slowly converges to it.

4.2.2 Demand shock

We now analyze the same interest rate shock studied in the baseline model. Imagine the economy starts with an interest rate r^H and suddenly switches to an interest rate $r^L < r^H$. We show that, after the shock, the city starts growing slowly and aggregate price overshoot in the short run and then, in contrast with the standard adjustment cost model, go back to a permanently higher level.

As in a standard adjustment cost model, the price in the marginal location I_t overshoots on impact and then declines back to the initial constant level, once the steady state is reached and the city does not grow anymore. However, the dynamics of the prices in all the other locations are more complex as Figure 13 shows.

Panel (a) of Figure 13 shows the price behavior in different locations over time. The dynamics are rich and differ across neighborhoods. Developed neighborhoods experience price dynamics that are qualitatively similar to the standard adjustment cost model, with an initial spike in prices followed by a gradual return to the initial price level. Neighborhoods that develop late in the process do not experience overshooting at all, and their prices display a gradual increasing path toward their long run level. Finally, intermediate neighborhoods show hump-shaped dynamics, with a gradual increase followed by a gradual decrease toward the long run level. Initially these neighborhoods grow and the local externality makes them more desirable, but eventually the city expands reducing the pressure of housing demand and driving prices down. Panel (b) of Figure 13 zooms in on the locations that are closer to the city center,

which were already developed at the initial steady state, and shows that prices in the less developed neighborhoods react more than in the developed ones. This is due to our gentrification mechanism and it is what drives the amplification result.

Let us now compare the aggregate house price behavior in our model with a standard adjustment cost model with no externalities. In the benchmark adjustment cost model, where $\beta = 0$, the rent is going to be the same in all locations and equal to $\alpha (N/2)^{1-\alpha} I_t^{\alpha-1}$. The equilibrium is then simply characterized by equations (14) together with

$$p_t = \sum_{j=0}^{\infty} \left(\frac{1}{1+r}\right)^j \alpha \left(\frac{N}{2}\right)^{1-\alpha} I_t^{\alpha-1}.$$

After a decrease in the interest rate, construction slowly increases to the new steady state level and aggregate price increase on impact and then decline to the same constant initial level. In our model, the average price level is given by

$$P_t = \frac{1}{2I_t} \int_{-I_t}^{I_t} p_t(i)\, di.$$

Here, the price in the marginal neighborhood goes back to the marginal cost $c'(0)$, but the aggregate price level of prices is going to be permanently higher. This is due to the fact that, after the shock, the city starts expanding, so that new neighborhoods are going to be born with positive prices and neighborhoods that were not fully developed are going to grow achieving permanently higher price levels. In the short run, there are two different forces at play. First, prices in the less developed neighborhoods grow faster than in the standard model, because agents living there expect their neighborhood to grow and the neighborhood externality to be stronger. However, the same mechanism speeds up construction in reaction to a housing demand shock, given that agents who live in less developed neighborhoods are willing to pay higher prices. This makes the prices in developed neighborhoods increase less than aggregate price in the standard model, because the housing supply adjusts faster. Different parametrization can generate an amplified or a dampened effect on aggregate prices in the short run.

Figure 14 shows an example where the reaction of aggregate prices to the demand shock is amplified in comparison to the standard model. The figure represents the behavior of aggregate prices after a decline in the interest rate both in our model and in the standard adjustment cost model, showing both the level effect and the amplification effect on aggregate prices.

4.3 Model with heterogeneous households

We now consider an extension of the baseline model where we refine the nature of the externality by assuming that the city has a fixed size and that there are two types of households: rich and

poor. Now rich households like to live next to rich neighbors, but dislike living near poor neighbors. This extension of the model allows us to go deeper in the understanding of our second empirical fact. As we discussed above, the baseline model is able to rationalize the fact that cities with higher housing price growth also display greater neighborhood convergence, by assuming either that different cities are hit by shocks of different size or that different cities are hit by a common shock but react differently because of some form of heterogeneity. This extension of the model allows us to put more structure on the second type of story and gives us more predictions that we test in the next section. In particular, this model implies that the reaction of housing prices is affected by the stage of development of a city which is related to the degree of gentrification. Hence a shock of similar size can have different effects on different cities that are at different stages of development (e.g. New York and Charlotte in the recent price boom), but also on the same city in different time periods.

4.3.1 Set up

Consider a variant of the economy described in Section 4.1.1 which is now populated by two types of infinitely-lived agents: a continuum of rich of measure N^R and a continuum of poor of measure N^P. The city has now a fixed dimension and is represented by a segment $[-\bar{I}, \bar{I}]$ and agents are fully mobile and can live in different locations $i \in [-\bar{I}, \bar{I}]$. Let $n_t^s(i)$ denote the measure of agents of type $s = R, P$ who live in location i at time t and $h_t^s(i)$ the dimension of the house they choose to live in. Market clearing requires

$$\int_{-\infty}^{+\infty} n_t^s(i)\, di = N^s \text{ for } s = R, P. \tag{16}$$

In each location there is a maximum space that can be occupied by houses that is normalized to 1, that is,

$$n_t^R(i)\, h_t^R(i) + n_t^P(i)\, h_t^P(i) \leq 1 \text{ for all } i, t.$$

Locations now differ for the measure of both rich and poor who live in the neighborhood. In particular, we assume the following externality: rich households like to live in neighborhoods not only where more space is occupied by houses of rich households but also where there are less poor households.[20] Let $H_t(i)$ denote the total space occupied by rich households in the interval of radius γ, centered at i, that is,

$$H_t(i) = \int_{i-\gamma}^{i+\gamma} h_t^R(j)\, n_t^R(j)\, dj.$$

[20] We assume that the externality is asymmetric for tractability. The same mechanism would go through if rich households like to live closer to more rich neighbors, rather than to more space occupied by rich.

Also let $B_t(i)$ denote the total measure of poor agents in the same interval,

$$B_t(i) = \int_{i-\gamma}^{i+\gamma} n_t^P(j)\, dj.$$

Households have separable utility for non-durable consumption and housing services. We assume that they enjoy linear utility from consumption, while their utility from housing services is represented by a strictly concave utility function. Moreover, rich households enjoy their housing services more if more space around is occupied by houses of other rich households (higher $H_t(i)$), and less if there are more poor households around (higher $B_t(i)$). More specifically, the preferences of rich households are represented by the utility function $c + \phi^R h^\alpha H(i)^\beta (A + B(i))^{-\eta}$, where $\phi^R, \alpha, \beta, \eta$ and A are positive scalars. The preferences of poor households are simply $c + \phi^P h^\alpha$, where ϕ^P is a positive scalar smaller than ϕ^R. The parameter ϕ^s captures the willingness of the agents to pay for housing services. More developed cities typically have higher ϕ^s both for $s = P, R$. The fact that $\phi^P < \phi^R$ implies that poor agents have smaller willingness to pay for housing than rich ones.

The constructor and the intermediaries behave as in the baseline model, except that rich and poor agents live in different types of houses. The marginal cost to build houses for agents of type s is equal to C^s, with $C^R > C^P$. If the constructor wants to construct a house for rich agents in a space occupied by a house for a poor agent, he has both to convince the poor to leave and to pay C^R. Hence both prices and rents are going to be contingent on the type of house and we denote them with $p_t^s(i)$ and $R_t^s(i)$, the price and the rent, respectively, for a house for agents of type s in location i at time t.

Competition among intermediaries requires that prices satisfy

$$p_t^s(i) = R_t^s(i) + \left(\frac{1}{1+r}\right) p_{t+1}^s(i) \text{ for all } t, i, s. \tag{17}$$

4.3.2 Equilibrium

An equilibrium is now a sequence of rent and price schedules $\left\{R_t^R(i), R_t^P(i), p_t^R(i), p_t^P(i)\right\}_{i \in R}$ and of allocations $\left\{n_t^R(i), n_t^P(i), h_t^R(i), h_t^P(i)\right\}_{i \in R}$ such that 1) households maximize their utility, 2) the constructor maximizes his profits, 3) the intermediaries maximize their profits, and 4) markets clear.

At each time, households decide their non-durable consumption, in which location to live, and the dimension of the house they want to live in, taking as given the rental price and the neighborhoods characteristics. As in the baseline model, because of full mobility, the households' optimization problem reduces to a series of static problems. The problem of a poor household

is simply

$$\max_{c,h,i\in\mathcal{I}_t^P} c+\phi^P h^\alpha,$$
$$s.t.\ c+hR_t^P(i)\leq y,$$

where $\mathcal{I}_t^P$ is the set of locations where houses for poor households are constructed and the rent schedule $R_t^P(i)$ is taken as given. It follows that their optimal house size is

$$h_t^P(i)=\left(\frac{\alpha\phi^P}{R_t^P(i)}\right)^{\frac{1}{1-\alpha}}. \tag{18}$$

Moreover, poor households must be indifferent among living in different locations where houses for poor are constructed, that is, among all $i\in\mathcal{I}_t^P$. All these locations are the same for them, except potentially for the rental price. This immediately implies that the rent they pay must be the same in all the locations where they live, that is, $R_t^P(i)=R_t^P$ for all $i\in\mathcal{I}_t^P$.[21] From equation (18), this also implies that all poor households choose houses of the same size, that is, $h_t^P(i)=h_t^P$.

The problem for a rich household at time t can be written as

$$\max_{c,h,i\in\mathcal{I}_t^R} c+\phi^R h^\alpha H_t(i)^\beta (A+B_t(i))^{-\eta},$$
$$s.t.\ c+hR_t^R(i)\leq y,$$

where $\mathcal{I}_t^R$ is the set of locations where there are houses for rich households, and the household takes as given the functions $H_t(i)$ and $B_t(i)$ and the rental price schedule $R_t^R(i)$. Conditional on choosing to live in location i at time t, the optimal size of a house for a rich household is given by

$$h_t^R(i)=\left(\alpha\phi^R\right)^{\frac{1}{1-\alpha}} R_t^R(i)^{\frac{1}{\alpha-1}} H_t(i)^{\frac{\beta}{1-\alpha}} (A+B_t(i))^{-\frac{\eta}{1-\alpha}}. \tag{19}$$

Rich households choose to live in bigger houses in neighborhoods where the rental price is lower, there are more houses where other rich households live, and there are less poor households around. Moreover, since they have to be indifferent among all the locations where they live, it must be that the rental price compensate them for the differences in locations. In particular, the optimality conditions of the rich households require

$$R_t^R(i)=KH_t(i)^{\frac{\beta}{\alpha}}(A+B_t(i))^{-\frac{\eta}{\alpha}}.$$

that is, rents must be higher in neighborhoods with more houses for rich and fewer poor around.

[21] If there was a location with construction for poor households and no poor living there, the intermediaries would be willing to decline the rent to 0 inducing the poor to move into that location.

In equilibrium the rich households are concentrated in the city center, that is, $\mathcal{I}_t^R = [-I_t, I_t]$ for some $I_t > 0$, while the poor households are squeezed at the margin of the city, that is, $\mathcal{I}_t^P = [-\bar{I}, I_t) \cup (I_t, \bar{I}]$. The extreme assumption that the poor households do not gain any utility from living around rich households simplify the analysis, by ensuring that poor and rich households never live in the same location. However, what is crucial for our mechanism to work is simply that the externality is stronger for the rich than for the poor. Also, the assumption that $C^P = 0$ guarantees that poor households are spread in all the locations with no rich households. This assumption is clearly also disposable. Once again, given the symmetry of the model, from now on let us focus on the positive portion of the real line.

We can now derive the rental price paid by the poor households. Given that the poor households live in locations where there are no rich, it must be that $h_t^P(i)\, n_t^P(i)$ is either 0 or 1 and in particular is equal to 1 for all $i \in (I_t, \bar{I}]$, and hence in these locations n_t^P is also constant and equal to $1/h_t^P$. Using the market clearing condition (16) for the poor households, we can then derive the dimension of their houses

$$h_t^P = \frac{2\left(\bar{I} - I_t\right)}{N^P}. \tag{20}$$

Combining this expression with the optimality condition (18) we then obtain the rental price that the poor households pay at time t, that is,

$$R_t^P = \alpha \phi^P \left(\frac{N^P}{2\left(\bar{I} - I_t\right)} \right)^{1-\alpha}. \tag{21}$$

Next, we can derive the rents paid by the rich households in different locations. Given their preferences and the fact that they live in locations where there are no poor, again it must be that $h_t^R(i)\, n_t^R(i)$ is either equal to 1 or to 0. We can then easily derive the function $H_t(.)$ as follows:

$$H_t(i) = \begin{cases} 2\gamma & \text{for } i \leq I_t - \gamma \\ \gamma + I_t - i & \text{for } i \in (I_t - \gamma, I_t] \\ \max\{0, \gamma - i\} & \text{for } i \in (I_t, \bar{I}] \end{cases}. \tag{22}$$

Using the expression (20) for h_t^P and the fact that $n_t^P = 1/h_t^P$, we can also calculate $B_t(i)$, that is,

$$B_t(i) = \begin{cases} 0 & \text{for } i \leq I_t - \gamma \\ \frac{N^P(i+\gamma-I_t)}{2\left(\bar{I}-I_t\right)} & \text{for } i \in (I_t - \gamma, I_t + \gamma] \\ \frac{\gamma N^P}{\left(\bar{I}-I_t\right)} & \text{for } i \in \left[I_t + \gamma, \bar{I}\right] \end{cases}. \tag{23}$$

We can now rewrite the rental price schedule for the rich households at time t as

$$R_t^R(i) = \begin{cases} R_t^R(I_t)\, 2^{\frac{\beta}{\alpha}} \left[1 - \frac{N^P\gamma}{N^P\gamma + 2\left(\bar{I}-I_t\right)A}\right]^{-\frac{\eta}{\alpha}} & \text{for } i < I_t - \gamma \\ R_t^R(I_t) \left(1 + \frac{I_t - i}{\gamma}\right)^{\frac{\beta}{\alpha}} \left[1 - \frac{N^P(I_t - i)}{N^P\gamma + 2\left(\bar{I}-I_t\right)A}\right]^{-\frac{\eta}{\alpha}} & \text{for } i \in (I_t - \gamma, I_t] \end{cases}. \tag{24}$$

Next, using the market clearing condition (16) for the rich households and the optimality condition (19), we obtain

$$2\int_0^{I_t} \frac{R_t(i)^{\frac{1}{1-\alpha}}}{(\alpha\phi^R)^{\frac{1}{1-\alpha}} H_t(i)^{\frac{\beta}{1-\alpha}} (A + B_t(i))^{-\frac{\eta}{1-\alpha}}} di = N^R,$$

which together with conditions (22)-(24) implicitly define a function $\rho(I_t)$ such that $R^R(I_t) = \rho(I_t)$.

