

w o r k i n g
p a p e r

10 08R

**Endogenous Gentrification and
Housing-Price Dynamics**

Veronica Guerrieri, Daniel Hartley, and
Erik Hurst



FEDERAL RESERVE BANK OF CLEVELAND

Working papers of the Federal Reserve Bank of Cleveland are preliminary materials circulated to stimulate discussion and critical comment on research in progress. They may not have been subject to the formal editorial review accorded official Federal Reserve Bank of Cleveland publications. The views stated herein are those of the authors and are not necessarily those of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System.

Working papers are available on the Cleveland Fed's website at:

www.clevelandfed.org/research.

Endogenous Gentrification and Housing-Price Dynamics

Veronica Guerrieri, Daniel Hartley, and Erik Hurst

In this paper, we begin by documenting substantial variation in house-price growth across neighborhoods within a city during citywide housing price booms. We then present a model which links house-price movements across neighborhoods within a city and the gentrification of those neighborhoods in response to a citywide housing-demand shock. A key ingredient in our model is a positive neighborhood externality: individuals like to live next to richer neighbors. This generates an equilibrium where households segregate based upon their income. In response to a citywide demand shock, higher-income residents will choose to expand their housing by migrating into the poorer neighborhoods that directly abut the initial richer neighborhoods. The in-migration of the richer residents into these border neighborhoods will bid up prices in those neighborhoods, causing the original poorer residents to migrate out. We refer to this process as “endogenous gentrification.” Using a variety of data sets and using Bartik variation across cities to identify city-level housing demand shocks, we find strong empirical support for the model’s predictions.

Keywords: gentrification, housing-price dynamics, housing-consumption externalities.

JEL codes: R12, R21, I32.

*Original version July 2010. First revision February 2012. Second revision March 2012.

Note: The online robustness appendix that accompanies this paper is available at <http://www.clevelandfed.org/research/workpaper/2010/wp1008R2-appendix.pdf>.

Veronica Guerrieri is at the University of Chicago Booth School of Business (Veronica.Guerrieri@Chicago-Booth.edu). Daniel Hartley is at the Federal Reserve Bank of Cleveland (Daniel.Hartley@clev.frb.org). Erik Hurst is at the University of Chicago Booth School of Business (Erik.Hurst@ChicagoBooth.edu). The authors would like to thank seminar participants at Chicago, Cleveland State, Duke Conference on Housing Market Dynamics, Harvard, MIT, Oberlin, Ohio State, Queen’s University Conference on Housing and Real Estate Dynamics, Rochester, Stanford, Summer 2010 NBER PERE meeting, Tufts, UCLA, UIC, University of Akron, USC, Wharton, Winter 2010 NBER EFG program meeting, Wisconsin, and the Federal Reserve Banks of Atlanta, Boston, Chicago, Cleveland, Minneapolis, and St. Louis. They are particularly indebted to Daron Acemoglu, Gary Becker, Hoyt Bleakley, V. V. Chari, Raj Chetty, Morris Davis, Fernando Ferreira, Ed Glaeser, Matt Kahn, Larry Katz, Jed Kolko, Guido Lorenzoni, Erzo Luttmer, Enrico Moretti, Kevin Murphy, Matt Notowidigdo, John Quigley, Esteban Rossi-Hansberg, Jesse Shapiro, and Todd Sinai for their detailed comments on previous drafts of this paper. Guerrieri and Hurst would like to acknowledge financial support from the University of Chicago’s Booth School of Business.

1 Introduction

It has been well documented that there are large differences in house price appreciation rates across U.S. metropolitan areas.¹ For example, according to the Case-Shiller Price Index, real property prices increased by over 100 percent in Washington DC, Miami, and Los Angeles between 2000 and 2006, while property prices appreciated by roughly 10 percent in Atlanta and Denver during the same time period. Across the 20 MSAs for which a Case-Shiller MSA index is publicly available, the standard deviation in real house price growth between 2000 and 2006 was 42 percent. Such variation is not a recent phenomenon. During the 1990s, the Case-Shiller cross-MSA standard deviation in house price growth was 21 percent.

While most of the literature has focused on trying to explain cross-city differences in house price appreciation, we document that there are also substantial *within*-city differences in house price appreciation. For example, between 2000 and 2006 residential properties in the Harlem neighborhood of New York City appreciated by over 130 percent, while residential properties less than two miles away, in midtown Manhattan, only appreciated by 45 percent. The New York City MSA, as a whole, appreciated by roughly 80 percent during this time period. Such patterns are common in many cities. Using within-city price indices from a variety of sources, we show that the average *within*-MSA standard deviation in house price growth during the 2000 - 2006 period was roughly 20 percent. Similar patterns are also found during the 1990s and 1980s. As is commonly discussed in the popular press, these large relative movements in property prices within a city during city-wide property price booms are often associated with changing neighborhood composition. Returning to the Harlem example, a recent New York Times article discussed how Harlem residents have gotten richer during the period when its house prices were substantially appreciating.²

Our goals in this paper are threefold. First, we set out to document a new set of facts about the extent and nature of within-city house price movements during city-wide housing price booms. The house price appreciation for the city as a whole is just a composite of the house price movements within all the neighborhoods of the city. Therefore, understanding the movements in house prices across neighborhoods within a city is essential for understanding house price movements for the entire city. Using a variety of different data sources, we show that there are substantial differences across neighborhoods within a city with respect to their house price growth when the city as a whole experiences a housing price boom.

Moreover, we show that there is a systematic pattern in this variation. In particular, we document three facts that are robust across time and data sources with respect to within-city house price movements. First, during city-wide housing price booms, neighborhoods with low initial housing prices appreciate at much greater rates than neighborhoods with high initial prices. Second, the variation in housing price appreciation rates among low housing price neighborhoods

¹See, for example, Davis et al. (2007), Glaeser et al. (2008), Van Nieuwerburgh and Weill (2010), and Saiz (2010).

²See the article “No Longer Majority Black, Harlem Is In Transition” from the January 5th, 2010 New York Times.

is much higher than the variation in housing price appreciation rates for higher housing price neighborhoods. Finally, we show that the larger the city-wide housing price boom, the greater is the difference in housing price appreciation rates between low house price and high house price neighborhoods. Regardless of the interpretation we give to some of these facts in later sections, we feel these facts alone are an interesting contribution to the literature on spatial variation in housing price growth.

Our second goal is to develop a spatial model of a city that links within-city neighborhood housing price dynamics with gentrification. We represent a city as the real line and each point on the line is a location. Agents are fully mobile across locations and there is a representative firm that can build houses in any location at a fixed marginal cost. The key ingredient of the model is that agents are heterogeneous in their income and all agents prefer to live close to richer neighbors. The relevance of such a neighborhood consumption externality in determining house prices is supported by the recent empirical work of Bayer et al. (2007) and Rossi-Hansberg et al. (2010). We show that there exists an equilibrium with full income segregation where the high income residents are concentrated all together and the low income residents live at the periphery. The sorting, as in Becker and Murphy (2003), is the result of the neighborhood externality where all agents are willing to pay more to live closer to rich neighbors. Poorer residents are less willing to pay high rents to live in the rich neighborhoods, so in equilibrium they live farther from the rich. Within the model, house prices achieve their maximum in the rich neighborhoods and decline as one moves away from them, to compensate for the lower level of the externality. For the neighborhoods that are far enough from the rich, there is no externality, and house prices are equal to the marginal cost of construction.

One of the main contributions of our model, and the basis for our subsequent empirical work, is to explore the dynamics of house prices across neighborhoods in response to city-wide housing demand shocks. Although there is no aggregate supply constraint and the city can freely expand, average house prices increase in response to an increase in city-wide housing demand because of gentrification. In particular, the neighborhoods that endogenously gentrify are the poor neighborhoods on the border of rich neighborhoods. For concreteness, we say that a neighborhood gentrifies when some poor residents are replaced by richer ones, increasing the extent of the neighborhood externality. For example, we consider a city hit by an increase in labor demand and a subsequent wave of migration (Blanchard and Katz, 1992). The richer migrants prefer to locate next to the existing richer households. As a result, they bid up the land prices in the poor neighborhoods that are next to the rich neighborhoods causing the existing poor residents to move out and the city as a whole to expand.

To sum up, our mechanism implies that unexpected permanent shocks to housing demand lead to permanent increases in house prices at the city level although the size of the city is completely elastic. This happens because gentrification bids up the value of the land in the gentrifying neighborhoods. Moreover, our model predicts that, in response to a positive city-wide housing demand shock, land prices in poor neighborhoods that are in close proximity to the rich neighborhoods

appreciate at a faster rate than both richer neighborhoods and other poor neighborhoods. We also find that average price growth within the city is affected both by the size of the housing demand shock and by the particular shape of preferences, technology, and income distribution within the city.

Our third goal is to provide explicit evidence showing that our endogenous gentrification mechanism is an important determinant of within-city variation in house price growth in response to city-wide housing demand shocks. We do this in multiple ways. To begin, we provide an additional fact about within-city neighborhood house price appreciation during city-wide housing booms. In particular, we show that, as our theory predicts, among all the poor neighborhoods it is the poor neighborhoods that are next to the rich neighborhoods that appreciate the most during city-wide housing booms. This result holds in the 1980s, 1990s, and 2000s and holds using a variety of different measures of neighborhood housing price appreciation. Moreover, these results are robust to including controls for distance to the city's center business district, the average commuting time of neighborhood residents, and proximity of the neighborhood to fixed natural amenities such lakes, oceans, and rivers. Again, these results are consistent with the first order predictions of our model.

We then use a Bartik-style instrument to isolate exogenous city level housing demand shocks (Bartik, 1991) and show that it is the housing prices in poor neighborhoods next to rich neighborhoods that appreciate the most in response to the exogenous city-wide housing demand shocks. Our Bartik shock predicts expected income growth in a city between periods t and $t + k$ based on the initial industry mix in that city at time t and the change in industry earnings for the entire U.S. between t and $t + k$. For example, in response to a one standard deviation Bartik shock, poor neighborhoods within the city which directly border a rich neighborhood have housing prices that appreciate roughly 7.0 percentage points (compared to a mean appreciation rate of 24.0 percent) more than otherwise similar poor neighborhoods within the city that are more than 3 miles away from rich neighborhoods. Again, these results hold controlling for distance to the center business district and proximity to fixed natural amenities within the city.

Finally, we explicitly show that the neighborhoods that appreciate the most during the exogenous city-wide housing demand shock also gentrify. Gentrification - the out migration of poor residents and the in migration of rich residents - is the key mechanism for the within-city house price dynamics we highlight. For this analysis, we again explore the within-city response to a Bartik-style shock. In particular, we show that in response to an exogenous city-wide demand shock, poor neighborhoods close to rich neighborhoods experience larger increases in neighborhood income, larger increases in the educational attainment of neighborhood residents, and larger declines in the neighborhood poverty rate than do otherwise similar poor neighborhoods that are farther away from the rich neighborhoods. For example, average neighborhood income grows by roughly 1.7 percentage points (compared to a mean growth rate of 14.9 percent) more in response to a one standard deviation Bartik shock for poor neighborhoods that border the rich neighborhoods than it does for otherwise similar poor neighborhoods that are more than 3 miles away from the rich neighborhoods. Lastly, we highlight that during both the 1980s and 1990s, most of the

poor neighborhoods that did in fact gentrify by some ex-post criteria were neighborhoods that were directly bordering existing rich neighborhoods.

As noted above, a key ingredient in our model is the existence of neighborhood consumption externalities in that individuals get utility from having rich neighbors relative to poor neighbors. Although, we do not explicitly model the direct mechanism for the externality, we have many potential channels in mind. For example, crime rates are lower in richer neighborhoods. If households value low crime, individuals will prefer to live in wealthier neighborhoods. Likewise, the quality and extent of public goods may be correlated with the income of neighborhood residents. For example, school quality - via peer effects, parental monitoring, or direct expenditures - tends to increase with neighborhood income. Finally, if there are increasing returns to scale in the production of desired neighborhood amenities (number and variety of restaurants, easier access to service industries such as dry cleaners, movie theaters, etc.), such amenities will be more common as the income of one's neighbors increases. Although we do not take a stand on which mechanism is driving the externality, our preference structure is general enough to allow for any story that results in higher amenities being endogenously provided in higher income neighborhoods.

Our work adds to the large literature on neighborhood gentrification.³ Some of this literature highlights correlates with neighborhood gentrification. For example, both Kolko (2007) and Brueckner and Rosenthal (2008) emphasize that the age and quality of the housing stock within a poor neighborhood is an important predictor of whether or not that poor neighborhood ever gentrifies. Additionally, there is a separate strand of work that emphasizes the importance of spatial dependence - either theoretically or empirically - in predicting neighborhood gentrification. For example, 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). Kolko (2007) finds that poor neighborhoods bordering richer neighborhoods in 1990 had larger income growth between 1990 and 2000 than otherwise similar poor neighborhoods that were next to other poor neighborhoods. Our addition to this literature is that we propose a model that explains both of these facts and then formally test the model's predictions. During periods of declining city-wide housing demand in urban areas (like the suburbanization movement during the 1960s), the richer neighborhoods on the border of the rich areas will be the first to contract. Conversely, during periods of positive increases in city-wide housing demand (like that associated with the migration back to cities during the 1990s), the poor neighborhoods bordering the richer neighborhoods will be the first to gentrify.⁴

Our work also complements recent papers which have highlighted the theoretical and empirical importance of residential consumption externalities. For example, our theoretical model builds upon the insights of Benabou (1993) which looks at neighborhood sorting within a city where

³For a recent review of this literature, see Kolko (2007).

⁴There is a separate literature looking at the effect of direct public policies on neighborhood gentrification. See, for example, Busso and Kline (2007), Kahn et al. (2009), Rossi-Hansberg et al. (2010), and Zheng and Kahn (2011). Our work complements this literature by highlighting gentrification that is not the result of government policy but instead endogenously results from the actions of private agents responding to city-wide housing demand shocks.

there are human capital externalities and the work of Becker and Murphy (2003) which looks at neighborhood sorting in a world with exogenous income groups where all agents have a preference to live around richer neighbors. From a theoretical standpoint, our work adds to this literature by examining the dynamics of sorting and house prices in response to city-wide housing demand shocks thereby generating a gentrification process. Recent empirical work that has documented that cities are not only centers of production agglomeration, but also centers of consumption agglomeration include Glaeser et al. (2001), Autor et al. (2010), and Rossi-Hansberg et al. (2010). Most relevant for our work is the recent paper by Bayer et al. (2007) which empirically documents the importance of neighborhood consumption externalities by showing that individuals are willing to pay more to have more highly educated and wealthier neighbors, all else equal.

2 Data

Our primary measure of within-city house price growth comes from the Case-Shiller zip code level price indices.⁵ The Case-Shiller indices are calculated from data on repeat sales of pre-existing 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 Case-Shiller 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 the most recent data we have access to 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 a few things that we would like to point out about the Case-Shiller indices. First, the Case-Shiller zip code level indices are only available for certain zip codes in certain metropolitan areas. For some of our analysis, we focus our attention only on the zip codes within the main city in the MSA. For example, we look at the patterns within the city of Chicago instead of just the broad Chicago MSA. When doing so, we only use the MSAs where the main city within the MSA has at least 10 zip codes with a usable house price index.⁶ Second, we only use information for the zip codes where the price indices were computed using actual transaction data for properties within the zip code. Some of the Case-Shiller zip code price indices were calculated using imputed data or data from some of the surrounding zip codes. We exclude all such zip codes from our analysis.⁷ Third, the Case-Shiller index has the goal of measuring the change in land prices by removing structure fixed effects using their repeat sales methodology. However, this methodology

⁵The zip code indices are not publicly available. Fiserv, the 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 local movements in housing prices. See, for example, Mian and Sufi (2009). Unfortunately, we only have the data through 2008 and, as a result, we cannot systematically explore within-city house price patterns during the recent bust. We have been unsuccessful in our attempts to secure the post 2008 data from Fiserv.

⁶We list the MSAs and cities used in our Case-Shiller analysis in Appendix Table A1.

⁷As a result, the Case-Shiller zip codes that we use in our analysis do not cover the universe of zip codes within a city. Only about 50 percent of the zip codes in the city of Chicago, for example, have housing price indices computed using actual transaction data. The fraction in other cities is closer to 100 percent. A more complete discussion of the zip codes with imputed house price data can be found in the NBER working paper version of the paper: Guerrieri et al. (2010).

only uncovers changes in land prices if the attributes of the structure remain fixed over time. If households change the attributes of the structure via remodeling or through renovations, the change in the house prices uncovered by a repeat sales index will be a composite of changes in land prices and of improvements to the housing structure. When the Case-Shiller index is constructed steps are taken to minimize the effect of potential remodeling and renovations.⁸

We augment our results using information on the percent change in median house price at the neighborhood level from the 1980, 1990, and 2000 U.S. Censuses.⁹ The primary benefit of the Census data is that it is available at very fine levels of spatial aggregation. In particular, we can examine within-city differences in housing price dynamics at both the level of zip codes and census tracts. We compute within-zip code or within-census tract appreciation rates by computing the growth in the median house price across similarly defined levels of disaggregation between 1980 and 1990 and between 1990 and 2000. The Census data, however, are not without limitations. Unlike the repeat sales methodology of the Case-Shiller index, the Census data is simply the growth in the median house price within a zip code or census tract. As a result, it may be confounding movements in land prices with movements in structure quality for the median house. Moreover, the median house value, in terms of quality, could be changing over time. For example, as low quality housing gets demolished, the median price in a neighborhood may increase with no change in either land prices or structure attributes for the remaining properties. We can partially address this limitation by including controls for the changes in neighborhood housing stock characteristics when using this measure. In the NBER working paper version of our paper, we show that the zip code level price indices from Case-Shiller and the Census data track each other very closely. As a result, we feel confident in using the Census data to explore house price dynamics at the sub-zip code level.

Finally, throughout the paper, we compute MSA level house price appreciation rates using Federal Housing Finance Agency (FHFA) metro level housing price indices if the Case-Shiller house price series is not available for the MSA. For the MSAs where both data sets exist, the Case-Shiller and FHFA data track each other nearly identically.

3 New Facts About Within City House Price Dynamics

In this section, we outline a series of new facts about the nature of housing price dynamics across different neighborhoods within a city (MSA) during city (MSA) wide housing price booms. Unlike

⁸For more information on the construction of the Case-Shiller indices see the Standard and Poor's web-site which documents their home price index construction methodology. See <http://www.caseshiller.fiserv.com/about-fiserv-case-shiller-indexes.aspx>. In the NBER working paper version of the paper, we also document all the main empirical patterns in the paper using the Zillow house price index. The Zillow index, at least partially, overcomes some of the deficiencies of the Case-Shiller index in that it allows the broad attributes of the structure (e.g., square footage, number of bed rooms, etc.) to change over time. The patterns we document using the Case-Shiller index are nearly identical to the patterns we find using the Zillow index for the MSAs and time periods where both indices overlap.