As in the baseline model, in equilibrium prices are constant. Hence, the arbitrage conditions 17 require that prices satisfy

$$p_t^s(i) = \frac{1+r}{r} R_t^s(i) \text{ for all } t, i, s. \tag{25}$$

Combining this with (21) immediately gives the price schedule for the poor households

$$p_t^P = \alpha\phi^P \left(\frac{1+r}{r}\right) \left(\frac{N^P}{2(\bar{I} - I_t)}\right)^{1-\alpha}. \tag{26}$$

Moreover, we know that if the constructor wants to expand the rich neighborhoods by increasing I_t, he must convince the poor who live in the marginal location to move out and then he has to pay the additional cost C^R. That is, in location I_t the price of a rich house must be

$$p_t^R(I_t) = p_t^P(I_t) + C^R.$$

and hence in equilibrium

$$R_t^R(I_t) = \alpha\phi^P \left(\frac{N^P}{2(\bar{I} - I_t)}\right)^{1-\alpha} + \frac{r}{1+r} C^R. \tag{27}$$

Combining this expression with $R_t^R(I_t) = \rho(I_t)$ yields the equilibrium size of the rich neighborhoods I_t, which, together with (24) allows us to fully characterizing the rent schedule $R_t^R(i)$, and then, using (25), the price schedule for the rich households.

In equilibrium the rich households live in the city center, while the poor are located at the boundary of the city. Given that the city has a fixed size, as the city develops and the size of rich neighborhoods increases, the poor are going to be squeezed at the margins. This makes more difficult to further gentrify the marginal neighborhoods where the rich live, because the cost of pushing the poor out increases. This model delivers the result that gentrification is more costly the more developed the city is. This is consistent with the fact that house prices in cities at different stage of development react differently. In particular, prices in more developed cities tend to react more to compensate for the higher cost of gentrification.

4.3.3 Demand shock

Consider two different cities: A and B. City A is more developed and have higher ϕ^P and ϕ^R than city B. Both cities start with an interest rate $r = r^H$ and are hit by a shock that decreases the interest rate to $r^L < r^H$. Both economies switch immediately to a new equilibrium with more rich neighborhoods (higher I) and higher prices both for the poor and the rich. The question is where prices grow faster and how they grow differently in different neighborhoods.

First of all, equation 26 shows that prices in the poor neighborhoods increase in both cities, but they increase more in city A. This implies that in order to expand their neighborhoods the rich in city A must be willing to pay more to evict the poor. Moreover, the poor are going to be more concentrated in each location, so that generates higher bad externality. This implies that the closer the rich live to the poor, the more they have to be compensated for that. As a result, the price schedule in city A not only displays a higher level but also a steeper slope. This generates a stronger negative relationship between the initial level of price in a location and its price growth rate.

Figure 15 shows the price growth rate in different locations as a function of the initial price level for the two cities (dotted lines), together with the OLS regression lines (solid lines). As our second stylized fact testifies, higher price growth is associated with more convergence. In particular, this model predicts that the more developed city is the one that exhibits more gentrification, and hence grow more and experiences an higher degree of neighborhood convergence.

5 Additional Empirical Evidence

The model in the previous section makes additional predictions about both the spatial patterns of housing price changes during housing price booms and the reason that housing prices change so dramatically in lower priced neighborhoods. In this section, we examine the empirical validity of these predictions. We start by exploring the geographic location of the housing price changes. In particular, we ask whether poor neighborhoods next to rich neighborhoods were more likely to experience house price increases. We then examine whether the initial low priced areas that experienced rapid housing price growth showed evidence of gentrification. In the last part of this section, we bring in new data on neighborhood building patterns to explore whether we see more building in the gentrifying neighborhoods. The results in this section offer strong support of the model outlined in the previous section. Large increases in housing demand within a local area induce poorer neighborhoods that neighbor richer neighborhoods to gentrify.

5.1 The Spatial Patterns of Price Movements

The top panel of Figure 16 shows a map of the neighborhoods within the city of Chicago. The underlying map is the same as Panel A of Appendix Figure A2 (discussed above). Using the complete data from the Chicago Tribune, we identify two types of neighborhoods within the city of Chicago. The first set of neighborhoods are the high property price neighborhoods in year 2000. Specifically, these neighborhoods were in the top quartile of all Chicago neighborhoods with respect to average housing prices in year 2000 (as measured by the 2000 U.S. Census). These high initial priced neighborhoods are indexed in Panel A of Figure 16 with darker shading. The second set of neighborhoods includes those neighborhoods that experienced the highest growth in property prices between 2000 and 2006 (as measured by our Chicago Tribune Price Index). Again, we highlight only those neighborhoods in the top quartile of all Chicago neighborhoods with respect to the growth in property prices. On Figure 16, these neighborhoods are shaded light grey. The single neighborhood that falls into both categories, Logan Square, is shaded in black.

The results in Panel A of Figure 16 show the spatial analog to the results shown in Panel A of Figure 1. In this figure, we show that the areas that grew fastest in Chicago were the poor neighborhoods that directly neighbored the initially rich neighborhoods. This is exactly what was predicted in spatial equilibrium model discussed above. During the property price boom of the early 2000s, the rich neighborhoods of Chicago expanded outwards (to the west and the south). Notice, the extreme south side of Chicago (neighborhoods Pullman (50), South Deering (51), East Side (52), West Pullman (53), Riverdale (54) and Hegewisch (55)) did not grow as fast as the neighborhoods directly to the west and south of the initially high priced areas. Yet, the neighborhoods on the extreme south side of Chicago (the non-shaded areas) were similar in terms of demographics in 2000 to the neighborhoods that experienced rapid price appreciation from 2000-2006 (the grey shaded areas). For example, using Census data, the mean income of the 32 south side neighborhoods that did not experience rapid growth from 2000 - 2006 was about $49,000 while the mean income of the 14 southside neighborhoods that did was about $36,000. High growth neighborhoods on the the southside had an average poverty rate of 34%, while other neighborhoods on the southside had an average poverty rate of 19%. The high growth southside neighborhoods were on average 65% African American, while the other southside neighborhoods were on average 54% African American.

In sum, the demographics of the southside neighborhoods that experienced rapid growth from 2000 to 2006 were not wildly different than the demographics of the southside neighbor-

hoods that did not grow rapidly. If anything the neighborhoods that grew were a little bit poorer and more heavily African American. Collectively, for Chicago, our results show that the poor neighborhoods that were close to the rich neighborhoods were much more likely to experience house price gains than the equally poor (or just slightly wealthier) neighborhoods that were not close to the rich neighborhoods.

Our results for Chicago are not unique. The spatial proximity of high price growth neighborhoods to initially high price level neighborhoods is found in nearly all cities where convergence was pronounced. In Panel B of Figure 16, we show similar patterns for New York City using the Furman data. Again, the dark shaded areas are the top quartile of NYC neighborhoods in 2000 with respect to initial average residential prices. The grey shaded areas are the top quartile of NYC neighborhoods that experienced the most rapid growth in residential prices between 2000 and 2006. The black shaded neighborhoods are the ones that were both in the high initial level and the high initial growth. There are 3 such neighborhoods in New York (the far upper west side of Manhattan - Morningside and Hamilton Heights - and the river neighborhoods of Brooklyn - Brooklyn Heights, Carroll Gardens, Gowanus, Park Slope, and Red Hook) during this time period. Like Chicago, however, the patterns show that the rapidly appreciating neighborhoods in New York City during the recent housing price boom were the ones that were directly to the north and west of the initially high priced neighborhoods of Manhattan.

Panel C of Figure 16 shows the same type of results for the city of Charlotte (using the detailed Charlotte transaction data). Similar to Chicago and New York City, Charlotte's high housing price growth areas (zip codes in this case) are adjacent to its initial high price areas.

Appendix Figure A3 shows additional spatial diffusion pictures for other areas. For Panels A - D we use metro areas (instead of cities) using the Case-Shiller data for the 2000-2006 period. We use metro areas because many of the zip codes in the center city are not covered by the Case-Shiller data. Respectively, Panels A - D of Figure A3 shows the diffusion plot for L.A., San Francisco, Washington, and Atlanta. Zip codes within the metro area not covered by the Case-Shiller Index are left blank. Otherwise, the shadings are the same as above. The dark shaded areas are the zip codes within the top quartile of the highest level of prices in 2000. The grey areas are the neighborhoods within the top quartile of housing price growth between 2000 and 2006. The black areas are the neighborhoods that are both in the highest level and which experience the highest growth. Panel E shows the same convergence patterns for New York during the 1984-1989 period using the Furman data. The spatial patterns are consistent across all areas. Collectively, the results show that the growth in prices is highly correlated spatially

as predicted by the model. The areas that grew the fastest were the poorer neighborhoods that initially bordered the rich neighborhoods.

5.2 Did the High Growth Neighborhoods Gentrify?

Within one version of the model, the mechanism by which initially lower priced neighborhoods experienced a rapid increase in home prices is through gentrification. A demand shock for housing induces richer households to move outward to the fringe of the high priced neighborhoods thereby expanding the high priced neighborhoods into the lower price neighborhoods. As the rich people move into the formerly poorer neighborhoods, the positive externality from living in that neighborhood increases, driving up the equilibrium price of land in that neighborhoods.

To test for evidence of this version of the model, we examine whether the neighborhoods that experienced rapidly growing home prices actually gentrified. To answer this question, we need very detailed demographic and income data for neighborhoods within a city. We use two types of data to perform this analysis. For the 2000-2006 period, we use statistics on average income by zip code compiled from tax records by the IRS.[22] For earlier periods, we use data from the 1980, 1990, and 2000 U.S. Census. The Census data is ideal given that it provides a variety of demographic and income measures at low levels of aggregation. The IRS data provides consistent measures of the following income variables for all zip codes in the country for the years 1998, 2001, 2002, and 2004 - 2006: the number of returns filed from people living in the zip code, the sum of total adjusted gross income (AGI) for the entire zip code, and the number of filers in the zip code with AGI under \$10,000, between \$10,000 and \$25,000, between \$25,000 and \$50,000 and over \$50,000.[23] When measuring potential gentrification during the 2000-2006 period, we use the IRS data from 1998 - 2006 period. We start in 1998 because it is the first year of data we can analyze prior to 2000.[24]

To analyze whether a potential neighborhoods experiences signs of gentrification, we estimate the following regression:

$$y_{t,t+k}^{i,j} = \omega_0 + \omega_1 \overset{g}{HP}{}_{t,t+k}^{i,j} + \delta X_t^{i,j} + \mu^j + \nu_{t,t+k}^{i,j} \tag{28}$$

where $y_{t,t+k}^{i,j}$ is some measure of gentrification in neighborhood i of city/metro area j during the period t to $t+k$, $\overset{g}{HP}{}_{t,t+k}^{i,j}$ is the growth rate in house prices within the neighborhoods during t and $t+k$ (as defined as in section 3), $X_t^{i,j}$ is a vector of controls for neighborhood i of city/metro

[22] See the Data Appendix for details of the data.

[23] The IRS data also includes information about salaries and wages, charitable contributions, and the number of Schedule C, F, and A returns. We do not use such measures in our analysis.

[24] Our results are unchanged if we focus on the 2001 - 2006 period.

area j during period t and μ^j is a vector of city/metro area fixed effects. In this sub-section, we explore whether various measures of gentrification are related to changes in house prices during the 2000-2006 period, the 1990-2000 period, and the 1980-1990 period. The key to this analysis is defining potential measures of gentrification.

Table 1 shows the results of the estimation of (28) for the pooled sample of metropolitan areas covered by the Case-Shiller data. For the results of in Table 1, we focus on the 1998 to 2006 period and use the IRS data to define our measures of gentrification.[25] For our measure of gentrification, we focus both on the change in the fraction of low income households living in the zip code between 1998 and 2006 (columns 1 and 2) and the change in the fraction of high income households living in the zip code between 1998 and 2006 (column 3). We define the fraction of "low income" households in a zip code in two ways: the fraction of households who reported having their AGI being below $10,000 (column 1) and the fraction who reported having their AGI being below $25,000 (column 2). We define the fraction of "high income" households as being the fraction of households who reported having their AGI being above $50,000. Our measure of $\overset{g}{HP}{}^{i,j}_{t,t+k}$ is the change in house prices at the zip code level as measured by the Case-Shiller index.

In the first row of Table 1, we estimate (28) omitting the vector of additional controls. The results show that neighborhoods that experienced substantial price appreciation between 2000 and 2006 also experienced a dramatic decline in the fraction of poorer households living in the neighborhood. In particular, a 100 percent increase in house prices within a neighborhood was associated with a 4.9 percentage point decline in the fraction of households who had AGI less that $10,000 per year and was associated with a 7.4 percentage point decline in the fraction of households who had AGI less than $25,000 per year. For the pooled data across the metropolitan areas covered by Case-Shiller, roughly 20 percent of all respondents had AGI less than $10,000 and roughly 45 percent had AGI less than $25,000. Also seen from Table 1, neighborhoods that experienced rapid house price growth experienced a slight increase in the fraction of households in the neighborhood that had AGI above $50,000 per year (column 3).

In rows 2 and 3 of Table 1, we include controls to proxy for initial differences in average AGI across the different neighborhoods. Specifically, in row 2, we include controls for the initial share of the variable in question. For example, in column (1), we include the share of households with AGI less than $10,000 in neighborhood i in 1998. Similarly for columns (2) and (3), we include the fraction of households with AGI less than $25,000 and the fraction of households

[25] We start in 1998 given that was the first year that the IRS had income data prior to 2000.

with AGI more than \$50,000, respectively. In row 3, we include 4 separate 1998 neighborhood level controls in each regression: the fraction of households with AGI less than \$10,000, the fraction of households with AGI less than \$25,000, the fraction of households with AGI above \$50,000 and the mean level of AGI in the neighborhood. As seen from these results, even controlling for initial differences in the level of income in these neighborhoods, neighborhoods that experienced high growth rate in house prices experienced a non-trivial decline in the fraction of poor households living in the neighborhood.

In column 4 of Table 1, we tried a different measure of gentrification: the percentage increase in average income within the zip code. As seen from the table, this measure is less correlated with changes in house prices. In hindsight, this is not surprising given the large divergence in income experienced between wealthier households and either median or poorer households from the mid-1990s through the mid-2000s[26]. The wealthy households who lived in the initially high priced areas in 2000 got much richer during the 2000-2006 period. So, even if some of those people moved into neighboring areas, the increase in average income in the high priced neighborhoods could still be large relative to the increase in income in the gentrifying neighborhoods.

To strengthen the results in Table 1, we re-estimate (28) on potentially gentrifying neighborhoods from earlier periods. The benefit of using the earlier periods is that we can then use more diverse and more localized measures of gentrification. The results are shown In Table 2. The four gentrification measures explored in Table 2 (i.e., our measures of $y^{i,j}_{t,t+k}$) are: (1) the percentage change in median income, (2) the percentage point change in the poverty rate, (3) the percentage point change in the fraction Black, and (4) the percentage point change in the vacancy rate. In rows (1) - (3), our measure of $\overset{g}{HP}{}^{i,j}_{t,t+k}$ is the percentage change in the Case-Shiller index between 1990 and 2000. As a result, we again restrict our analysis to only those zip codes covered by Case-Shiller. In rows (1) and (2), we look at gentrification patterns within the cities of Denver and Portland, respectively, during the 1990s. Again, these were the two cities that experienced the greatest appreciation in prices during the 1990s. In row (3), we examine all metro areas covered by Case-Shiller during the 1990s. For the regressions reported in this row, we estimate (28) pooling together all the metro areas and including metro area fixed effects. Consistent with the results from Table 1, neighborhoods that appreciated the fastest seemed to gentrify. Relative to other areas within the city, high house price appreciation neighborhoods experienced higher growth in median income, a bigger decline in the poverty rate, a fall in the fraction of neighborhood populated by Blacks, and a fall in the vacancy rate.

[26]See, for example, Piketty and Saez (2007) and Kaplan and Rauh (2009)

Rows (4)-(5) are analogous to row (3) except our measure of $\overset{g}{HP}{}^{i,j}_{t,t+k}$ is now the percentage change in median census tract house prices (using Census data). In row (4), we examine the relationship between our measures of gentrification and the change in median housing prices between 1990 and 2000 within census tracts of a metropolitan area. In row (5), we restrict our analysis to the 1980-1990 period. As with the results in Figure 8, we only include metro areas that have at least 50 consistently measured census tracts between 1980 (1990) and 1990 (2000) with non-missing house price information. Again, we find that neighborhood price appreciation within a city is correlated with observable measures of gentrification. For all results in Table 2, we exclude the vector of additional controls (X). However, we re-estimated all the specification including the initial level of the variable being examined as a control. The results were quantitatively similar.

5.3 Building Within Gentrifying Areas

In our last empirical sub-section, we explore building patterns within gentrifying neighborhoods. There is no systematic collection of building behavior at the neighborhood level across multiple cities for a long time period. As a result, in this sub-section, we restrict our analysis to the neighborhoods within Chicago during the 1993-2004 period. We use this time period because the city of Chicago maintained a database with all Chicago building permits issued during that time period. For each permit in the database, it was recorded whether or not the permit was for residential housing (as opposed to non-residential). For residential properties, the number of units in the proposed building was also tracked. Lastly, we only focus on permits for "new construction'. At this point we ignore all permits for additions, alterations, repairs, and demolition. Each permit is associated with an address. We use the addresses to map the permits into the 77 Chicago neighborhoods. (These are the same neighborhoods as shown in Panel A of Appendix Table A2).

Using the permit data, we make measures of gross building activity within each neighborhood for each year. The gross number of units added to a neighborhood in each year is defined as the total units planned to be built from permits issued in a given year using the information from the "new construction permits". See the Data Appendix for a complete discussion of the Chicago permit data and for our construction of gross new units added by neighborhood.[27]

[27]For ease of exposition, we are going to refer to the amount of planned units built from the permit data as being actual building behavior. We are well aware of the possibility that not all building permits will be executed. However, such differences between building and actual construction are typically small. For example, for the U.S. as whole between the 1999 and 2004 period, the U.S. Census Bureau estimated that 97.5 percent of all residential building permits resulted in a housing start (i.e., construction started) and 94 percent of permits resulted in a housing completion (i.e., construction was completed). See `http://www.census.gov/const/www/`

In Figure 18, we show the building patterns across three groups of Chicago neighborhoods for the 1993 - 2004 period. The three groups of neighborhoods are the same as the three groups of neighborhoods discussed in Panel A of Figure 16: the high initial price neighborhoods in 2000, the high growth neighborhoods between 2000 and 2006, and all other neighborhoods. The solid line is for the initially high priced neighborhoods, the long dashed line is for the high growth neighborhoods and the dotted line is for the other neighborhoods. The left hand panel shows the level of permits issued in each year and the right hand panel shows the share of all permits during a given year that were issued in each of the three groups of neighborhoods. For example, in 1993, roughly 900 new units were added to the high priced region while only roughly 250 and 500 new units were added to the high growth and other neighborhoods respectively. As seen from Panel B, this implies that in 1993, 47 percent were built in the high price neighborhoods, 21 percent were built in the high growth neighborhoods, and 32 percent were build in the other neighborhoods.

Throughout the entire time period, most of the gross addition of units in Chicago occurred within the initially high priced neighborhoods. However, as seen from Panel B, there was a rapid shift in the building patterns within the City of Chicago between 1998 and 2004. In the high growth neighborhoods, the share of gross units built went from roughly 15 percent (in 2000) to over 30 percent (in 2004). As these neighborhoods experienced rapid price increases, these neighborhoods experienced gross additions to their housing stock.

6 Discussion

The model and empirical results have implications for many related literatures. In this section, we briefly discuss the relationship of our results to these literatures.

There is a well developed literature using spatial equilibrium models to explain differences in house prices within a metropolitan area. See, for example, Alonso (1964), Mills (1967) and Muth (1969). These models focus on the trade-off between transportation costs (measured in distance to the center city business district) and house prices. A separate literature using spatial equilibrium models developed to explain price differentials across metropolitan areas. See, for example, Rosen (1979) and Roback (1982). In these models, places are distinguished by the amenities they provide and by the wages paid to workers. In high amenity cities (e.g., warm weather cities), wages are low or housing prices are high.[28]

We contribute to this broad literature by introducing neighborhood externalities into a spa-

`nrcdatarelationships.html` for more details.

[28] Glaeser (2007) offers a thorough survey of this literature.

tial equilibrium model. The importance of neighborhood externalities has emphasized recently. For example, Glaeser, Kolko, and Saiz (2001) present evidence showing that cities are centers of consumption and that bigger cities offer more consumption opportunities. They conclude that "the role of urban density in facilitating consumption is extremely important and understudied." Rossi-Hansberg, Sartre, and III (2009) use data from revitalization programs in Richmond to show that residential externalities are important. They show that the neighborhoods adjacent to the areas that were targeted by the program experienced considerable increases in land value relative to similar sites in a control neighborhood. Becker and Murphy (2003) present a model of neighborhood sorting with two types of individuals in two fixed size neighborhoods. In the Becker-Murphy model, all individuals prefer to live around one of the types. In this set up, neighborhoods provide an externality to individuals to the extent that there is a greater share of good types residing in the neighborhood.

The innovation of our paper is to embed these neighborhood externalities into a standard spatial equilibrium model and show how such preference affect the evolution of house prices across various types of neighborhoods with respect to shocks to housing demand. Given recent housing price dynamics, there has been a renewed interest in understanding why house prices cycle and why the size of the cycle differs across different locations. A prominent explanation to explain why housing prices increased more in one area relative to another for a similar sized demand shock is that different areas have different short-to-medium run housing supply elasticities. In areas where supply can adjust easily, the effect of demand shocks on prices are mitigated via increased building. In areas where it is harder or more costly to adjust supply, increased demand shocks will raise the equilibrium price of housing. If long run housing supply elasticities are bigger than short run housing supply elasticities, an increase in housing demand followed by the subsequent housing supply adjustment will lead to equilibrium housing price cycles.[29]

What are these potential short run supply constraints and why do they differ across locations? Through a series of papers, Glaeser shows that strict regulations on new building is an important deterrent to housing supply adjustments.[30]. In a recent paper, Saiz (2009) shows that water barriers and the steepness of land gradients, in addition to regulation, serve as binding barriers to the adjustment of housing supply. Gyourko (2009) provides a detailed survey of barriers to supply adjustment. Our paper adds to the literature of why housing prices can

[29]For a formalization of these concepts, see Topel and Rosen (1988) and Glaeser and Gyourko (2006).

[30]See, for example, Glaeser and Gyourko (2003), Glaeser, Gyourko, and Saks (2005b), Glaeser, Gyourko, and Saks (2005a), and Glaeser and Ward (2009)

evolve differentially within a given area for a given demand shock. The ability to expand neighborhoods and still maintain neighborhood externalities acts like a barrier to adjusting supply. As seen from section 4, even in a model with no adjustment costs, housing demand shocks can have a sustained effect on prices in a neighborhood depending on the concentration of poor people in neighborhoods adjacent to the rich neighborhoods. Additionally, we show that the existence of positive neighborhood externalities can amplify housing price dynamics relative to a similar model with standard adjustment costs.

In a separate literature, Gyourko, Mayer, and Sinai (2006) develop a model to explain the phenomenon of "superstar cities" - cities that had house price appreciation rates well above the national average. The cornerstone of their model is there are certain locations (in fixed supply) that have amenities that are more desirable than other the amenities in other locations. Given these high amenity locations are in fixed supply, rising national incomes raise the prices in these locations as high income individuals compete to live in these locations. Our model suggests another reason for such superstar cities (which could complement the results in Gyourko, Mayer, and Sinai (2006)). Our model predicts that the amenities in these superstar cities will actually grow over time due to the neighborhood externalities. As more rich people move into the superstar cities, the amenities provided from living around more rich people could put additional upward pressure on housing prices in these cities. Even though the discussion in our paper focused on explaining cross neighborhood housing prices within a city over time, the model could be augmented naturally to explain cross city differences in housing prices over time.

Our paper also speaks to the large literature on the gentrification of urban areas. Recent work has discussed the role of the following in explaining gentrification: the increased consumption benefits from living in a city (Glaeser, Kolko, and Saiz (2001)), the age of a city's housing stock (Rosenthal (2008)) and Brueckner and Rosenthal (2008)) and direct public policy initiatives via community redevelopment programs (Busso and Kline (2007), Rossi-Hansberg, Sartre, and III (2009)).[31]. Our work adds to this literature by showing how generic shocks to housing demand within a city can result in the gentrification of neighborhoods.

Despite the broad literature on gentrification, very little work emphasizes the importance of spatial dependence - either theoretically or empirically - in predicting the spatial patterns of gentrification. There are two notable exceptions. Brueckner (1977) finds that urban neighborhoods in the 1960s that were in close proximity to rich neighborhoods got relatively poorer between 1960 and 1970 (as measured by income growth). Conversely, Kolko (2007) finds the

[31]For a recent survey of the theoretical and empirical work on gentrification, see Kolko (2007)

opposite pattern. Kolko finds that 1990 neighborhoods bordering richer neighborhoods grew faster between 1990 and 2000 than 1990 neighborhoods bordering poorer neighborhoods.

Our model reconciles both of these results. During periods of declining housing demand in urban areas (like the suburbanization movement during the 1960s), neighborhoods within an urban area will diverge - with the divergence being most pronounced for the areas neighboring the high priced areas. Conversely, during urban renewals (like what was witnessed during the 1990s), neighborhoods within the urban area will converge - with the convergence being most pronounced for areas neighboring the high priced areas. In short, our model makes very specific predictions about the spatial patterns of gentrification when urban centers are faced with either positive or negative housing demand shocks.

Very few existing studies have examined within city (or metropolitan area) price movements. Of those that exist, most perform their within city analysis by comparing the appreciation rates of "high end" properties to the appreciation rate of "low end" properties. Moreover, these studies tend to focus their analysis on a given market (or a very small set of markets) during a given time period.[32] Two papers, however, look specifically at differences in house prices appreciation across zip codes within a metropolitan area. Case and Mayer (1996) look at differential movements in prices within cities of the Boston metro area between 1982 and 1992 while Case and Marynchenko (2002) look at differential trends in prices across different zip codes within the Boston, Chicago, and Los Angeles metro areas during the 1983 to 1993 period. No systematic relationships emerged from these studies.

Recently, Mian and Sufi (2009) documented that the zip codes within the U.S. that had increased access to subprime mortgages during the recent period also experienced higher levels of property price appreciation during this period. They took this finding as additional evidence that the housing demand shock experienced during 2000 and 2006 was larger for poorer (subprime) households than it was for richer households. As poor households suddenly received access to credit, they bid up the price of properties in poorer neighborhoods (where they lived) relative to rich neighborhoods. While our paper does not speak directly to the differential demand shock experienced between rich and poor households during recent periods, our spatial equilibrium model with neighborhood externalities shows that different prices movements between rich and poor neighborhoods cannot - by itself - be used as evidence of different demand shocks for rich and poor households. In terms of our model, we show that even in a world

[32] See, for example, Mayer (1993) and Poterba (1991) who analyzed Oakland, Dallas, Chicago, and Atlanta between 1970 and the mid 1980s, Case and Shiller (1994) who analyzed Boston and Los Angeles during the 1980s, and Smith and Tesarek (1991) who analyzed Houston during the 1970s and 1980s.

with a common housing demand shock (across people with different incomes), initially poor neighborhoods will appreciate at a higher rate as richer households expand out of their original neighborhoods into newer neighborhoods. Given all the empirical work we present in this paper, it is likely that the reason prices appreciated more in initially low priced neighborhoods during the recent period was do, at least in part, to the mechanisms presented in our paper.

7 Conclusion

In this paper, we have focused our attention on within city price movements to learn about the nature and extent of housing price dynamics. We start by provide a variety of empirical data showing the relationship between the initial level of housing prices and the extent to which housing prices appreciate during housing price booms and busts. We find that during housing price booms, initially low priced neighborhoods appreciate at a much greater level than do initially high priced neighborhoods. This fact is robust across the use of different housing price series and across different time periods. We also find that the there is more convergence across neighborhoods when the city as a whole experiences a larger housing boom. In other words, larger city wide housing booms are associated with a greater convergence of neighborhoods within the city. We also find some evidence for the converse. Larger city wide housing busts are associated with a greater divergence of neighborhoods within the city.

We then present a spatial equilibrium model of property price movements across neighborhoods within a city to explain the above facts. There are two key features of our model. First, we analyze a linear city. Second, we model within city neighborhood externalities. In particular, we assume that local amenities are increasing in the number of people one lives around. The more people in the local area, the greater the number and the greater the variation in services or public goods (restaurants, museums, coffee shops, dry cleaners, etc.) that can be provided. In some of our specifications, we further refine the local agglomeration effects by making them being positively related with the number of "rich" people in the neighborhood and negatively related to the number of "poor" people. In doing this, we are capturing the fact that density - per se - does not yield better amenities but, it is the density of a certain type of people that generates the positive consumption amenities. The importance of urban density in facilitating neighborhood externalities has been recently emphasized in the work of Glaeser, Kolko, and Saiz (2001) and in Becker and Murphy (2003). Our innovation is to embed these externalities into a model of neighborhood development within a city and show how such preference affect the evolution of house prices across various types of neighborhoods with respect to shocks to

housing demand.

The existence of such externalities changes the nature of housing price dynamics in a variety of situations. Aside from predicting the patterns documented in the first part of the paper (cross neighborhood convergence when faced with a positive shocks to housing demand), in the presence of adjustment costs, our model generates persistently higher level of city-wide housing prices in the long run and may generate an amplified reaction of aggregate prices in the short run. In particular, we explore an extension of the model with gentrification. We show that cities at a more advanced stage of development feature more gentrification in reaction to similar demand shocks, hence experiencing bigger housing price booms and more neighborhood convergence. This suggests not only that different cities may react differently to a common demand shock, but also that the same city, at different stages of its development, may experience differential housing price responses to housing demand shocks of similar size. The reason for this is that as the rich neighborhoods expand, the poor get condensed into the remaining neighborhoods. If the rich have a distaste of living with the poor, it becomes more difficult to expand the rich neighborhood the more concentrated the poor become.

In summary, the existence of neighborhood externalities, provide a new set of predictions of how a given sized demand shock will 1) affect housing prices across neighborhoods within a city, 2) affect housing prices across different cities, and 3) affect housing prices within a city at different stages of its development.

In the last part of the paper, we test the predictions of our spatial equilibrium model in a variety of different ways. All empirical tests support the predictions of the model. In particular, we find strong evidence of the spatial patterns predicted by the model in the sense that it is the poor neighborhoods directly next to the rich neighborhoods that increased their prices most. Additionally, we find evidence that these high growth neighborhoods actually gentrified.