⁹Most of the tract-level Census data that we use comes from the Neighborhood Change Database which is distributed by GeoLytics. The Neighborhood Change Database provides variables from the 1970, 1980, and 1990 Censuses that have been re-weighted for the 2000 tract boundaries.

previous attempts to study within-city house price movements, we analyze these patterns simultaneously for a large number of cities and for multiple time periods.¹⁰ As we show, there are many systematic patterns that emerge with respect to house price dynamics across neighborhoods within a city during city-wide housing price booms.

3.1 Fact 1: Within City House Price Growth Variation is Large

Table 1 shows the degree of between- and within-MSA variation in house price appreciation separately during the 2000-2006 period (row 1) and the 1990-2000 period (row 2). Columns 1 and 2 focus on cross-MSA variation in house price appreciation for comparison to the within-MSA or within-city variation. When focusing on the cross-MSA variation, we use FHFA data (Column 1) and Case-Shiller data (Column 2).¹¹ As seen from Table 1, there is large variation in price appreciation across MSAs during the 1990s and the 2000s. This is consistent with the well documented facts discussed in Davis et al. (2007), Glaeser et al. (2008), Van Nieuwerburgh and Weill (2010), and Saiz (2010). Specifically, the cross-MSA standard deviation in house price growth using the FHFA data was 33 percentage points during the 2000-2006 period and 17 percentage points during the 1990-2000 period.

The next three columns of Table 1 show within-city or within-MSA, cross-zip code variation in house price appreciation for the same time periods. For columns 3 and 4, we use data from the Case-Shiller indices and show the results for all available zip codes within the MSA (column 3) and then for all available zip codes in the main city of the MSA (column 4). In column 5, we show the results for the within-city cross-zip code standard deviation in house price appreciation using the Census data for the 1990-2000 period. When using the Census data in column 5, we restrict the sample to be the same as the Case-Shiller sample. The results in these columns show that the within-MSA variation during the 2000-2006 period was about one half as large as the cross-MSA variation but was still substantial at 18 percentage points. During the 1990-2000 period, the within-city variation was of the same order of magnitude as the cross-city variation at about 15 percentage points.

The final two columns of the table show within-city cross-census tract variation for 1990-2000 period using the Census data. The sample in column 6 is restricted to census tracts that overlap with the 496 zip codes used in the sample for column 5. Column 7, broadens the sample of census tracts to include all tracts in all cities that contain at least 30 census tracts in 1990. As one would expect, the within-city variation increases as the level of our definition of a neighborhood gets smaller. For example, the cross-census tract variation in house price growth in the the 1990s was roughly fifty percentage points. Collectively, the results in Table 1 show that within-MSA variation

¹⁰Papers examining within city house price movements for a given city or small set of cities during the 1970s and 1980s include Poterba (1991), Mayer (1993), Case and Shiller (1994), Case and Mayer (1996), and Case and Marynchenko (2002). Ferreira and Gyourko (2011) build upon our work and provide additional facts about the timing of booms and busts at the neighborhood level during the 2000s.

¹¹For reference, the house price appreciation rates using the FHFA MSA level index for the 1990-2000 and 2000-2006 periods for each MSA in our Case-Shiller sample are shown in Appendix Table A1.

in house price growth is around the same order of magnitude as the cross-MSA variation that has received so much attention in the literature.

3.2 Fact 2: Initially Low Price Neighborhoods Within a City Appreciate More than High Price Neighborhoods During City-Wide Housing Booms

The next fact we wish to highlight is shown in Figure 1 and Table 2. Figure 1 plots the house price appreciation rate in each zip code within the New York MSA between 2000 and 2006 (using the Case-Shiller data) against the median house price for the same zip codes in year 2000 (from the Census). As seen from the figure, there is a sharp negative relationship between the initial level of housing prices within the zip code and the subsequent appreciation rate in the zip code. On average, zip codes with lower initial housing prices within the MSA appreciated at roughly twice the rate as zip codes with higher initial housing prices within the MSA during this period.

Our choice of showing New York in Figure 1 is done for illustrative purposes. Table 2 shows the relationship between the initial median housing price and the subsequent housing price growth across neighborhoods within the city/MSA for a large selection of cities and metro areas during different time periods. Specifically, Table 2 shows the mean growth rate in property prices over the indicated time period for neighborhoods in different quartiles of the initial house price distribution within the city or metro area. The last column shows the p -value of the difference in house price appreciation rates between the properties that were initially in the top (column 1) and bottom (column 4) quartiles of the housing price distribution within the city or metro area. This table is the analog to the scatter plot shown in Figure 1. In all cases, the initial level of housing prices used to define the quartiles in period t is defined using the median level of reported house price for the neighborhood from the corresponding U.S. Census (i.e., 2000, 1990, or 1980 depending on the time period studied). The house price appreciation is measured using the Case-Shiller index.

We can conclude a few things from the results in Table 2. First, the patterns found in Figure 1 for New York for the 2000-2006 period are also found in a wide variety of other cities and MSAs during the same period. Second, as seen from Table 2, these within-city patterns are not limited to the recent period. During the 1990s, Denver and Portland experienced large housing price booms, and it was the low priced neighborhoods that appreciated at much higher rates than the high priced neighborhoods. Likewise, during the 1980s, Boston experienced a large housing price boom during which the low priced neighborhoods appreciated at much higher rates than the high priced neighborhoods. Finally, there is also some evidence that poor neighborhoods fall the most during city-wide housing price busts. For example, within San Francisco and Boston during the 1990s, the poorer neighborhoods contracted slightly more relative to the richer neighborhoods.¹²

Are the results shown in Table 2 representative of the patterns in a broader sample of cities?

¹²In Guerrieri et al. (2012) we present a case study of Detroit examining the protracted bust experienced there from 1980 through the late 2000's. In that paper, we find patterns that are consistent with those observed in San Francisco and Boston in the 1990's. Those patterns show the reverse of the gentrification patterns documented during booms within this paper.

The answer is definitely yes. To illustrate this, we estimate:

$$\frac{\Delta P_{t,t+k}^{i,j}}{P_t^{i,j}} = \mu_j + \omega_1 \ln(HP_t^{i,j}) + \epsilon_{t,t+k}^{i,j} \quad (1)$$

where $\Delta P_{t,t+k}^{i,j}/P_t^{i,j}$ is the growth in housing prices between period t and $t+k$ within neighborhood i in city or MSA j using the various house price series and $HP_t^{i,j}$ is the median house price in neighborhood i in city or MSA j in year t as measured by the U.S. Census. Given that we also include city or MSA fixed effects, μ_j , all of our identification comes from variation across neighborhoods within a city/MSA. The variable of interest from this regression is ω_1 which estimates the relationship between initial median house prices in the neighborhood and subsequent neighborhood housing price growth. We run this regression using different neighborhood house price series and for different time periods. For all specifications, we weight the data using the number of owner occupied housing units in the neighborhood during period t (from the Census).

To conserve space, we do not show the results of this regression in the main text. However, in the online robustness appendix that accompanies this paper, we show the results of this specification for different time periods, different measures of house price growth, and different samples. The results across the different specifications are very consistent. For cities experiencing a city-wide housing price boom, it is the neighborhoods with the initially low housing prices that appreciate the most. For example, during the 2000-2006 period, restricting the sample to all zip codes with a Case-Shiller house price index, and using the Case-Shiller index to measure zip code housing price growth, our estimate of ω_1 is -0.24 with a standard error of 0.05.

One concern one may have about the results in Figure 1, Table 2, and the regressions results from equation (1) is that they are driven by transitory measurement error or temporary shocks. For example, a neighborhood that got a temporary negative shock to house prices today would have both lower house prices today and a higher growth rate between today and tomorrow as the temporary negative shock abated. This story is not, however, responsible for the results we document. To illustrate this fact, we can re-estimate equation (1) instrumenting $HP_t^{i,j}$ with house prices in the neighborhood 10 years earlier. Specifically, our estimate of ω_1 is -0.26 with a standard error of 0.05 when we instrument for 2000 neighborhood house price levels using 1990 neighborhood house price levels. Notice, this estimate is nearly identical to the OLS estimate reported in the prior paragraph.

In the robustness appendix, we also show that the difference between the house price appreciation of initially low price neighborhoods within the city and initially high price neighborhoods within the city increases with the size of the city-wide housing price boom.

3.3 Fact 3: The Variance in Appreciation Rates is Also Higher for Initially Low Price Neighborhoods During City-Wide Housing Booms

Returning to Figure 1, another feature of the data for the New York MSA is that the house price appreciation rate among initially low priced neighborhoods exhibits substantially more variability than the house price appreciation rates among initially high priced neighborhoods. In particular, the standard deviation of housing price growth between 2000 and 2006 for neighborhoods in the lowest initial house price quartile for the New York MSA was 29 percent while the standard deviation of house price growth during the same time period for neighborhoods in the top initial house price quartile for the New York MSA was only 6 percent. The difference is significant at less than 1 percent level.

This difference in variability of growth rates between initially low priced neighborhoods and initially high priced neighborhoods within a city during a city-wide housing price boom is a robust feature of the data across the many cities in our sample. Again, we formally document these facts in the online robustness appendix that accompanies the paper. When cities experience housing price booms, the variability in house price growth among initially low price neighborhoods is much higher than the variability of house price growth among initially high price neighborhoods. Pooling the MSAs in our sample, the standard deviation of housing price growth between 2000 and 2006 for neighborhoods in the lowest initial house price quartile within each MSA was 61 percent while the standard deviation of house price growth in the top initial house price quartile was 46 percent. This difference is also significant at the less than 1 percent level.