One area on which our paper is silent is the welfare implications across individuals of shocks to housing demand (such as a decrease in interest rates). In our model with bounded cities and both rich and poor agents, the poor are strictly worse off from the reduction in interest rates. A full model – potentially incorporating commuting costs – could be used to quantitatively assess the welfare reduction of the poor from a reduction in interest rates such that it stimulates a large increase in housing demand. Although we did not attempt such an analysis in this paper, we feel this is an important area for future research.

References

Alonso, William. 1964. *Location and Land Use*. Harvard University Press.

Becker, Gary and Kevin Murphy. 2003. *Social Economics: Market Behavior in a Social Environment*. Harvard University Press.

Brueckner, Jan K. 1977. "The Determinants of Residential Succession." *Journal of Urban Economics* 4 (1):45–59.

Brueckner, Jan K. and Stuart S. Rosenthal. 2008. "Gentrification and Neighborhood Housing Cycles: Will America's Future Downtowns be Rich?" .

Busso, Matias and Patrick Kline. 2007. "Do Local Economic Development Programs Work? Evidence from the Federal Empowerment Zone Program." Unpublished.

Case, Karl E. and Maryna Marynchenko. 2002. "Home Price Appreciation in Low-and Moderate-Income Markets." *Low-income homeownership: Examining the unexamined goal* :239–256.

Case, Karl E. and Christopher J. Mayer. 1996. "Housing Price Dynamics within a Metropolitan Area." *Regional Science and Urban Economics* 26 (3-4):387–407.

Case, Karl E. and Robert J. Shiller. 1989. "The Efficiency of the Market for Single Family Homes." 79:125–137.

———. 1994. "A Decade of Boom and Bust in the Prices of Single-Family Homes: Boston and Los Angeles, 1983 to 1993." *New England Economic Review* :40–40.

Foote, Christopher L., Kristopher Gerardi, and Paul S. Willen. 2008. "Negative equity and foreclosure: Theory and evidence." *Journal of Urban Economics* .

Glaeser, Edward L. 2007. "The Economic Approach to Cities." NBER Working Paper 13696.

Glaeser, Edward L. and Joseph Gyourko. 2003. "The Impact of Building Restrictions on Housing Affordability." *Economic Policy Review* 9 (2).

———. 2006. "Housing Dynamics." NBER Working Paper 12787.

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

———. 2005b. "Why Is Manhattan So Expensive? Regulation and the Rise in Housing Prices." *The Journal of Law and Economics* 48 (2):331–369.

Glaeser, Edward L., Jed Kolko, and Albert Saiz. 2001. "Consumer city." *Journal of Economic Geography* 1 (1):27.

Glaeser, Edward L. and Bryce A. Ward. 2009. "The Causes and Consequences of Land Use Regulation: Evidence from Greater Boston." *Journal of Urban Economics* 65 (3):265–278.

Gyourko, Joseph. 2009. "Housing Supply." *Annual Review of Economics* 50 (1):295–318.

Gyourko, Joseph, Christopher J. Mayer, and Todd M. Sinai. 2006. "Superstar cities." NBER Working Paper 12355.

Kaplan, Steven N. and Joshua Rauh. 2009. "Wall Street and Main Street: What Contributes to the Rise in the Highest Incomes?" *Review of Financial Studies* .

Kiyotaki, Nobuhiro, Alexander Michaelides, and Kalin Nikolov. 2008. "Winners and Losers in Housing Markets." Centre for Dynamic Macroeconomic Analysis, CDMA Conference Paper Series, Number 0705.

Kolko, Jed. 2007. "The Determinants of Gentrification." .

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

Mian, Atif and Amir Sufi. 2009. "Forthcoming. The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis." *Quarterly Journal of Economics* .

Mills, Edwin S. 1967. "An Aggregative Model of Resource Allocation in a Metropolitan Area." *The American Economic Review* :197–210.

Muth, Richard F. 1969. *Cities and Housing.* Univ. of Chicago Press.

Piketty, Thomas and Emmanuel Saez. 2007. "Income and Wage Inequality in the United States, 1913-2002." *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-speaking Countries* :141.

Poterba, James M. 1991. "House Price Dynamics: The Role of Tax Policy and Demography." *Brookings Papers on Economic Activity* :143–203.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *The Journal of Political Economy* 90 (6):1257.

Robert E. Lucas, Jr and Esteban Rossi-Hansberg. 2002. "On the internal structure of cities." *Econometrica* :1445–1476.

Rosen, Sherwin. 1979. "Wage-based Indexes of Urban Quality of Life." *Current issues in urban economics* 3.

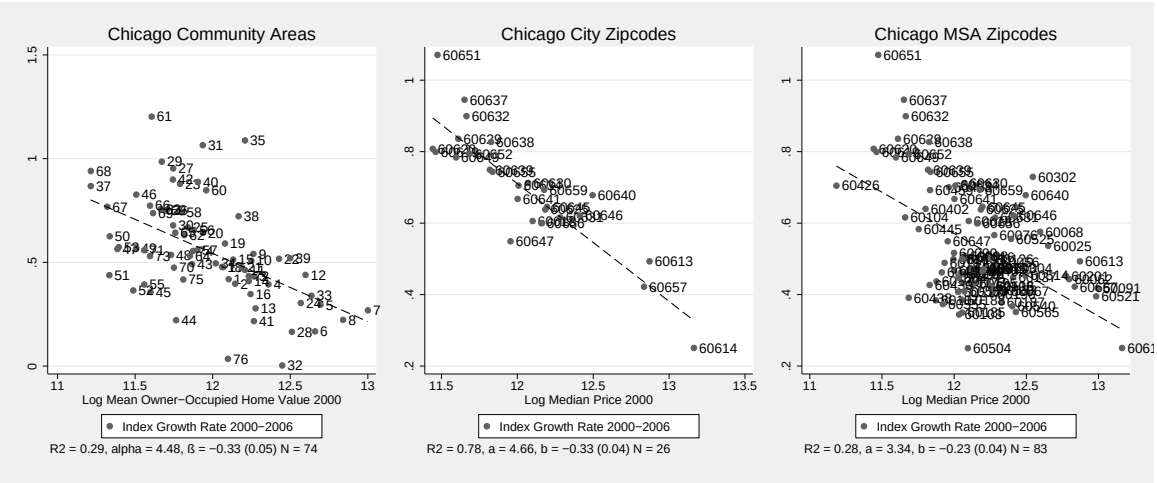
Rosenthal, Stuart S. 2008. "Old homes, Externalities, and Poor Neighborhoods. A Model of Urban Decline and Renewal." *Journal of Urban Economics* 63 (3):816–840.

Rossi-Hansberg, Esteban, Pierre-Daniel Sartre, and Raymond Owens III. 2009. "Firm Fragmentation and Urban Patterns." *International Economic Review* 50 (1):143–186.

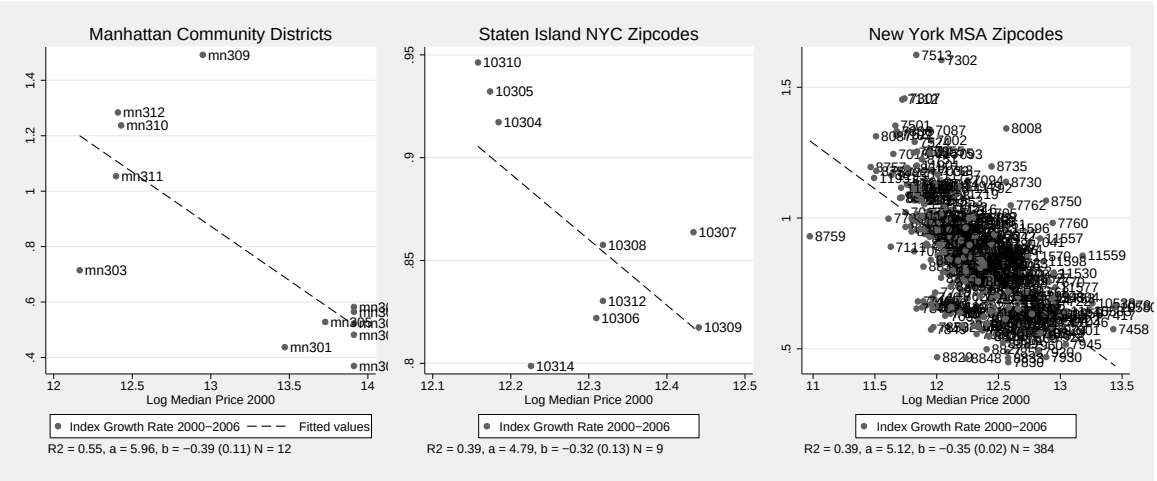
Saiz, Albert. 2009. "Forthcoming. On Local Housing Supply Elasticity." *Quarterly Journal of Economics* .

Smith, Barton A. and William P. Tesarek. 1991. "House Prices and Regional Real Estate Cycles: Market Adjustments in Houston." *Real Estate Economics* 19 (3):396–416.

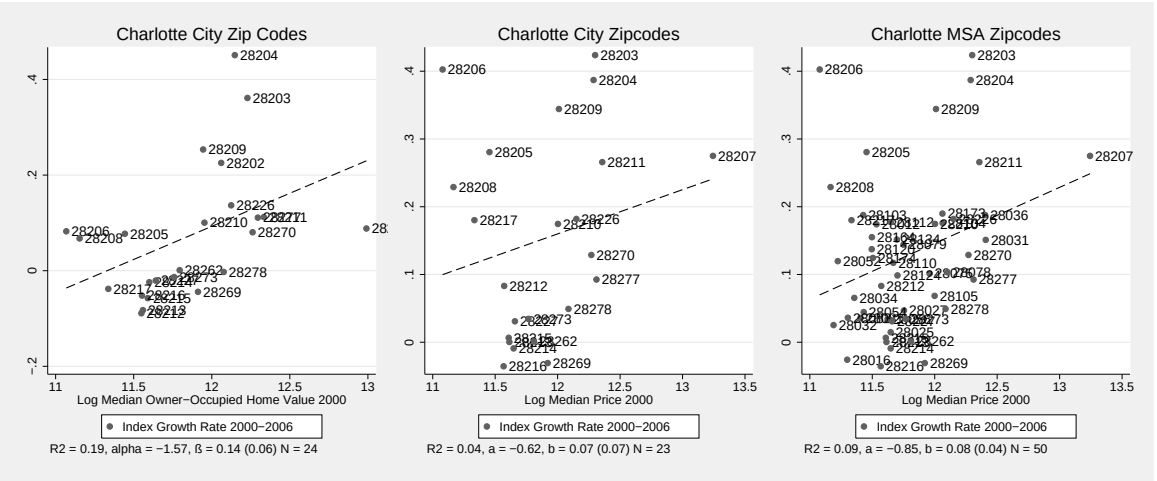
Topel, Robert and Sherwin Rosen. 1988. "Housing Investment in the United States." *The Journal of Political Economy* :718–740.