3.4 Fact 4: Poor Neighborhoods Closer to Rich Neighborhoods Appreciate More than other Poor Neighborhoods During City-Wide Housing Booms

What explains the increased variation in house price appreciation across the poorer neighborhoods? In this subsection, we highlight the role of proximity to richer neighborhoods as being an important determinant of house price appreciation of poorer neighborhoods within a city during city-wide housing price booms. Moreover, we show that the relationship between the proximity to richer neighborhoods and house price appreciation of poorer neighborhoods remains strong even after we control for proximity to jobs within the city and to fixed natural amenities within the city.¹³

¹³There are many theories that can explain within city differences in house price appreciation. For example, if cities are viewed as centers of production agglomeration, as in the classic work by Alonso (1964), Mills (1967), and Muth (1969), neighborhoods that are close to jobs will have higher land prices than neighborhoods that are farther away. Likewise, Rosen (1979) and Roback (1982) show that land prices within the city can differ based on their proximity to a desirable fixed natural amenity. In this section, we show that proximity to rich neighborhoods is an important determinant of within city house price movements above and beyond proximity to jobs and proximity to fixed natural amenities within the city.

To begin, we describe the data by estimating the following regression:¹⁴

$$\frac{\Delta P_{t,t+k}^{i,j}}{P_t^{i,j}} = \mu_j + \beta_1 \ln(\text{Dist}_t^{i,j}) + \Gamma X_t^{i,j} + \Psi Z_t^{i,j} + \epsilon_{t,t+k}^{i,j} \quad (2)$$

where $\ln(\text{Dist}_t^{i,j})$ measures the log of the distance (in miles) to the nearest zip code in the city that resides in the top quartile of neighborhoods with respect to median housing prices in period t .¹⁵ The variable of interest in the above regression is β_1 , the coefficient on $\ln(\text{Dist}_t^{i,j})$. All of our regressions also include city fixed effects, μ_j . As a result, our identification comes from within-city variation. We report heteroscedasticity robust standard errors that are clustered at the city level.

When estimating the above regression, our sample only includes low housing price neighborhoods within the city. We define low housing price neighborhoods as those neighborhoods whose median housing price at time t is in the bottom half of neighborhoods with respect to median housing prices across all neighborhoods in city j at time t .¹⁶ As above, we use the Census data to define the level of period t median housing prices for each neighborhood when segmenting the sample. The vector $X_t^{i,j}$ includes a series of variables designed to control for initial differences across the neighborhoods. These controls include the log of median household income of residents in neighborhood i in period t , the log of the median initial house price in neighborhood i in period t , the fraction of the residents in neighborhood i in period t that are African American, and the fraction of the residents in neighborhood i in period t that are Hispanic. When the Census data is used to compute housing price appreciation, we also include a vector of variables to proxy for the change in structure quality within the neighborhood between t and $t+k$.¹⁷

We also include a vector $Z_t^{i,j}$ which is designed to control for the other potential mechanisms which can generate differential price movements across neighborhoods within a city. Specifically, we control for the average distance to the closest center business district within the city as reported by the 1982 Census of Retail Trade.¹⁸ The Census data provide another measure of proximity to jobs in that they track how long it takes for individuals in the neighborhood to get to work. Given this, we also include the mean commuting time of individuals within neighborhood i during period t as an additional control. Finally, we control for the distance to fixed natural amenities like major lakes, rivers, and oceans that are within 10 miles of the city.

¹⁴Given that neighborhoods within a city have different amounts of homeowners or potential housing market transactions, all regressions are weighted by the number of owner-occupied housing units in the neighborhood during period t .

¹⁵We measure distance from the centroid of each neighborhood.

¹⁶Sometimes in the text we will refer to these neighborhoods as “poor neighborhoods”. We do this for expositional ease. We also used an income based measure to define poor neighborhoods. Given the very high correlation between neighborhood average income and neighborhood housing prices, the results are broadly consistent if we segment neighborhoods by initial income as opposed to initial house prices.

¹⁷These controls include: the change in the fraction of homes in the tract that are single-family-detached, the change in the fraction that have zero or one bedrooms, the change in the fraction that have two bedrooms, the change in the fraction that have three bedrooms, the change in the fraction built in the past 5 years, the change in the fraction built between 5 and 20 years ago, the change in the fraction built between 20 and 40 years ago, and the change in the fraction built between 40 and 50 years ago.

¹⁸The CBD data can be found at <http://www.census.gov/geo/www/cbd.html>.

Table 3 shows the results of the above regressions using different time periods, different housing price appreciation measures, and different levels of aggregation. The first two columns show the results for the 2000-2006 period where we use Case-Shiller house price indices. In column 1, we exclude the Z vector of controls in order to gauge their impact when in column 2 we include both the X and Z vectors of controls. The specific sample for the results in columns 1 and 2 is all zip codes which were (1) in Case-Shiller cities where a Case-Shiller index exists and (2) in the bottom half of zip codes within in the city in 2000 with respect to median house prices. There are 236 such zip codes.

The results in columns 1 and 2 show that there is some systematic variation in house price appreciation rates among the poor neighborhoods during the 2000-2006 period. In particular, it is the initially low price neighborhoods in 2000 which were in close proximity to the high price neighborhoods that appreciated more than otherwise similar initially low price neighborhoods. These results hold even after controlling for proximity to the city's Central Business District (CBD), average commuting times, distance to fixed natural amenities and the X vector of neighborhood controls (column 2).¹⁹ In terms of economic magnitudes, the estimates are non-trivial. For example, the results in column 2 suggest that low priced neighborhoods that were roughly 4 miles away from higher price neighborhoods appreciated at 12.4 percentage point lower rates than low priced neighborhoods that were roughly 1 mile away from higher priced neighborhoods ($0.062 * 2$, p-value < 0.01). Given that the average neighborhood house price appreciation rate for the neighborhoods in our sample during this period was roughly 90 percent, the estimated relationship with distance to high price neighborhoods is non trivial.

In columns 3 - 6, we show similar results for the 1990 - 2000 period. All specifications in these columns control for both the full vector of X and Z controls. In columns 3 and 4, we use the Case-Shiller data on Case-Shiller zip codes. The difference between the two columns is that in column 4 we also include an additional regressor: $\ln(Dist_t^{i,j}) * Bust_{t,t+k}^j$ where $Bust_{t,t+k}^j$ is an indicator variable taking the value of one if city j experienced non-positive housing price growth between t and $t + k$. We do not include this variable in the 2000-2006 period because all cities in our sample experienced a positive house price increase. However, as seen from Appendix Table A1, some cities in our Case-Shiller sample experienced real housing price declines during the 1990s. As seen from column 4, the relationship between house price growth among poor neighborhoods close to and far from high price neighborhoods differs depending on whether the city experienced a positive or non-positive housing demand shock during the period. In particular, the cities that did not experience a housing price boom had very little difference in housing price growth between poor neighborhoods that were close to high price neighborhoods and poor neighborhoods that were farther away ($-0.067 + 0.070$). However, for cities that experienced a positive city-wide housing

¹⁹In a recent paper, Glaeser et al. (2012) focus on house price appreciation of neighborhoods close to city centers. They show neighborhoods close to city centers appreciate more than other neighborhoods during the 2000s - particularly if poverty is concentrated in the city center. Building on our methodology, they find that about one-third of the effect of the house price growth of neighborhoods close to the center city can be attributed to the endogenous gentrification story that we highlight.

price increase, the estimated relationship in the 1990s mirrored what we found in the 2000 - 2006 period. The estimated coefficient on log distance during the 1990s was -0.067

In columns 5 and 6, we show that the results are roughly consistent using the Census data during the 1990-2000 period. We define neighborhoods as census tracts and use different samples. In column 5, the sample is all census tracts within only the 28 Case-Shiller cities. In column 6, the sample is census tracts in any U.S. city that has at least 30 census tracts contained within the city. There were 173 U.S. cities in 1990 that met this condition.²⁰ Column 6 shows that the main results still hold when examining price movements at the level of census tracts and that the results are not simply limited to Case-Shiller cities. In column 7, we show the results for the 1980-1990 period. The specification in column 7 is analogous to the one shown in column 6 aside from the fact that it looks at the 1980-1990 period for the 110 U.S. cities in 1980 that had at least 30 census tracts contained within the city. The patterns in the 1980s are similar to those found in the 1990s and early 2000s. It is important that the results are similar between the Case-Shiller and Census housing price measures. In Section 5, we explore the response of within-city house price dynamics to exogenous city-wide housing demand shocks as a test of our endogenous gentrification theory. To get enough power, we need to use the large samples shown in columns 6 and 7 for which only the Census housing price measures are available.²¹

In summary, the results of this section show that (1) there is a tremendous amount of variation across neighborhoods within a city with respect to house price appreciation during a city-wide house price boom, (2) it is the poor neighborhoods that systematically appreciate more than the rich neighborhoods during a city-wide house price boom, (3) the variance in house price growth is also higher among the poor neighborhoods during a city-wide house price boom, and (4) among the poor neighborhoods it is the poor neighborhoods that are in close proximity to the richer neighborhoods that systematically appreciate at the highest rates during a city-wide house price boom. Again, we think these facts are interesting in their own right. Additionally, these facts will be consistent with the theory of endogenous gentrification that we develop in the next section.

4 Model

In this section, we develop a spatial model of housing prices across neighborhoods within a city consistent with the facts documented in the previous section and based on a positive neighborhood externality: people like to live next to richer neighbors. We do not micro-found the source of this externality and leave the model flexible enough to encompass alternative possible stories behind the preference for richer neighborhoods, such as lower crime rates, higher school quality, and more positive neighborhood amenities. Whatever micro-foundation one prefers, the presence of such an

²⁰In the online robustness appendix, we specify in detail all our sample criteria when using the expanded set of census tracts. In particular, we discuss how we select census tracts that are consistently defined over time.