A: Chicago

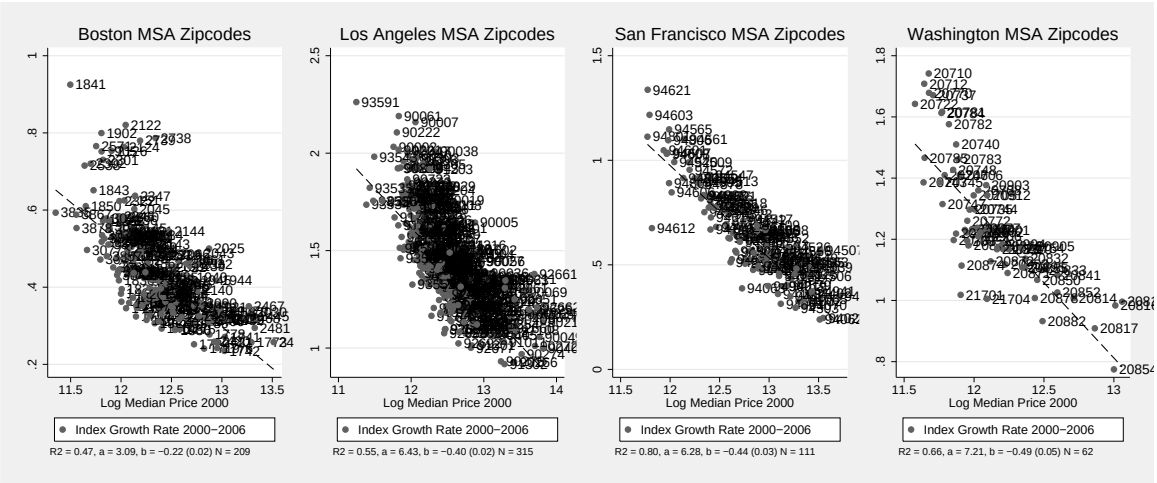


B: New York

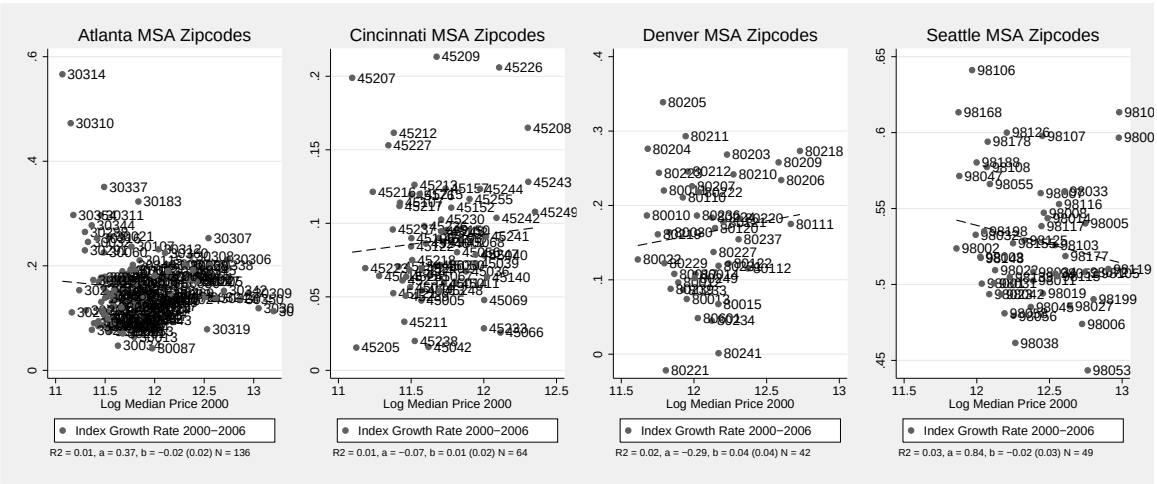


C: Charlotte

Figure 1: Convergence or lack of: 2000-2006

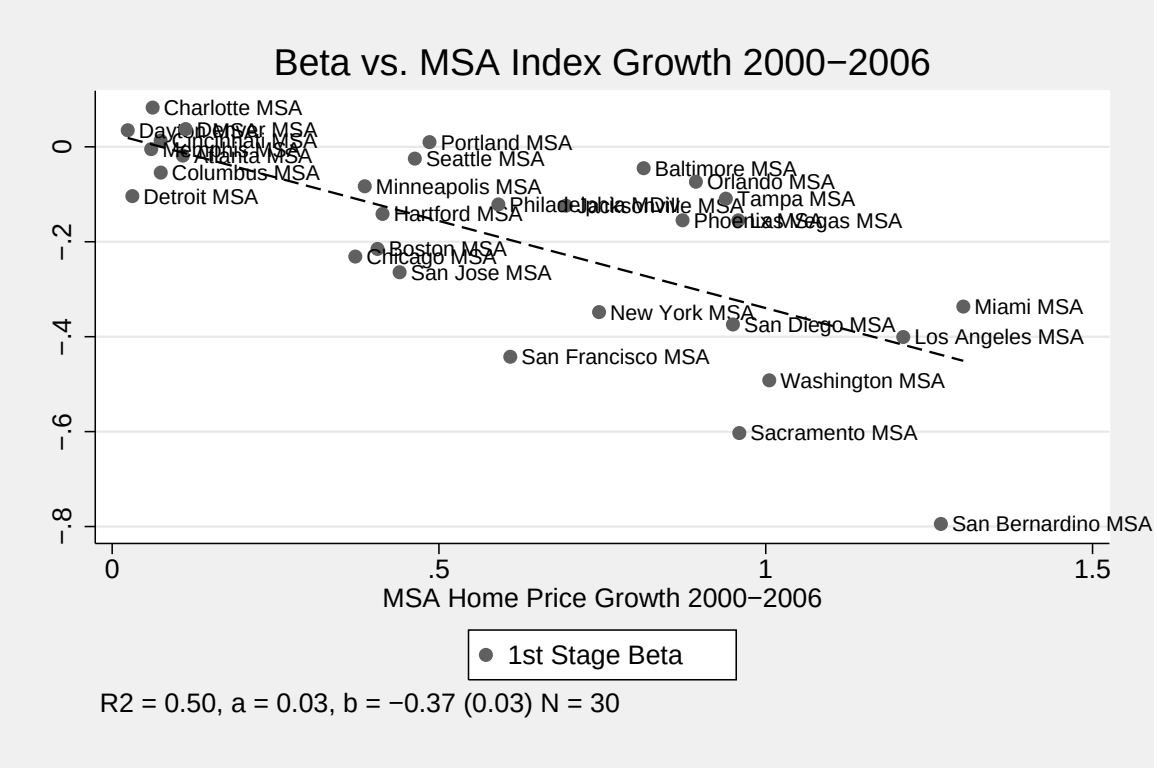


A: Other cities that had strong convergence.

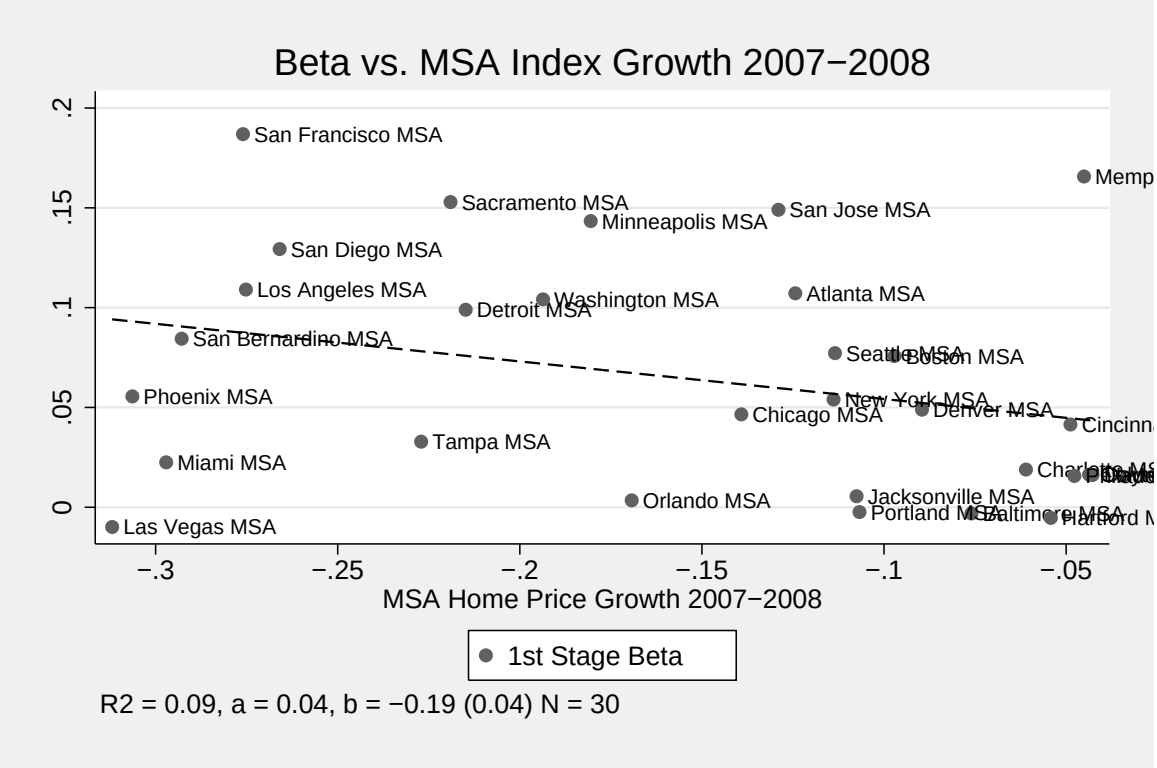


B: Other cities that weaker convergence.

Figure 2: Other Metro Areas using Case-Shiller Index by Zip code 2000-2006

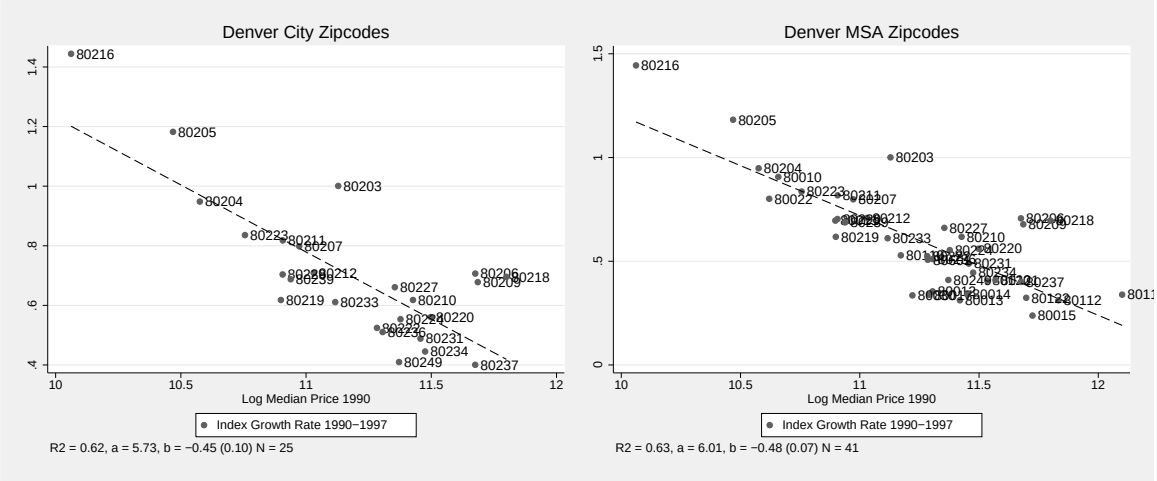


A: Beta vs. Growth 2000-2006

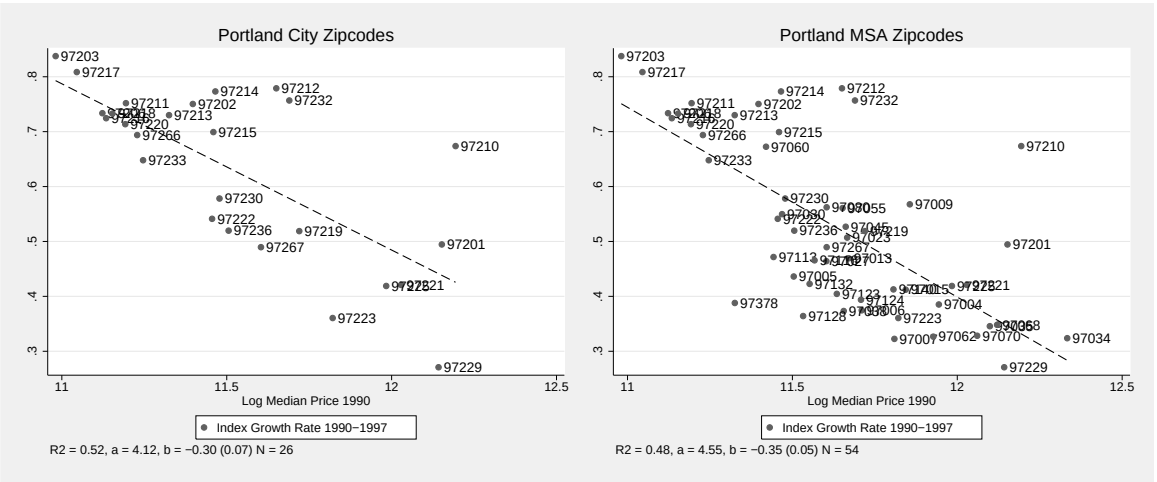


B: Beta vs. Growth 2007-2008

Figure 3: Summary Plots Showing Convergence and Divergence



A: Denver



B: Portland

Figure 4: Convergence 1990-1997

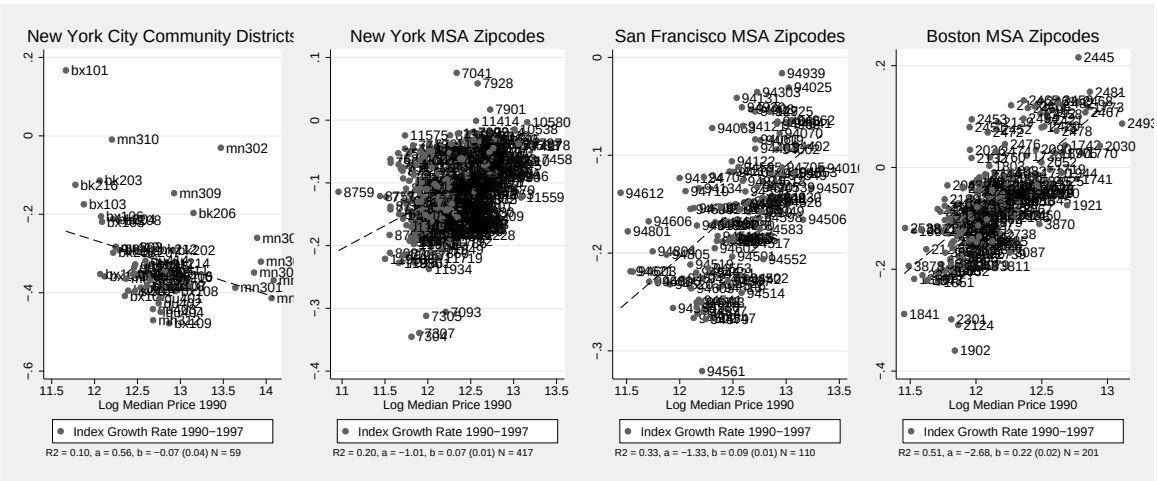


Figure 5: Divergence 1990-1997

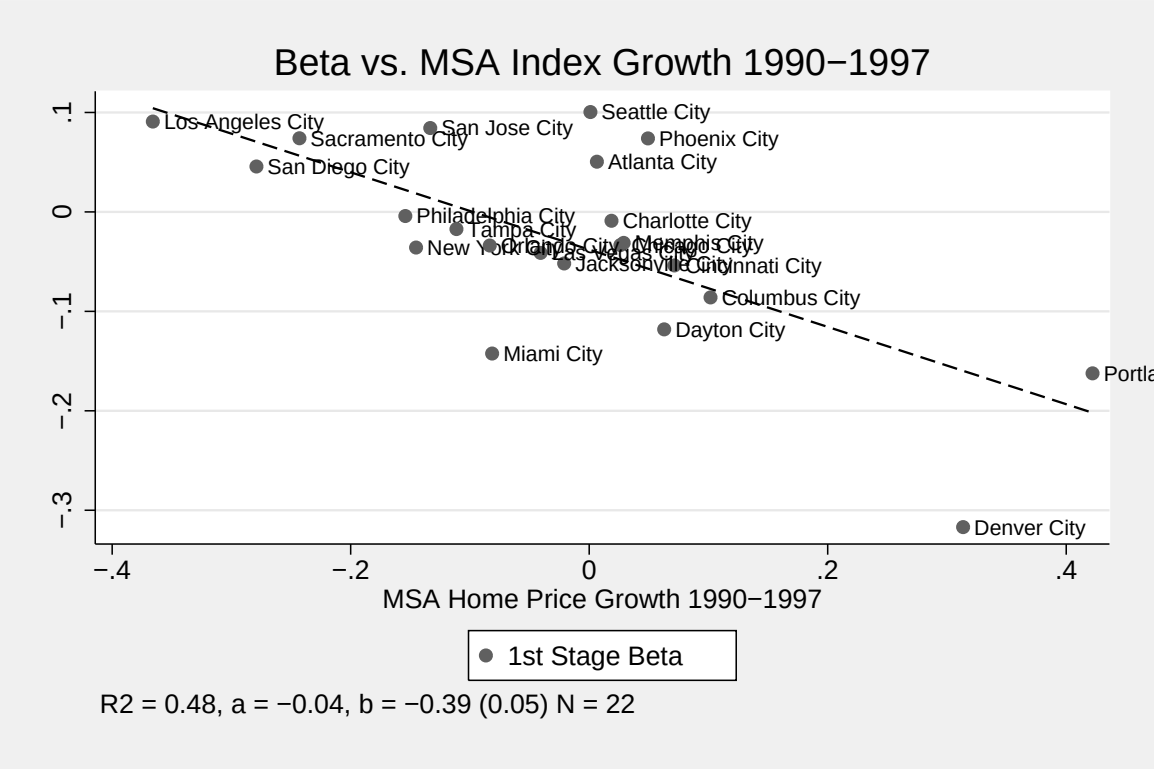


Figure 6: Summary of Convergence and Divergence 1990-1997

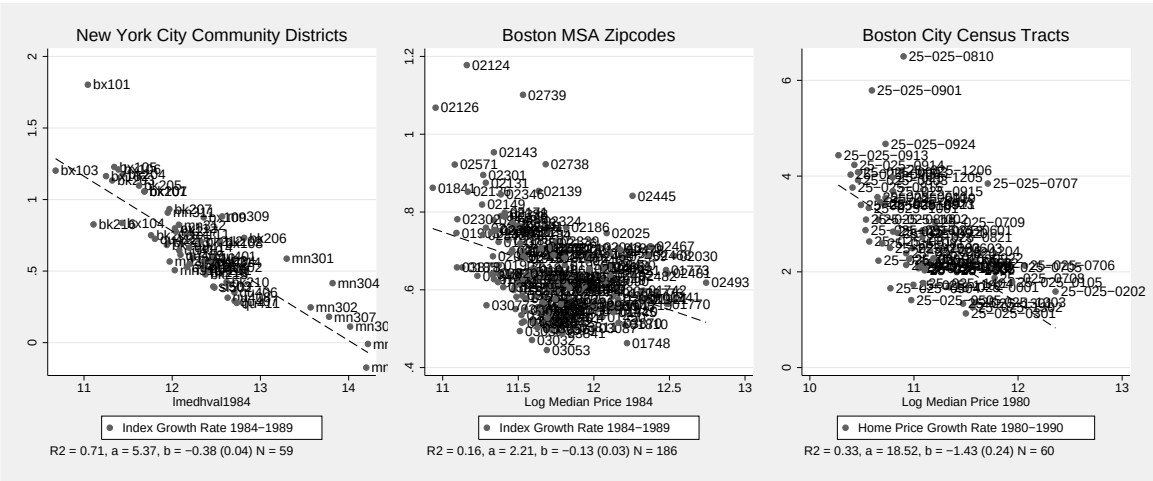


Figure 7: Convergence 1984-1989

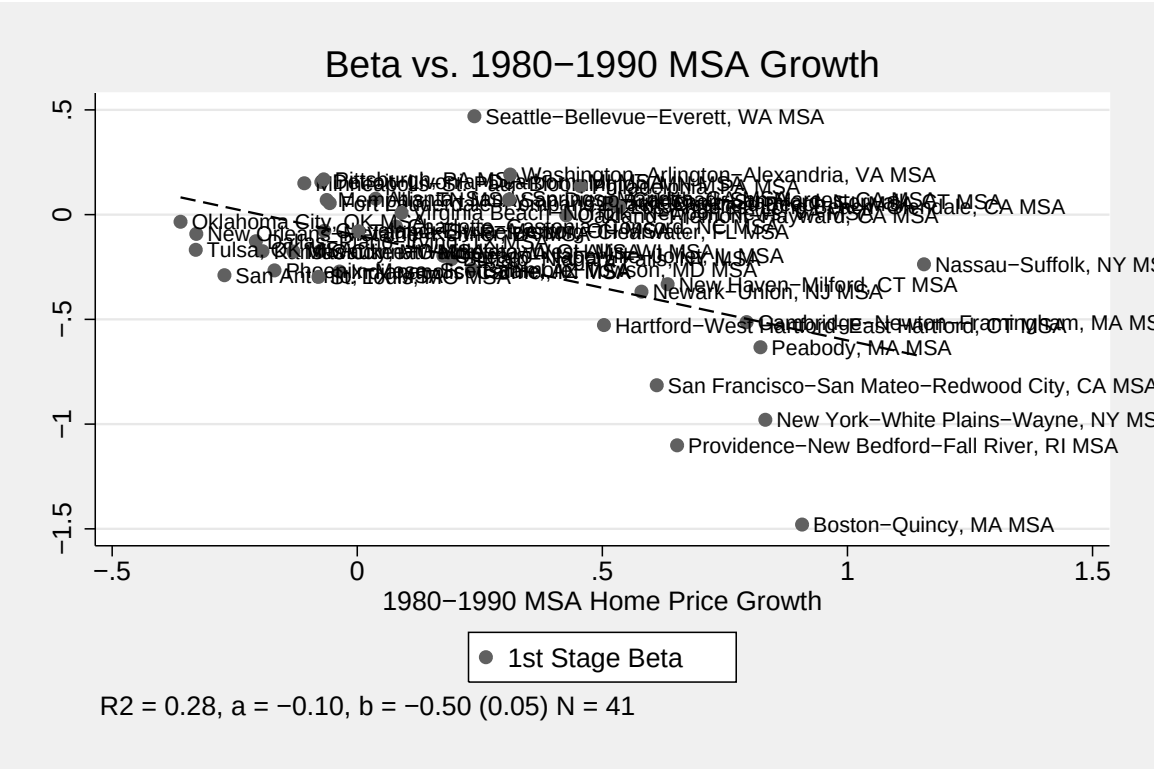


Figure 8: Convergence in the 1980's

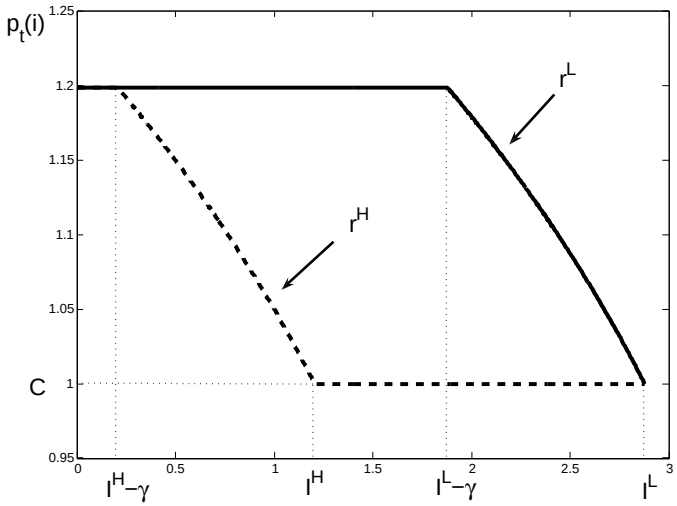


Figure 9: Reaction of the price schedule to a decline in the interest rate. We set $\alpha = .77$, $\beta = .2$, $C = 1$, $r^L = .04$, and $r^H = .05$.

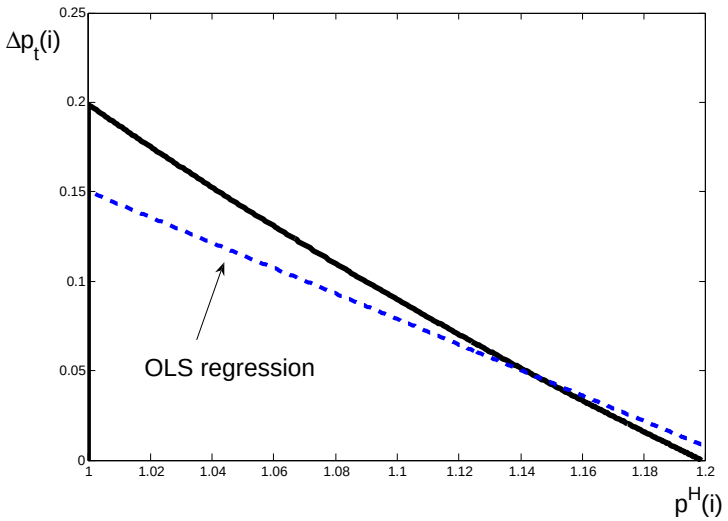


Figure 10: Price growth rate across locations as a function of the initial price level. We set $\alpha = .77$, $\beta = .2$, $C = 1$, $r^L = .04$, and $r^H = .05$.

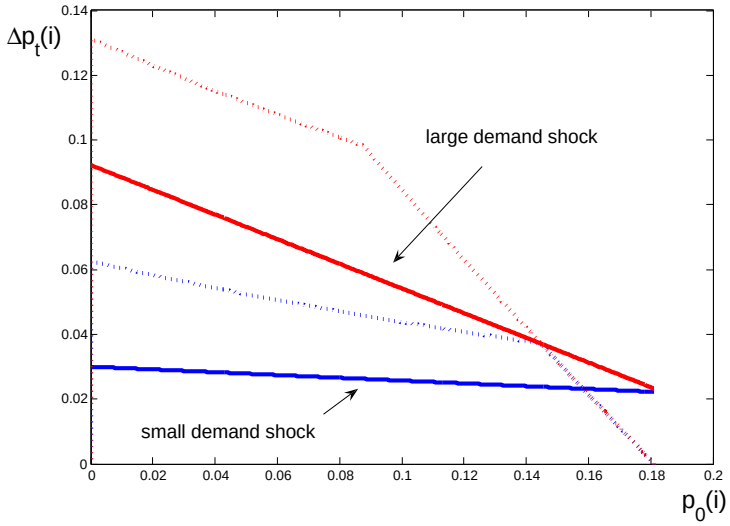


Figure 11: Price growth rate across locations as a function of the initial price level in reaction to a decline of the interest rate from .05 to .0475 (blue dots) and from .05 to .045 (red dots). The associated solid lines represent the OLS regressions. We set $\alpha = .77$, $\beta = .2$, $\phi = 1$, and $C = 1$.

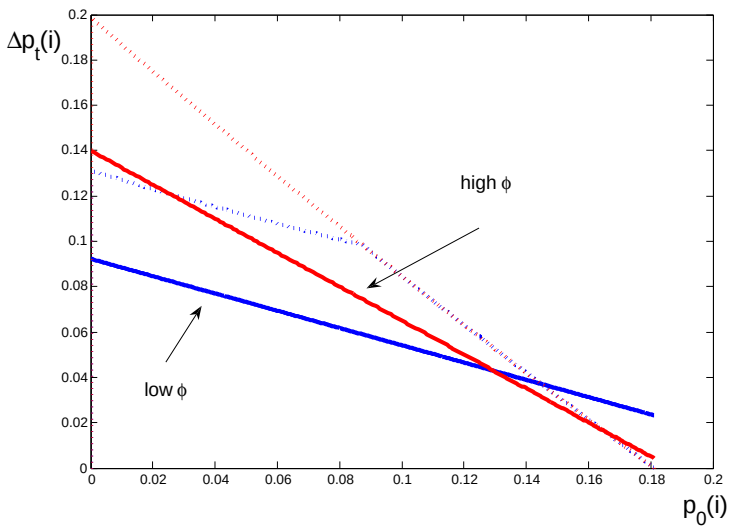


Figure 12: Price growth rate across locations as a function of the initial price level in reaction to a decline of the interest rate from .05 to .045 in a city with $\phi = 1$ (blue dots) and one with $\phi = 1.2$ (red dots). The associated solid lines represent the OLS regressions. We set $\alpha = .77$, $\beta = .2$, $C = 1$.

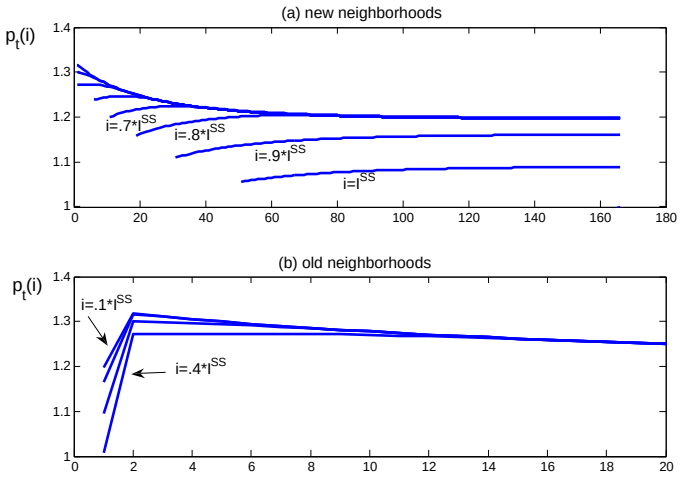


Figure 13: The figure shows the price behavior in different locations in reaction to a decline in the interest rate. Panel (a) shows neighborhoods that existed at the initial steady state, while panel (b) shows neighborhoods that are born as a result of the shock. We set $\alpha = .77$, $\beta = .2$, $C_1 = 1$, $C_2 = 1$, $\phi = 2$. Also we set $r^L = .04$ and $r^H = .05$.

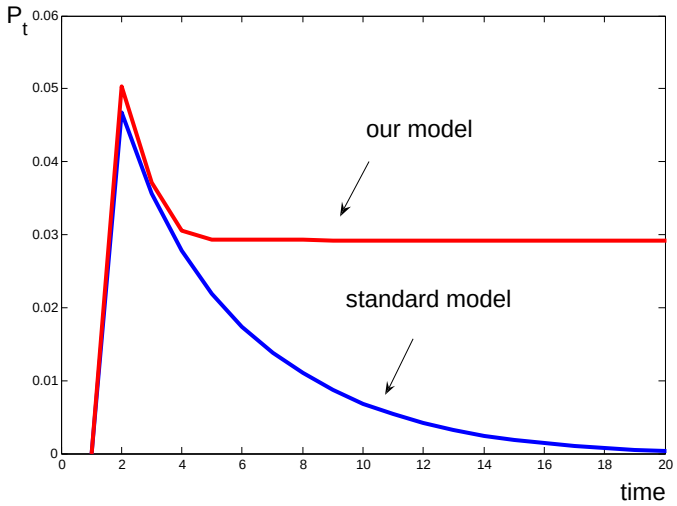


Figure 14: The figure shows the behavior of aggregate house prices after a positive demand shock in our model (red line) and in the standard adjustment cost model (blue line). We set $\alpha = .77$, $\beta = .2$, $C_1 = 1$, $C_2 = .1$, $\phi = 2.5$. Also we set $r^L = .04$ and $r^H = .05$.

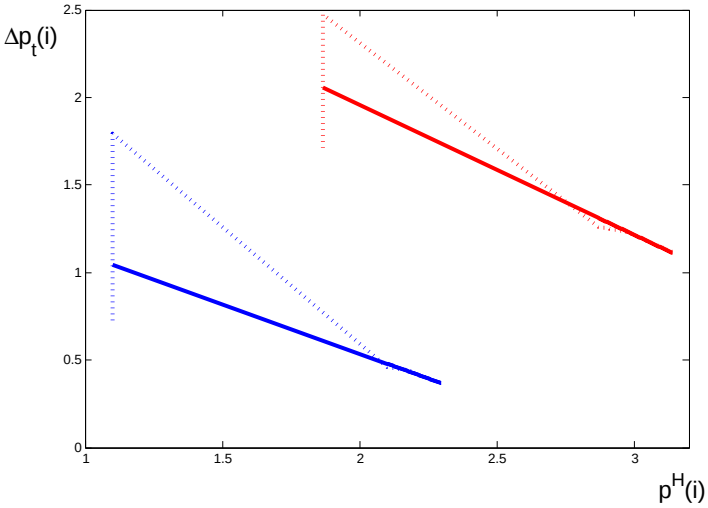
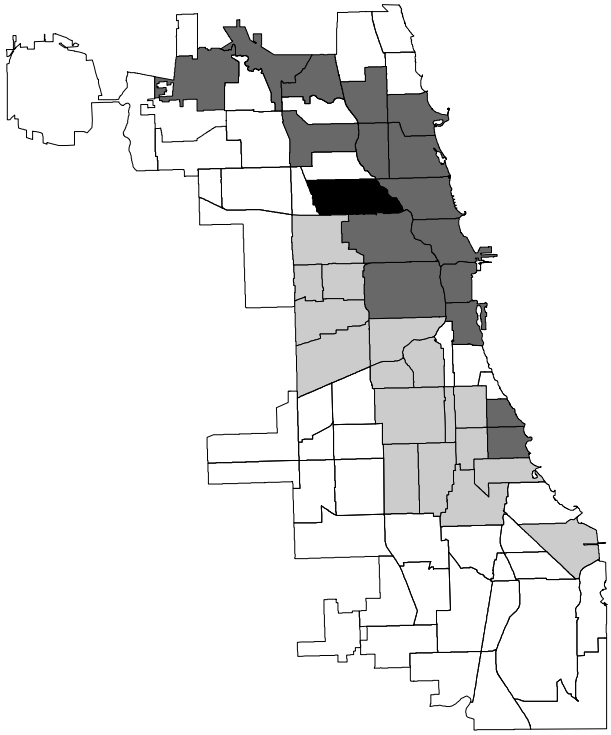
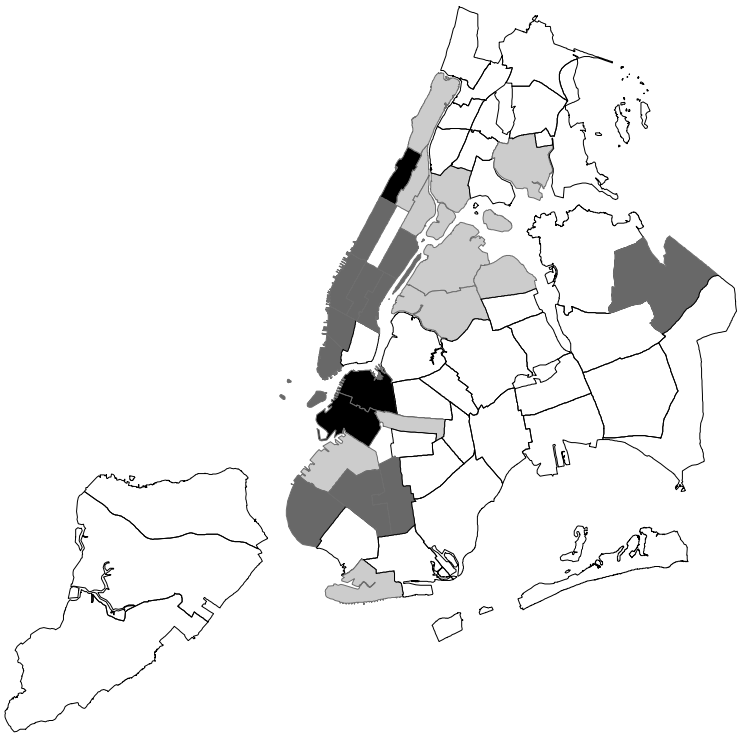


Figure 15: Consider two cities at different stages of development: city A is more developed (red lines) and city B less (blue lines). For both cities the figure shows the price growth rate, in reaction to a demand shock of the same size, as a function of the initial price level in the different locations (dotted lines) and the corresponding OLS regression (solid lines).