²¹We also performed a series of additional robustness specifications on our results. These results are shown in our online robustness appendix. For example, Brueckner and Rosenthal (2008) show that the age and quality of the housing stock could be an important determinant of which neighborhoods will subsequently gentrify. Our results are robust to the inclusion of the initial age of the housing stock in our specifications.

externality generates a gentrification process in response to a city-wide increase in housing demand. We are interested in exploring the relationship between gentrification and house price dynamics in response to city-wide housing demand shocks.

4.1 Set up

Time is discrete and runs forever. We consider a city populated by N_t infinitely lived individuals comprised of two types: a continuum of rich households of measure N_t^R and a continuum of poor households of measure N_t^P . Each period households of type s , for $s = R, P$, receive an exogenous endowment of consumption goods equal to y^s , with $y^P < y^R$.²²

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 $n_t^s(i)$ the measure of households of type s who live in location i at time t and by $h_t^s(i)$ the size of the house they choose. In each location, there is a maximum space that can be occupied by houses which is normalized to 1,²³ that is,

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

Moreover, market clearing requires

$$\int_{-\infty}^{+\infty} n_t^s(i) di = N_t^s \text{ for } s = R, P. \quad (3)$$

The key ingredient of the model is that there is a positive location externality: households like to live in areas where more rich households live. Each location i has an associated neighborhood, given by the interval centered at i of fixed radius γ . Let $H_t(i)$ denote the total space occupied by houses of rich households in the neighborhood around location i ,²⁴ that is,

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

²²The assumption that there are only two types of households (rich and poor) is for simplicity. One could extend the model to allow for a continuum of income types. The model's implications would then depend on the shape of the income distribution and on the way in which the externality is modeled. In particular, the externality would be equal to some weighted average of the income of the households who live in each neighborhood.

²³Our notion of space is uni-dimensional: if there is need for more space to construct houses we assume that the neighborhoods have to expand horizontally. We could enrich the model with a bi-dimensional notion of space, by allowing a more flexible space constraint in each location. For example, we could imagine some form of adjustment cost to construct in each location, so that in reaction to a demand shock the city can expand both in the horizontal and in the vertical dimension. Our model is the extreme case with infinite adjustment cost on the vertical dimension and no adjustment costs on the horizontal dimension. Our mechanism would go through if we allow some convex adjustment costs to the vertical margin.

²⁴An alternative is to define the neighborhood externality $H_t(i)$ as the measure of rich households living in the neighborhood around location i (or even as their average income). However, this would make the model less tractable without affecting the substance of the mechanism. A more interesting extension would be to relax the assumption that a neighborhood has a fixed size and make the concept of a neighborhood more continuous. Again the main mechanism of the model would survive this change, but the price schedule would look smoother.

Households have non-separable utility in non-durable consumption c and housing services h . The location externality is captured by the fact that households enjoy their consumption more if they live in locations with higher $H_t(i)$. The utility of a household of type s located in location i at time t is given by $u^s(c, h, H_t(i))$, where $u(\cdot)$ is weakly concave in c and h . For tractability, we assume that u takes the following functional form: $u^s(c, h, H) = c^\alpha h^\beta (A + H)^{\delta^s}$, where α , β , and δ^s are non negative scalars and A is a constant that prevents utility from being zero when H takes the value of zero.²⁵ Moreover, we assume that $\delta^R \geq \delta^P$, so that rich households who generate the externality benefit from it at least as much as poor households. We want to stress that all of the implications of our model go through even if $\delta^R = \delta^P$.

On the supply side, there is a representative firm who can build housing in any location $i \in (-\infty, +\infty)$. There are two types of houses: rich houses (type R) and poor houses (type P). Each type of household only demands houses of his own type. The marginal cost of building houses of type s is equal to C^s , with $C^R \geq C^P$. If the firm wants to convert houses of type \tilde{s} into houses of type s , he has to pay $C^s - C^{\tilde{s}}$. The (per square foot) price of a house for households of type s in location i at time t is equal to $p_t^s(i)$. Hence there is going to be construction in any empty location i as long as $p_t^s(i) \geq C^s$. Moreover, if the firm wants to construct a house of type s in a location occupied by a house of type \tilde{s} , he has to pay the converting cost and the additional cost of convincing households of type \tilde{s} to leave. Hence, there is going to be construction of houses of type s in any location occupied by agents of type \tilde{s} if $p_t^s(i) \geq C^s - C^{\tilde{s}} + p_t^{\tilde{s}}(i)$.

Finally, there is a continuum of risk-neutral competitive intermediaries who own the houses and rent them to the households. The intermediaries are introduced for tractability. If we allowed the households to own their houses, nothing would change in steady state, but the analysis of a demand shock would be more complicated.²⁶ The (per square foot) rent for a house of type s in location i at time t is denoted by $R_t^s(i)$. 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 equations hold:

$$p_t^s(i) = R_t^s(i) + \left(\frac{1}{1+r} \right) p_{t+1}^s(i) \text{ for all } t, i, s. \quad (5)$$

4.2 Equilibrium

An equilibrium is a sequence of rent and price schedules $\{R_t^R(i), R_t^P(i), p_t^R(i), p_t^P(i)\}_{i \in R}$ and of allocations $\{n_t^R(i), n_t^P(i), h_t^R(i), h_t^P(i)\}_{i \in R}$ such that households maximize utility, the representative firm maximizes profits, intermediaries maximize profits, and markets clear.

Because of full mobility, the household's maximization problem reduces to a series of static

²⁵Davis and Ortalo-Magné (2010) show that a Cobb-Douglas relationship between housing consumption and non-housing consumption fits the data well along a variety of dimensions.

²⁶When the economy is hit by a positive demand shock, we will show that house prices appreciate by different amounts in different locations. If households own their houses this would introduce an additional source of heterogeneity in wealth which would complicate the analysis.

problems. The problem of households of type s at time t is simply

$$\begin{aligned} \max_{c, h, i \in \mathcal{I}_t^s} & c^\alpha h^\beta [A + H_t(i)]^{\delta^s}, \\ \text{s.t. } & c + hR_t^s(i) \leq y^s, \end{aligned}$$

where households take as given the function $H_t(i)$, the rent schedule $R_t^s(i)$, and the set \mathcal{I}_t^s of locations where houses for type- s households are available. Hence, conditional on choosing to live in location i at time t , the optimal house size is

$$h_t^s(i) = \frac{\beta}{\alpha + \beta} \frac{y^s}{R_t^s(i)} \text{ for all } t, s, i \in \mathcal{I}_t^s. \quad (6)$$

Households choose to live in bigger houses in neighborhoods where the rental price is lower and, conditional on a location, richer households choose bigger houses. 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, agents of type s have to be indifferent among living in different locations where houses of their type are available at time t , that is, in all $i \in \mathcal{I}_t^s$.²⁷ Then it must be that

$$U_t^s(i) \equiv \alpha^\alpha \beta^\beta \left(\frac{y^s}{\alpha + \beta} \right)^{\alpha + \beta} \frac{[A + H_t(i)]^{\delta^s}}{R_t^s(i)^\beta} = \bar{U}_t^s \text{ for all } t, s, i \in \mathcal{I}_t^s. \quad (7)$$

This, in turn, requires that

$$R_t^s(i) = K^s [A + H_t(i)]^{\frac{\delta^s}{\beta}} \text{ for all } t, s, i \in \mathcal{I}_t^s, \quad (8)$$

for some constant K^s . This expression is intuitive, as rents must be higher in locations with a stronger externality. Moreover, rich households are more willing to pay higher rents for a given locational externality, all else equal.

Proposition 1 *If $\delta^R \geq \delta^P$, there exists an equilibrium with full segregation. If $C^R = C^P$, an equilibrium with full segregation exists if and only if $\delta^R \geq \delta^P$.*

This proposition proves that if $\delta^R \geq \delta^P$ there exists an equilibrium with full separation. In particular, the proof proceeds by constructing an equilibrium with full segregation, where the rich households are concentrated in the city center, while the poor households live at the periphery of the city. This is the equilibrium we will focus on in the rest of the analysis. However, there may be other equilibria with full segregation with more centers of agglomeration of the rich households.²⁸ Moreover, let us highlight that there may be other types of equilibria, e.g. we can construct an equilibrium with partial segregation, where intervals with only poor people alternate with intervals

²⁷If there was a location with construction of type s and no type s households living there, the intermediaries would be willing to decrease the rent to 0 inducing households of type s to move into that location.

²⁸It is interesting to notice that, as long as these centers are far enough from each other, the implications in terms of house prices are isomorphic to the equilibrium we focus on.

where poor and rich people coexist.²⁹

Let us now proceed to the construction of our equilibrium. As a normalization, we choose point 0 as the center of the city. Then, $\mathcal{I}_t^R = [-I_t, I_t]$ and $\mathcal{I}_t^P = [-\bar{I}_t, -I_t) \cup (I_t, \bar{I}_t]$, for some $\bar{I}_t > I_t > 0$. Both the size of rich neighborhoods, I_t , and the size of the city, \bar{I}_t , are equilibrium objects. Given that such an equilibrium is symmetric in i , from now on, we can restrict attention to $i \geq 0$.

Since rich households live in locations where there are no poor, it must be that $h_t^R(i) n_t^R(i)$ is either equal to 1 or to 0 and is equal to 1 for all $i \in [0, I_t]$. Then, we can easily derive the externality function $H_t(\cdot)$ as follows:

$$H_t(i) = \begin{cases} 2\gamma & \text{for } i \in [0, I_t - \gamma] \\ \max\{\gamma + I_t - i, 0\} & \text{for } i \in (I_t - \gamma, \bar{I}_t] \end{cases}. \quad (9)$$

That is, neighborhoods close to the city center are richer and enjoy the maximum degree of externality, while the farther a location is from the center the smaller the strength of the externality. Figure 2 shows the externality $H_t(i)$ for a given t as a function of the location. If $\bar{I}_t > I_t + \gamma$, there are going to be locations at the margin of the city where the externality has zero effect. From now on, we assume that the measure of poor households, N_t^P , is sufficiently large so that $\bar{I}_t > I_t + \gamma$.

Combining (8) and (9), we obtain

$$K^R = R_t^R(I_t) (A + \gamma)^{-\frac{\delta_R}{\beta}} \quad \text{and} \quad K^P = R_t^P(\bar{I}_t) A^{-\frac{\delta_P}{\beta}}, \quad (10)$$

so that we can rewrite the rent schedules as

$$R_t^R(i) = R_t^R(I_t) \left(1 + \frac{\min\{\gamma, I_t - i\}}{A + \gamma}\right)^{\frac{\delta_R}{\beta}} \quad \text{for } i \in [0, I_t], \quad (11)$$

$$R_t^P(i) = R_t^P(\bar{I}_t) \left(1 + \frac{\max\{\gamma + I_t - i, 0\}}{A}\right)^{\frac{\delta_P}{\beta}} \quad \text{for } i \in (I_t, \bar{I}_t]. \quad (12)$$

From the optimizing behavior of the representative firm, it must be that the price of a poor house at the boundary of the city is equal to the marginal cost C^P .³⁰ Moreover, the price of a rich house at the boundary of the rich neighborhoods must be equal to the price of a poor house, which is the compensation needed to vacate poor households living there, plus the additional cost of transforming a poor house to a rich one. This implies that $p_t^P(\bar{I}_t) = C^P$ and $p_t^R(I_t) = p_t^P(I_t) + C^R - C^P$. In equilibrium prices are constant over time and hence arbitrage conditions (5) require that for each location $i \in \mathcal{I}_t^s$ prices satisfy

$$p_t^s(i) = \frac{1+r}{r} R_t^s(i) \quad \text{for all } t, i, s. \quad (13)$$

²⁹However, we can show that there is no equilibrium with full integration, that is, where poor and rich agents simultaneously live in every occupied location.

³⁰We assume that the economy starts with no housing and a fixed measure of poor and rich agents N^P and N^R . Then at date 0 there is positive construction which pins down the housing price at the boundary of the city. See footnote 30 for the case of a negative shock to the measure of agents if the economy starts with a positive stock of housing.

Combining these conditions we obtain

$$R_t^P(\bar{I}_t) = \frac{r}{1+r}C^P \text{ and } R_t^R(I_t) = R_t^P(I_t) + \frac{r}{1+r}(C^R - C^P), \quad (14)$$

where, from (8) and (10), we have

$$R_t^P(I_t) = \frac{r}{1+r}C^P \left(\frac{A + \gamma}{A + \max\{\gamma + I_t - \bar{I}_t, 0\}} \right)^{\frac{\delta_P}{\beta}}. \quad (15)$$

Combining these last two expressions with (11), (12), and (13) allows us to determine the rent and the price schedules as a function of I_t and \bar{I}_t only. Figure 2 also shows the shape of the price schedule as a function of the location.

In our full segregation equilibrium, the rich households are concentrated in the city center, while the poor are located at the periphery. Equilibrium prices reflect the fact that locations that are further away from the center of the rich enclave and closer to the space occupied by poor households are less appealing. In particular, prices are the highest in the center of the rich neighborhoods. As we move away from the center, prices start declining because the space in the neighborhood occupied by rich households goes down. This segregation equilibrium is sustained by the fact that the poor are unwilling to lower their non housing consumption by paying higher rent to get the larger neighborhood externality.

To complete the characterization of the equilibrium, we need to determine the size of the city, \bar{I}_t , and the size of the rich neighborhoods, I_t . Using market clearing (3) together with the optimal housing size (6) and the fact that $\mathcal{I}_t^R = [-I_t, I_t]$ and $\mathcal{I}_t^P = [-\bar{I}_t, -I_t) \cup (I_t, \bar{I}_t]$, we obtain the following expressions for I_t and \bar{I}_t :

$$I_t = \gamma + (A + 2\gamma)^{-\frac{\delta_R}{\beta}} \left\{ \frac{\beta}{\alpha + \beta} \frac{y^R N^R}{2K^R} - \frac{\beta}{\delta_R + \beta} (A + \gamma)^{-1} \left[(A + \gamma)^{\frac{\beta + \delta_R}{\beta}} - (A + 2\gamma)^{\frac{\beta + \delta_R}{\beta}} \right] \right\} \quad (16)$$

$$\bar{I}_t = I_t + \gamma + A^{-\frac{\delta_P}{\beta}} \left\{ \frac{\beta}{\alpha + \beta} \frac{y^P N^P}{2K^P} - \frac{\beta}{\delta_P + \beta} A^{-1} \left[A^{\frac{\beta + \delta_P}{\beta}} - (A + \gamma)^{\frac{\delta_P}{\beta} + 1} \right] \right\}. \quad (17)$$

As intuition suggests, the rich neighborhoods cover a larger portion of the city when N_t^R (the number of rich people) or y^R (the income of rich people) are higher, and when the marginal cost of construction C^R or the interest rate r are lower. Moreover, the city overall is bigger when there are more rich households when the rich households are richer, when there are more poor households or when the poor are richer. Likewise, the city is larger when the marginal cost of construction C^P or the interest rate are lower.

Finally, we have to check that the households choose their location optimally, that is, we have to check that the rich would not prefer to move to a poor neighborhood and vice versa. More precisely, we need to prove that $U_t^R(i) \leq \bar{U}_t^R$ for all $i \in [I_t, \bar{I}_t]$ and $U_t^P(i) \leq \bar{U}_t^P$ for all $i \in [0, I_t]$ where $U_t^s(i)$ is defined in expression (7). In the Appendix, we show that both these conditions are satisfied if $\delta^R \geq \delta^P$, completing the proof of the Proposition.

4.3 Demand shock

We are now interested in analyzing how house prices, both at an aggregate and at a disaggregate level, react to an unexpected increase in the demand for housing.³¹ We will do so by focusing on the equilibrium with full segregation where all rich agents live in a connected interval, which we constructed in the previous section.

In equilibrium, the aggregate price level is given by

$$P_t = \frac{2}{\bar{I}_t} \int_0^{I_t} p_t^R(i) di + \frac{2}{\bar{I}_t - I_t} \int_{I_t}^{\bar{I}_t} p_t^P(i) di,$$

where, from the analysis in the previous section,

$$p_t^R(i) = \left[C^P \left(1 + \frac{\gamma}{A} \right)^{\frac{\delta_P}{\beta}} + C^R - C^P \right] \left(1 + \frac{\min\{\gamma, I_t - i\}}{A + \gamma} \right)^{\frac{\delta_R}{\beta}} \quad \text{for } i \in [0, I_t], \quad (18)$$

$$p_t^P(i) = C^P \left(1 + \frac{\max\{\gamma + I_t - i, 0\}}{A} \right)^{\frac{\delta_P}{\beta}} \quad \text{for } i \in (I_t, \bar{I}_t], \quad (19)$$

with I_t and \bar{I}_t given by (16) and (17).

For concreteness, we analyze the economy's reaction to a migration shock, but the price dynamics are equivalent if we consider any shock that increases housing demand, such as a positive income shock or a reduction in the interest rate. Imagine that at time $t+1$ the economy is hit by an unexpected and permanent increase in the population N . Let us assume that the measure of both rich and poor households increase proportionally, that is, $N_{t+1}^s = \phi N_t^s$ with $\phi > 1$ for $s = P, R$. We now show that the aggregate level of house prices permanently increases and prices in locations with a higher initial price level typically react less than prices in locations where houses are cheaper to start with and which are closer to the expensive neighborhoods. The new rich households moving into the city want to live close to other rich households, so that the poor neighborhoods close to the rich ones get gentrified and the poor households who used to live there move towards the periphery. This is what we refer to as endogenous gentrification. The house prices in gentrified neighborhoods are driven up due to our externality.

Let us define the function $g_t(\cdot) : [C^P, \bar{p}] \mapsto [1, \infty)$, where $g_t(p)$ denotes the average gross growth rate between time t and $t+1$ in locations where the initial price is equal to p , that is,

$$g_t(p) = E_{t+1} \left[\frac{p_{t+1}(i)}{p_t(i)} \mid p_t(i) = p \right],$$

and $\bar{p} \equiv [C^P (1 + \gamma/A)^{\frac{\delta_P}{\beta}} + C^R - C^P] (1 + \gamma/(A + \gamma))^{\delta_R/\beta}$. The next proposition shows that after

³¹The long run reaction of house prices would be symmetric in the case of a negative shock if we introduce some degree of depreciation that is big enough relative to the shock. However, after a negative shock the economy would not jump to the new steady state immediately, but there would be some transitional dynamics. Contact the authors if you are interested in the full analysis of a negative demand shock in the presence of depreciation. Given our data, we only focus on housing booms resulting from positive housing demand shocks.

an unexpected permanent positive demand shock, the aggregate price level permanently increases and the price growth rate is higher in locations that had lower price levels initially, whenever prices are higher than the minimum level, consistent with Fact 2.³²

Proposition 2 *Suppose the city exhibits an equilibrium with full segregation where all rich agents live in a connected interval. Imagine that at time $t + 1$ the economy is hit by an unexpected and permanent increase in population, that is, $N_{t+1}^s = \phi N_t^s$ with $\phi > 1$ for $s = P, R$. Then there is a permanent increase in the aggregate price level P_t , and*

$$E_{t+1} \left[\frac{p_{t+1}(i)}{p_t(i)} \mid p_t(i) = \bar{p} \right] < E_{t+1} \left[\frac{p_{t+1}(i)}{p_t(i)} \mid p_t(i) < \bar{p} \right].$$

Moreover, if the shock is large enough, $g_t(p)$ is non-increasing in p for all $p > C^P$.