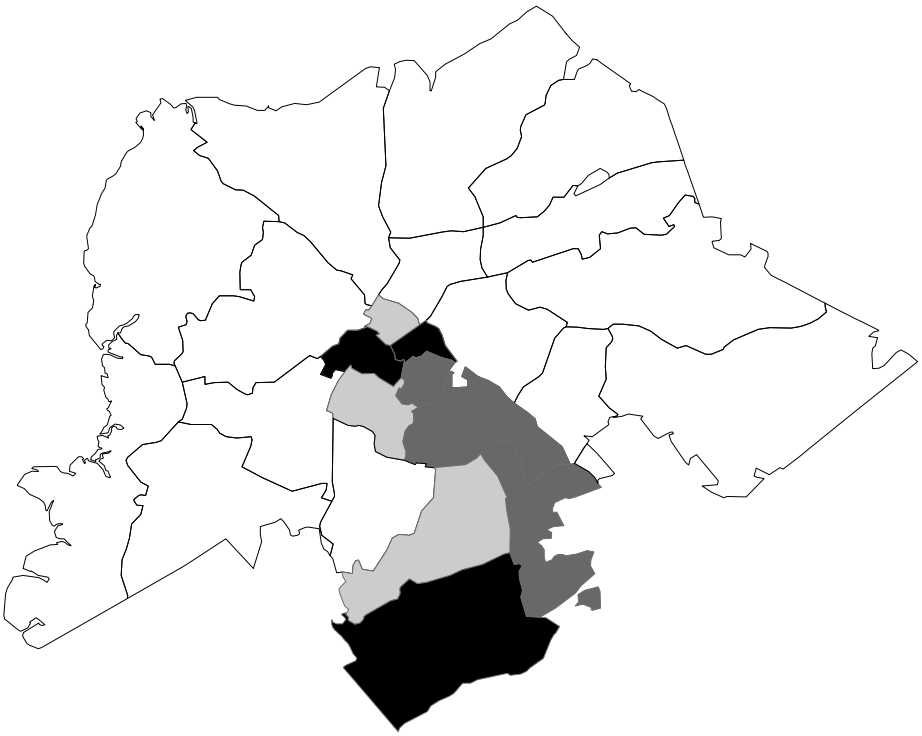


A: Chicago City



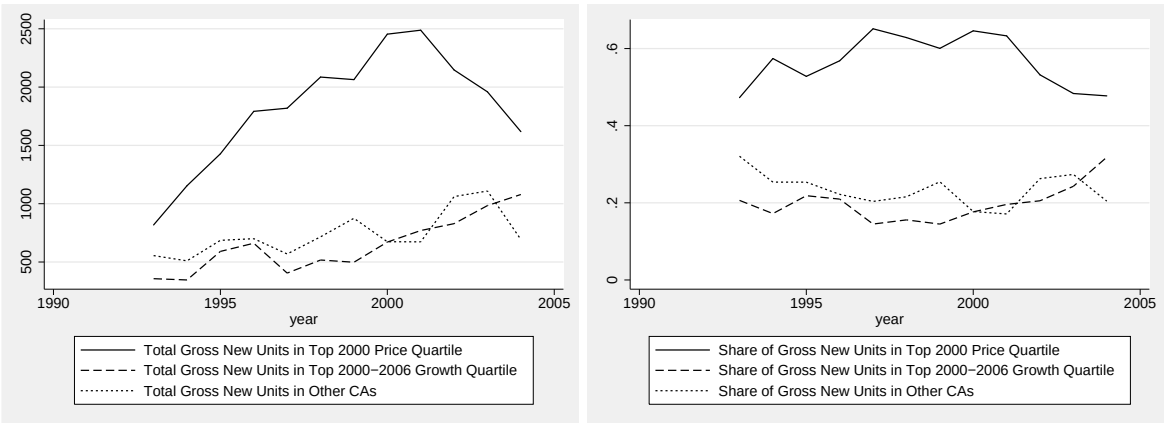
B: New York City

Figure 16: Diffusion Maps: dark shaded areas are high price in 2000, light shaded ares are high growth 2000 - 2006, black regions are both.



A: Charlotte City

Figure 17: Diffusion Maps: dark shaded areas are high price in 2000, light shaded ares are high growth 2000 - 2006, black regions are both.



A: Total Gross New Units B: Share of Gross New Units

Figure 18: Construction Patterns in Chicago Community Areas

Table 1: Correlation Between Zip Code Home Price and Income Growth

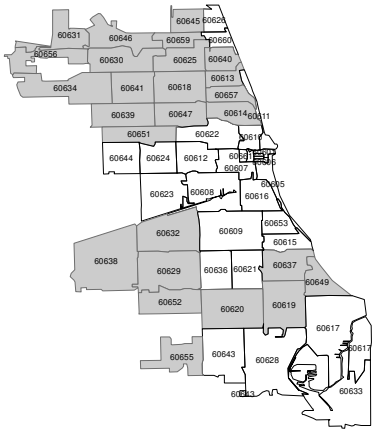
	(1) Δ Share < 10K	(2) Δ Share < 25K	(3) Δ Share > 50K	(4) Income Growth
$\overset{g}{HP}{}^{i,j}_{t,t+k}$	-0.049***	-0.074***	0.006*	-0.027*
	(0.002)	(0.004)	(0.003)	(0.015)
R^2	0.30	0.31	0.19	0.18
$\overset{g}{HP}{}^{i,j}_{t,t+k}$	-0.018***	-0.010**	-0.001	-0.022
	(0.002)	(0.004)	(0.005)	(0.022)
Init. Share/log Wage	Y	Y	Y	Y
R^2	0.61	0.52	0.19	0.18
$\overset{g}{HP}{}^{i,j}_{t,t+k}$	-0.012***	-0.011***	0.003	-0.019
	(0.002)	(0.004)	(0.005)	(0.020)
All Thresholds, Mean	Y	Y	Y	Y
R^2	0.63	0.52	0.23	0.23

Note: N = 2,738. Initial share of households below or above threshold income in 1998 or log mean wage in 1998 included as controls in middle panel specifications. All thresholds and log mean included in Bottom panel.

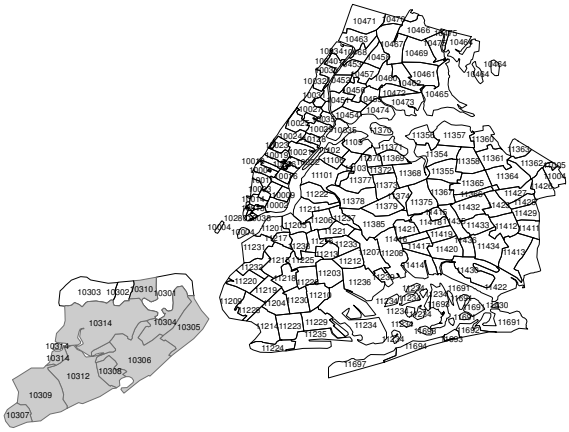
Table 2: Correlation Between Home Price and other Gentrification Measures

	(1) % Change in Med. Inc.	(2) Δ Poverty Rate	(3) Δ Frac. Afr. Amer.	(4) Δ Vac. Rate
Denver City				
1990-2000				
$\overset{g}{HP}{}^{i,j}_{t,t+k}$	0.167***	-0.047***	-0.051*	-0.026**
	(0.045)	(0.012)	(0.025)	(0.010)
N	25	25	25	25
R^2	0.52	0.68	0.31	0.35
Portland City				
1990-2000				
$\overset{g}{HP}{}^{i,j}_{t,t+k}$	0.237***	-0.051***	-0.020	-0.027***
	(0.061)	(0.015)	(0.016)	(0.009)
N	26	26	26	26
R^2	0.35	0.20	0.05	0.23
All C-S MSA Zip Codes				
1990-2000				
$\overset{g}{HP}{}^{i,j}_{t,t+k}$	0.125***	-0.034***	-0.047***	0.009*
	(0.016)	(0.004)	(0.007)	(0.005)
N	1,896	1,896	1,896	1,896
R^2	0.30	0.22	0.13	0.25
All Census MSA Tracts				
1990-2000				
$\overset{g}{HP}{}^{i,j}_{t,t+k}$	0.149***	-0.027***	-0.023***	-0.004**
	(0.019)	(0.003)	(0.003)	(0.002)
N	12,077	12,081	12,081	12,081
R^2	0.07	0.11	0.07	0.18
All Census MSA Tracts				
1980-1990				
$\overset{g}{HP}{}^{i,j}_{t,t+k}$	0.080***	-0.010***	-0.007***	-0.005**
	(0.012)	(0.002)	(0.002)	(0.002)
N	6,474	6,476	6,476	6,476
R^2	0.25	0.16	0.09	0.15

Note: Zip code or tract effects included.



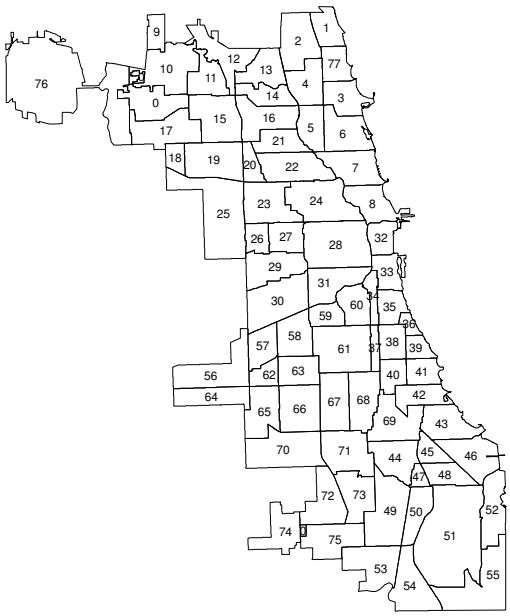
A: Chicago



B: New York City

C: Charlotte

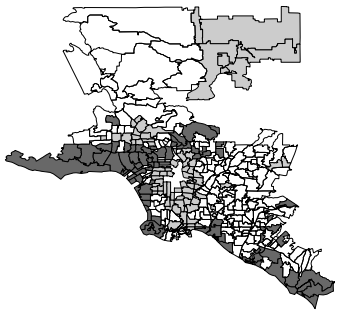
Figure A1: Shaded Zip codes are Covered by Case-Shiller Indices in 2005



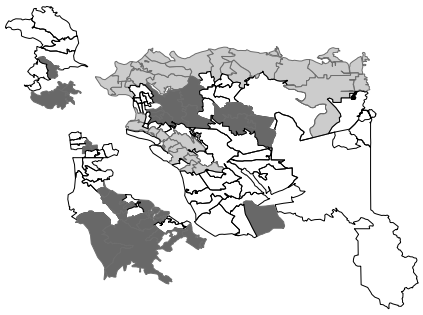
A: Chicago

B: New York City

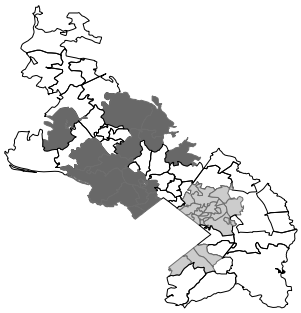
Figure A2: Chicago Community Areas and NYC Community Districts



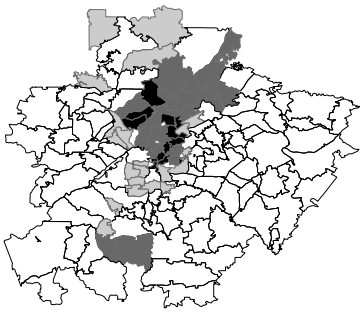
A: Los Angeles Metro 2000 - 2006



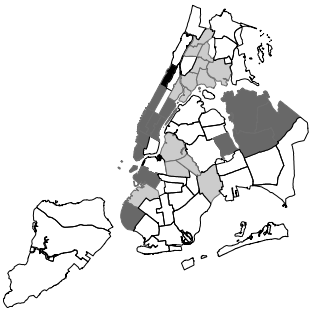
B: San Francisco Metro 2000 - 2006



C: Washington DC Metro (MD) 2000 - 2006



D: Atlanta Metro 2000 - 2006



E: New York City 1984-1989

Figure A3: Diffusion Maps: dark shaded areas are high price in initial year, light shaded ares are high growth over period listed, black regions are both.

D1 Data Appendix

D1.1 Chicago Deeds

Property records for Chicago, IL were downloaded from a section of the Chicago Tribune's website which gives access to Record Information Services' Property Transfers Database.[33] During the spring of 2008, we downloaded all records that were available on the website for property transfers pertaining to properties located in the City of Chicago through May 30, 2008. We dropped any records that had prices recorded as either zero or ten dollars. We also dropped any records which had not already been geocoded and which we were also unable to geocode. Finally, we dropped transactions in the top and bottom percentile of the price distribution in each year. (The top and bottom percentile of the price distribution for all years were transactions over $1,400,000 or under $17,600, respectively.[34]) Table D1 presents tabulations of the number of property transfers that remained by year. Next, we downloaded building characteristic data for each parcel identification number present in the Record Information Services' Property Transfers Database. We obtained this building characteristic data from the Cook County Tax Assessor's website.[35] While the Cook County Tax Assessor database contains a rich set of building characteristics for single-family and multi-family homes, only the age of the structure is available for condominiums. Table D2 shows summary statistics for the variables that are available for all property types. These include indicators for whether the property is a condominium or a multi-family building, and indicators for the age of the building. Table D3 presents estimates from the a regression of log price on these indicator variables and a vector of community area * year indicator variables. The excluded property type is single-family home and the excluded building age category is 120 years and older. To form the hedonic index, we evaluated the regression equation at the mean building characteristic values. Thus, differences in the index value are driven by differences in the estimates of the community area * year effects, and the index can be interpreted as representing the value of the mean structure type in different neighborhoods at different points in time.