Figure 3 illustrates the response of house prices in different locations to a positive demand shock (a proportional increase in population). Given that the city is symmetric, the figure represents only the positive portion of the real line. One can notice that both the size of the city \bar{I}_t and the size of the rich neighborhoods I_t expand, and prices remain constant at the two extremes: in the richest locations in the center of the city and far enough away from the rich neighborhoods. Most importantly, prices strictly increase in the rich neighborhoods on the border of the poor neighborhoods where the externality is below its maximum level and, even more, in the poor neighborhoods that are physically close to the rich neighborhoods. Clearly, this makes the aggregate level of prices in the city increase permanently.

The next proposition shows the main implication of our model: among the locations with initial level of price equal to C^P , the ones that appreciate the most are closer to the richer neighborhoods.

Proposition 3 *Suppose the city exhibits an equilibrium with full segregation where all rich agents live in a connected interval. Imagine that at time $t + 1$ the economy is hit by an unexpected and permanent increase in population, that is, $N_{t+1}^s = \phi N_t^s$ with $\phi > 1$ for $s = P, R$. Then*

$$\frac{d(p_{t+1}(i)/p_t(i))}{di} \leq 0 \text{ for } p_t(i) = C^P.$$

Among the poor neighborhoods, it is the poor neighborhoods in close proximity to the richer neighborhoods that should appreciate the most during a city-wide housing demand shock. This proposition underlies the variation in appreciation rates among the poorer neighborhoods. The poor neighborhoods next to the rich neighborhoods experience large price increases because they gentrify. Rich households expand into the neighborhood thereby increasing the desirability of being in those neighborhoods. This proposition is consistent with Facts 3 and 4 and lies at the heart of our following empirical work. An increase in city-wide housing demand - perhaps do to an in

³²See the online appendix for all the proofs that are not in the text.

migration of rich residents - will cause poor neighborhoods on the border of richer neighborhoods to gentrify. The empirical work that follows does, in fact, show strong support for this prediction.

Proposition 4 *Suppose the city exhibits an equilibrium with full segregation where all rich agents live in a connected interval. Imagine that at time $t + 1$ the economy is hit by an unexpected and permanent increase in population, that is, $N_{t+1}^s = \phi N_t^s$ with $\phi > 1$ for $s = P, R$. Then the growth rate in the aggregate price level is larger the larger is the increase in ϕ and, if the shock is large enough, $d^2 g_t(p) / dp d\phi \leq 0$ for all $p > C^P$ where the derivative is well-defined.*

This proposition shows that if two identical cities are hit by demand shocks of different sizes, the one hit by the larger shock is going to feature both a higher aggregate price growth rate and more price appreciation among the poor neighborhoods due to a higher degree of gentrification, consistent with the facts. It is also easy to show that two cities with different initial income composition react differently to the same demand shock. In particular, if the shock is large enough, the initially richer city is the one that features both a higher aggregate price growth rate and higher within-city house price convergence.

5 Housing Price Dynamics and Proximity to Rich Neighborhoods: Exogenous Housing Demand Shock

In Section 3, we established empirical relationships that are consistent with our theory of endogenous gentrification. In particular, we documented that during a city-wide housing boom it is the poor neighborhoods bordering the rich neighborhoods that appreciate the most. However, the prior descriptive results do not make any claims about causation. In this section, we directly explore whether an exogenous shock to housing demand in a city affects property prices differentially across neighborhoods within that city in a way that is consistent with our theory.

5.1 Exogenous Housing Demand Shock: Bartik Instrument

To measure exogenous shocks to local housing demand for each city j between t and $t + k$, we use the variation in national earnings by industry between t and $t + k$. This approach of imputing exogenous income shocks for local economies was developed by Bartik (1991) and has been used extensively by others in the literature as a measure of local labor demand shocks.³³ In doing so, we are explicitly equating local income shocks with local housing demand shocks. As shown by Blanchard and Katz (1992), such positive Bartik-type local income shocks cause an influx of population from other cities which puts upward pressure on city-wide housing prices.

To compute exogenous changes in city-level income, we use the initial industry composition of residents within the city's MSA in period t . Note, that even though we are examining house price dynamics within a given city, our estimate of the local demand shock is based on the industry

³³See, for example, Blanchard and Katz (1992), Notowidigdo (2010), and Charles et al. (n.d.).

mix of the MSA as a whole. Given the amount of commuting into and out of the city from the suburbs (in both directions), we feel that the MSA income shock is a broader proxy for city-wide changes in housing demand. Then, for each MSA^{*j*}, we compute the predicted income growth for the MSA using the initial industry shares and the growth in income for individuals in those industries between t and $t + k$ for the entire U.S. (excluding residents from MSA^{*j*}).

For our results examining the neighborhood response to local housing demand shocks, we focus on the 1980-1990 period. We analyze this period because there is significant variation across MSAs in predicted MSA-wide average income growth based on industry composition during this time. Specifically, we use the five percent samples from the 1980 and 1990 IPUMS data to compute MSA-level predicted income growth. For this procedure, we use two-digit industry classifications. Our measure of income is individual earnings. The only restrictions we place on the data are that the individual had to be employed full-time (worked 48 weeks or more in the prior year and usually worked more than 30 hours per week) and had to be between the ages of 25 and 55. Again, when computing the predicted income growth for a given MSA, we exclude the residents of that MSA in calculating the national income growth between 1980 and 1990 for each of the industries. To compute the MSA predicted income growth, we simply multiply the industry growth rate in earnings by the fraction of people between 25 and 55 in each MSA working full time in those industries in 1980.

There is a large amount of variation in actual income growth by industry between 1980 and 1990. For example, the Security, Commodity Brokerage, and Investment Companies industry had a real appreciation of annual earnings of roughly 59% during the 1980s. Likewise, the Legal Services industry had a real appreciation of annual earnings of 55%. On the other hand, the Trucking Services and Warehousing and Storage industry only had a real appreciation of annual earnings of 3%. As a result of differences in industry mixes across MSAs, there is a nontrivial amount of predicted income variation across the MSAs. To set notation, we define $\widehat{IncShock}_{t,t+k}^j$ to be the predicted income growth for the MSA corresponding to city j between t (1980) and $t + k$ (1990) based on the industry mix of residents in the MSA in 1980.

For our results in this section, our sample includes all MSAs that contain a city which has at least 30 census tracts within the city. This will be the same sample that we used in column 7 of Table 3. The mean of $\widehat{IncShock}_{t,t+k}^j$ across these MSAs was 19.5 percent with a standard deviation of 1.5 percent. As shown by others in the literature, the predicted Bartik income growth for the MSA does in fact predict actual MSA income growth. A simple regression of actual MSA income growth during the 1980s on $\widehat{IncShock}_{t,t+k}^j$ yields a coefficient on $\widehat{IncShock}_{t,t+k}^j$ of 2.26 (with a standard error of 0.60) and a F-stat of 14.01.

To examine the effect of exogenous housing demand shocks on within-city house price dynamics, we estimate a specification similar to (2):

$$\frac{\Delta P_{t,t+k}^{i,j}}{P_t^{i,j}} = \mu_j + \beta_1 \ln(Dist_t^{i,j}) + \beta_2 \ln(Dist_t^{i,j}) * \widehat{IncShock}_{t,t+k}^j + \Gamma X_t^{i,j} + \Psi Z_t^{i,j} + \epsilon_{t,t+k}^{i,j} \quad (20)$$

where $\Delta P_{t,t+k}^{i,j}/P_t^{i,j}$, $\ln(Dist_t^{i,j})$, $\widehat{IncShock}_{t,t+k}^j$, $X_t^{i,j}$, $Z_t^{i,j}$, and μ_j are defined as above. We are interested in β_2 , the coefficient on the interaction term. With this regression, we are asking whether, for a given sized city income shock, poor neighborhoods within the city in close proximity to rich neighborhoods within the city appreciate more than otherwise similar poor neighborhoods that are farther away. For our measure of housing price growth, we use the Census data and for the measure of neighborhood we use census tracts. As with the results above that use census house price data, we include controls for changes in the neighborhood housing stock characteristics as part of our X vector. Otherwise, the X vector is the same. Because our instrument is an estimated regressor, we bootstrap our standard errors.³⁴

5.2 Results

The results of estimating the above equation are shown in Table 4. In column 1, we estimate (20) as it is specified. The variable of interest is in the second row and provides an estimate of β_2 . As with the simple descriptive results shown in Table 3, an exogenous shock to city income results in house prices increasing more in poor neighborhoods that are in close proximity to rich neighborhoods (coefficient = -2.33 with a standard error of 0.49). To help interpret the economic magnitude, we consider the differential housing price response to otherwise similar poor neighborhoods which are 1 and 4 miles away from a rich neighborhood in response to the MSA receiving a one-standard deviation MSA-level predicted income shock of 1.5 percent. Given the estimated coefficient, a census tract that starts in the bottom half of the city-wide house price distribution in 1980 appreciates by 7.0 percentage points more when they are 1 mile from a high price census tract relative to an otherwise similar census tract that is 4 miles away ($-2.33 * 2 * 0.015$). This result is non-trivial and is in line with the general descriptive patterns shown in Table 3.