D1.2 Charlotte Deeds

Property records for Charlotte, NC were downloaded from the Mecklenburg County Real Estate Lookup System.[36]. Mecklenburg County contains the entire City of Charlotte and several

[33] Web Address: `http://chicagotribune.public-record.com/realestate`.

[34] All prices are left in nominal terms until hedonic index is computed. All prices are then converted to year 2000 dollars.

[35] Web Address: `http://cookcountyassessor.com/`.

[36] Web Address: `http://meckcama.co.mecklenburg.nc.us/relookup/`

smaller townships and unincorporated areas. We began by identifying all single-family and condominium parcel identification numbers that were contained in the system in July 2009. Next, we retrieved the full record for each parcel ID. Each record contains information about the owner (name, whether they are an individual, partnership, or corporation, mailing address), the legal description of the property (municipality, fire district, etc.), tax assessment value, tax exemptions, land use, building characteristics, sales history, and assessed value history. While the owner information contained the zip code of the owner's mailing address, the building characteristic information contained all of the address of building except for the zip code. If the building address matched the owner address, we assigned the building the same zip code as was listed in the owner's mailing address. If not, we attempted to geocode the address. We dropped any transactions related to properties for which we were unable to determine a zip code. We dropped any transactions which had a price equal to zero. We dropped transactions in the top and bottom percentile of the price distribution in each year. Table D4 presents tabulations of the number of property transfers that remained by year. The sharp increase in transactions from 1992 to 1994 leads us to believe that some part of the universe of transactions is missing prior to 1994, thus we drop any transactions prior to 1994. We also drop any transactions that occurred in 2009. Table D5 shows summary statistics for the property characteristic variables that are available for Mecklenburg County. These include the log of square footage, the log of heated square footage, the number of bathrooms, the number of bedrooms, whether the property is a condominium or single-family home, whether the property has a crawl space or a slab foundation, whether there are multiple structures on the property, whether there are multiple units in the property, whether the property has air duct heating, a heat pump or another type of heating, whether the property has central air conditioning, the number of stories, age of the building, type of exterior, the quality grade of the structure, and the type of fuel used for heating. Table D6 presents estimates from the a regression of log price on these variables and a vector of community area * year indicator variables. To form the hedonic index, we evaluated the regression equation at mean building characteristics values.

D1.3 Furman Center Repeat Sales Index for NYC

We use a repeat sales index for New York City community districts that was produced by NYU's Furman Center for Real Estate and Urban Policy. These data consist of a repeat sales index for each community district in New York City reported at an annual frequency and running from 1974 through 2008.[37]

[37] See `http://www.furmancenter.org` for more information.

D1.4 Census Data

We use tabulations from the United States Census from 1980, 1990, and 2000. Some data come from Summary File 1 (SF1) which contains 100 percent counts and information for all people and housing units, while other data come from Summary File 3 (SF3) which contains a 1-in-6 sample and is weighted to represent the entire population. In general, we obtained data from the 2000 Census from the American Factfinder Website.[38] However, zip code and census tract tabulations for the 1980 and 1990 Censuses were obtained from the Inter-University Consortium for Political and Social Research.[39] Finally, variables indicating census tracts that had not changed boundaries between 1980 and 1990 were obtained from the Neighborhood Change Database produced by Geolytics.[40] The data used for the x-axis of Figure 1 come Census 2000 SF3 tract level tabulations. For the x-axis of the left-most panel of Figure 1A we calculate mean owner-occupied home value by summing the aggregate owner-occupied home value for all census tracts in a community area and dividing by the sum of the number of owner-occupied housing units in the community area. For the x-axis of the left-most panel of Figure 1B we obtained 2000 Census tabulations of the median value of owner-occupied housing units which were specially tabulated for NYC community districts and are available from NYC's GIS department at `http://gis.nyc.gov`. The x-axis of all other plots in Figure 1 and all plots of Figure 2 are simply from the median owner-occupied housing value variable from the 2000 Census SF3 tabulation for 5-digit zip codes. The x-axes of Figures 4 and 5 also use the median owner-occupied housing value variable from the 2000 Census SF3 tabulation for 5-digit zip codes, but 1990 values have been imputed by deflating by the growth of the Case-Shiller zip code index from 1990 to 2000. A similar imputation has been performed for the left and center panels of Figure 7. These data also underly the first stage estimates of Figures 3 and 6.

The right-most panel of Figure 7 and the data underlying Figure 8 use census tract tabulations of the 1980 and 1990 censuses. For both years we used a variable from the SF3 file. However, the definition of the variable changed from 1980 to 1990. In 1980, the variable is the median value of non-condominium owner-occupied housing units, while the 1990 variable is the median value of all owner-occupied housing units (including condominiums). We restricted our analysis to census tracts for which the boundaries had not changed from 1980 to 1990.

[38] Web Address: `http://factfinder.census.gov/`.

[39] Web Address: `http://www.icpsr.umich.edu/icpsrweb/ICPSR/`.

[40] Web Address: `http://www.geolytics.com/USCensus,Neighborhood-Change-Database-1970-2000,Products.asp`.

D1.5 MSA-Level Housing Price Growth

Data on MSA-level housing price growth come from two publicly available repeat sales indices. Repeat sales indices for the 20 cities for which Chicago Mercantile Exchange Housing Futures and Options are traded come from Case-Shiller.[41]. For other cities we used the Federal Housing Finance Agency MSA repeat sales indices (formerly called the OFHEO indices).[42]

[41] Web Address: `http://www.macromarkets.com/csi_housing/sp_caseshiller.asp`.
[42] Web Address: `http://www.fhfa.gov/Default.aspx?Page=216`.

Table D1: Tabulation of Chicago Property Deed Data by Year

	(1) Number of Observations
2000	33,548
2001	33,894
2002	37,250
2003	51,436
2004	55,819
2005	51,274
2006	50,043
2007	42,024
2008	6,857
Total	362,145

Note: The tabulation for 2008 includes only transactions that occurred before May 30, 2008.

Table D2: Summary Statistics for Chicago Hedonic Regression Variables

	(1) Mean	(2) S.D.	(3) Min.	(4) Max.
Price	$271,068	$184,121	$10,184	$2,000,000
Condominium	0.26	0.44	0	1
Multi-Family	0.17	0.37	0	1
Age 0-4 Years	0.07	0.26	0	1
Age 5-9 Years	0.04	0.19	0	1
Age 10-14 Years	0.01	0.11	0	1
Age 15-19 Years	0.01	0.12	0	1
Age 20-29 Years	0.09	0.28	0	1
Age 30-39 Years	0.17	0.37	0	1
Age 40-49 Years	0.10	0.30	0	1
Age 50-59 Years	0.08	0.26	0	1
Age 60-69 Years	0.03	0.16	0	1
Age 70-79 Years	0.07	0.25	0	1
Age 80-89 Years	0.10	0.31	0	1
Age 90-99 Years	0.09	0.29	0	1
Age 100-109 Years	0.07	0.25	0	1
Age 110-119 Years	0.06	0.24	0	1

Note: 362,145 observations.

Table D3: Chicago Hedonic Regression

	(1) Coeff.	(2) S.E.
Condominium	-0.376***	(0.003)
Multi-Family	0.271***	(0.003)
Age 0-4 Years	0.229***	(0.008)
Age 5-9 Years	0.388***	(0.008)
Age 10-14 Years	0.303***	(0.010)
Age 15-19 Years	0.141***	(0.011)
Age 20-29 Years	0.027***	(0.008)
Age 30-39 Years	-0.028***	(0.008)
Age 40-49 Years	0.021***	(0.008)
Age 50-59 Years	0.046***	(0.008)
Age 60-69 Years	0.098***	(0.008)
Age 70-79 Years	0.111***	(0.008)
Age 80-89 Years	0.093***	(0.008)
Age 90-99 Years	0.093***	(0.008)
Age 100-109 Years	0.056***	(0.008)
Age 110-119 Years	0.014*	(0.008)

Note: Dependent variable is log price. N = 362,145 observations. R^2 = 0.46. Eicker-Huber-White standard errors in parentheses. Community area * year effects included.

Table D4: Tabulation of Mecklenburg County Property Deed Data by Year

	(1) Number of Observations
1980	819
1981	589
1982	601
1983	1,208
1984	1,369
1985	1,543
1986	2,298
1987	2,273
1988	2,399
1989	2,359
1990	1,224
1991	1,229
1992	1,228
1993	4,594
1994	11,260
1995	10,859
1996	12,371
1997	12,413
1998	14,608
1999	16,622
2000	16,277
2001	14,229
2002	12,087
2003	12,381
2004	15,643
2005	16,867
2006	18,617
2007	16,828
2008	7,894
2009	1,346
Total	235,035

Note: The tabulation for 2009 includes only transactions that occurred before August, 2009.

Table D5: Summary Statistics for Mecklenburg County Hedonic Regression Variables

	(1) Mean	(2) S.D.	(3) Min.	(4) Max.
Price	$190,454	$127,522	$19,500	$1,275,000
log sqft.	7.73	0.44	6.20	10.57
log heated sqft.	7.55	0.40	6.17	9.60
0 Baths	0.00	0.02	0	1
2 Baths	0.74	0.44	0	1
3 Baths	0.12	0.32	0	1
4+ Baths	0.03	0.16	0	1
0 Bed Rooms	0.01	0.10	0	1
2 Bed Rooms	0.10	0.31	0	1
3 Bed Rooms	0.55	0.50	0	1
4 Bed Rooms	0.29	0.45	0	1
5+ Bed Rooms	0.04	0.19	0	1
Condo	0.13	0.34	0	1
Crawlspace Foundation	0.54	0.50	0	1
Multiple Structures	0.00	0.05	0	1
Multiple Units	0.01	0.08	0	1
Air Duct Heat	0.91	0.28	0	1
Heat Pump	0.06	0.25	0	1
Central Air	0.95	0.22	0	1
1 Story	0.35	0.48	0	1
1.5 Stories	0.09	0.29	0	1
2 Stories	0.53	0.50	0	1
2.5+ Stories	0.02	0.15	0	1
Age 0-4 Years	0.41	0.49	0	1
Age 5-9 Years	0.15	0.35	0	1
Age 10-14 Years	0.10	0.30	0	1
Age 15-19 Years	0.07	0.25	0	1
Age 20-29 Years	0.08	0.28	0	1
Age 30-39 Years	0.06	0.25	0	1
Age 40-49 Years	0.06	0.24	0	1
Age 50-69 Years	0.06	0.23	0	1
Age 70-89 Years	0.01	0.12	0	1
Alum./Vinyl Exterior	0.27	0.45	0	1
Alum./Vinyl and Brick	0.11	0.32	0	1
Alum./Vinyl and Stone	0.00	0.07	0	1
Asbestos Siding	0.01	0.07	0	1
Cedar/Redwood Siding	0.01	0.09	0	1
Plywood Siding	0.02	0.13	0	1
Brick	0.26	0.44	0	1
HardiPlank Siding	0.01	0.09	0	1
Masonite	0.15	0.35	0	1
Masonite and Brick	0.05	0.21	0	1
Stucco	0.01	0.12	0	1
Wood on Sheeting	0.05	0.22	0	1
Wood Sheeting and Brick	0.03	0.16	0	1
Grade: Below Average	0.00	0.06	0	1
Grade: Custom	0.00	0.07	0	1
Grade: Good	0.21	0.41	0	1
Grade: Very Good	0.06	0.24	0	1
Grade: Excellent	0.01	0.11	0	1
Electric Heat	0.14	0.34	0	1
Gas Heat	0.85	0.35	0	1
Oil Heat	0.01	0.07	0	1
Solar Heat	0.00	0.01	0	1

Note: 208,956 observations.

Table D6: Mecklenburg County Hedonic Regression

	(1) Coeff.	(2) S.E.
log sqft.	0.320***	(0.006)
log heated sqft.	0.354***	(0.007)
0 Baths	0.071	(0.050)
2 Baths	0.084***	(0.002)
3 Baths	0.109***	(0.004)
4+ Baths	0.099***	(0.007)
0 Bed Rooms	-0.051***	(0.013)
2 Bed Rooms	-0.004	(0.009)
3 Bed Rooms	-0.002	(0.009)
4 Bed Rooms	0.014	(0.009)
5+ Bed Rooms	0.000	(0.010)
Condo	-0.077***	(0.002)
Crawlspace Foundation	0.034***	(0.001)
Multiple Structures	0.189***	(0.021)
Multiple Units	-0.231***	(0.013)
Air Duct Heat	0.015***	(0.006)
Heat Pump	0.024***	(0.006)
Central Air	0.118***	(0.004)
1 Story	0.046***	(0.009)
1.5 Stories	0.045***	(0.009)
2 Stories	0.023***	(0.009)
2.5+ Stories	-0.002	(0.010)
Age 0-4 Years	0.090***	(0.018)
Age 5-9 Years	0.080***	(0.018)
Age 10-14 Years	0.043**	(0.018)
Age 15-19 Years	0.005	(0.018)
Age 20-29 Years	-0.050***	(0.018)
Age 30-39 Years	-0.087***	(0.018)
Age 40-49 Years	-0.082***	(0.018)
Age 50-69 Years	0.027	(0.018)
Age 70-89 Years	0.064***	(0.019)
Alum./Vinyl Exterior	-0.054***	(0.005)
Alum./Vinyl and Brick	-0.035***	(0.005)
Alum./Vinyl and Stone	-0.049***	(0.009)
Asbestos Siding	-0.090***	(0.011)
Cedar/Redwood Siding	0.086***	(0.009)
Plywood Siding	-0.112***	(0.006)
Brick	0.033***	(0.005)
HardiPlank Siding	-0.042***	(0.011)
Masonite	-0.023***	(0.005)
Masonite and Brick	-0.009*	(0.005)
Stucco	0.050***	(0.010)
Wood on Sheeting	-0.044***	(0.005)
Wood Sheeting and Brick	-0.045***	(0.005)
Grade: Below Average	-0.182***	(0.015)
Grade: Custom	0.403***	(0.018)
Grade: Good	0.174***	(0.002)
Grade: Very Good	0.341***	(0.004)
Grade: Excellent	0.443***	(0.010)
Electric Heat	0.049**	(0.021)
Gas Heat	0.055***	(0.021)
Oil Heat	0.081***	(0.023)
Solar Heat	0.167**	(0.079)

Note: Dependent variable is log price. N = 208,956 observations. $R^2 = 0.82$. Eicker-Huber-White standard errors in parentheses. Zip code * year effects included.