In column 2, we re-run the same specification replacing the log distance variable with two dummy variables measuring the proximity to high housing price census tracts. We do this to explore in greater depth whether the relationship between housing price growth and proximity to richer neighborhoods declines monotonically as the poor neighborhoods become farther away from the rich neighborhoods. Specifically, we replace the log distance measure with dummies indicating whether the census tract was between 0 and 1 miles and between 1 and 3 miles, respectively, to the nearest census tract in the top quartile of the city-wide house price distribution in 1980. We are interested in the coefficient on the interaction between these dummies and the instrumented change in neighborhood income. The house price response to an exogenous city-wide income shock is positive and statistically different from zero for both distance ranges. Reassuringly, the house price response is four times as large for census tracts that are between 0 and 1 miles from the high housing price neighborhoods relative to census tracks between 1 and 3 miles from the high housing price neighborhoods (p-value of difference of the two coefficients = 0.02). Given the average size of

³⁴For each iteration in the bootstrap procedure, we sample with replacement from the 1980 and 1990 IPUMS data to calculate the first stage. Next, we sample with replacement from the census tracts in each of the cities in our sample to calculate the second stage. We report standard errors calculated by repeating this process 2,500 times. We also estimated the standard errors using only 500 repetitions. The results were very similar.

census tracts, almost all the initially poor census tracts within 1 mile of a rich census tract actually abut the rich neighborhood. In other words, the biggest responses in prices within a city to a city-wide housing demand shock are for those poor census tracts that border richer census tracts. The estimated magnitudes are also nontrivial. In response to a one standard deviation instrumented income shock, poor census tracts within 0 and 1 miles and within 1 and 3 miles appreciated by 6.1 and 1.5 percentage points more, respectively, than poor census tracts more than 3 miles away.

We wish to make four additional comments about the results in Table 4. First, given that we are including city fixed effects, all our results are identified off of within-city variation. Second, as with the results in Table 3, we are controlling for proximity to CBD, average commuting times, and proximity to fixed natural amenities. Given this, our results are being driven by proximity to rich neighborhoods above and beyond proximity to the center business district or fixed natural amenities within the city. Third, although not shown, the results hold broadly for the 1990s as well but power is more of an issue during that time period. Finally, we explored whether the responsiveness of house prices in poor neighborhoods that were close to rich neighborhoods to Bartik shock was greater in cities where housing supply was more inelastic. To do this, we further interacted our distance to rich neighborhoods and our distance measure multiplied by the Bartik shock with Saiz’s measure of MSA housing supply elasticity (Saiz, 2010). The point estimates of the triple interaction indicated that the price response of poor neighborhoods bordering rich neighborhoods was stronger in more inelastic cities. However, the standard error was much too large to be conclusive.

6 Housing Price Dynamics, Proximity to Rich Neighborhoods, and Evidence of Neighborhood Gentrification

In this section, we examine more deeply the main mechanism of our model: after a city-wide housing demand shock the poor neighborhoods next to the rich neighborhoods are the ones that appreciate the most because they are the ones where rich households migrate to. This implies that neighborhoods that experience higher house price appreciation should also show signs of gentrification. We show three sets of results documenting that the neighborhoods we focus on which experienced higher house price appreciation also showed signs of gentrification.

6.1 A Descriptive Analysis

To analyze whether neighborhoods that experienced a rapid growth in prices also experienced signs of gentrification, we estimate the following descriptive relationship:

$$\frac{\Delta Y_{t,t+k}^{i,j}}{Y_t^{i,j}} = \mu_j + \beta \frac{\Delta P_{t,t+k}^{i,j}}{P_t^{i,j}} + \Gamma X_t^{i,j} + \Psi Z_t^{i,j} + \epsilon_{t,t+k}^{i,j}$$

where $\Delta P_{t,t+k}^{i,j}/P_t^{i,j}$, $X_t^{i,j}$, $Z_t^{i,j}$, and μ_j are defined as above, and $\Delta Y_{t,t+k}^{i,j}/Y_t^{i,j}$, is the growth rate of median household income from t to $t+k$ in neighborhood i in city j . The regression asks whether

or not a neighborhood that experiences higher house price growth than other neighborhoods within the city also experiences higher income growth than other neighborhoods within the city. In this specification, we are equating neighborhood income growth with neighborhood gentrification. In the work below, we try to explore different and perhaps broader measures of neighborhood gentrification. However, given that the force in our model that drives the house price appreciation of poor neighborhoods that abut rich neighborhoods is the exodus of poor residents which are replaced by richer residents, we think neighborhood income growth is a good summary statistic for the mechanism we are trying to highlight.

The results of this regression, using different samples, different measures of housing price growth and different levels of aggregation for a neighborhood, are shown in Table 5. In columns 1 - 4, we look at the relationship between income growth and house price growth across neighborhoods during the 1990s. In column 5, we explore the relationship during the 1980s. In columns 1 and 2, we restrict our analysis to Case-Shiller zip codes using Case-Shiller house price data (column 1) and Census house price data (column 2), respectively, to compute housing price growth. In columns 3 - 5, we use Census house price data to compute housing price growth and define neighborhoods at the level of a census tract. In column 3, we explore census tracts in Case-Shiller cities. In columns 4 and 5, we explore all census tracts in cities that have at least 30 consistently measured census tracts over the decade. As with our analysis in Tables 3 and 4, we also restrict all samples to include only neighborhoods within the city that are in the bottom half of the city's house price distribution in period t . As a result, the samples used for columns 1-5 of Table 5 are analogous to samples used in columns 3-7 of Table 3, respectively.

The main take away from Table 5 is that there is a strong relationship between neighborhood income growth and neighborhood house price growth regardless of the house price measure, regardless of the level of aggregation and regardless of the time period. Although not shown, these results still hold even if we drop the X and Z vectors of neighborhood controls from the regressions.

6.2 An Ex-Post Analysis

In this subsection, we perform a different analysis to highlight the spatial nature of gentrification. In particular, we identify all neighborhoods within cities that ex-post can be classified as having gentrified according to some broad definition and we examine their spatial proximity to high income neighborhoods. Our goal is to illustrate that when gentrification occurs, it almost always occurs in poorer neighborhoods that border higher income neighborhoods.

For our analysis, we define a gentrifying neighborhood as a neighborhood within a city that starts with median neighborhood house prices in the bottom half of the city's house price distribution in period t and where the median real income of neighborhood residents grows by either 50 percent or 25 percent between t and $t + k$. We use our broadest sample of cities with at least 30 consistently measured census tracts for the 1980s and 1990s. Specifically, the samples we use are the same as those used in columns 6 and 7 of Table 3.

For this analysis, we simply regress a dummy variable for whether the initially poor neighbor-

hood gentrified by some income growth metric on distance dummies to the nearest rich neighborhood and city fixed effects. As above, we define rich neighborhoods as those neighborhoods that were in the top quartile of the city’s median house price distribution in period t . We define four dummy variables to measure the poor neighborhood’s proximity to the nearest rich neighborhood: between 0 and 0.5 miles, between 0.5 and 1 mile, between 1 and 2 miles, and between 2 and 3 miles. Finally, we run such regressions separately for two measures of gentrification: neighborhood real income growth greater than 50 percent during the decade and neighborhood real income growth greater than 25 percent during the decade.

The results of these regressions are shown in Table 6. The results are quite striking. Take, for example, the results where gentrification is defined as a poor neighborhood having average neighborhood income growth increasing by at least 50 percent during the decade. Between 1980 and 1990, 11% of all poor neighborhoods gentrified by this metric. The comparable number between 1990 and 2000 was 5.9%. During the 1980s, the probability of gentrification was 7.0 percentage points higher if the census tract was between 0 and 0.5 miles from a high house price neighborhood than for a poor census tract that was more than 3 miles away from a rich neighborhood (column 1 of Table 3, p-value < 0.01). The coefficient is large in economic magnitude. Given the base gentrifying rate was 11 percentage points, a poor census tract being within 0 and 0.5 miles was associated with a 64 percent increase in the probability of gentrification. During the 1990s, poor census tracts that were within 0.5 mile of a rich census tract were 97 percent more likely to gentrify than poor census tracts that were more than 3 miles away from the rich census tracts. Similar patterns are found in both decades if we define gentrification as neighborhood income growth increasing by 25%.

Poor census tracts that were within 0.5 miles of a rich census tract almost always abutted the rich census tract. Among the poor census tracts, as one moves farther away from the nearest rich census tract, the probability of gentrification declines monotonically. These results are also seen in Table 6.

The results are consistent with the housing price dynamics in our model: poor neighborhoods tend to gentrify only when they are in close proximity to existing rich neighborhoods. The results show that there is definitely a spatial nature to the gentrification process.

6.3 Exogenous Housing Demand Shock

In the final subsection, we complete our analysis by assessing whether exogenous city-wide demand shocks cause poor neighborhoods in close proximity to richer neighborhoods to endogenously gentrify. As we saw in Table 4, poor neighborhoods in close proximity to rich neighborhoods had house price growth that was larger than other poor neighborhoods in response to positive city-wide Bartik shocks. In this subsection, we show that these close neighborhoods were also more likely to experience signs of gentrification.

To look for signs of endogenous gentrification, we estimate the following:

$$G_{t,t+k}^{i,j} = \mu_j + \beta_1 \ln(Dist_t^{i,j}) + \beta_2 \ln(Dist_t^{i,j}) * \widehat{IncShock}_{t,t+k}^j + \Gamma X_t^{i,j} + \epsilon_{t,t+k}^{i,j} \quad (21)$$

where $G_{t,t+k}^{i,j}$ is a measure of gentrification in neighborhood i of city j between t and $t+k$. Specifically, we use three measures of $G_{t,t+k}^{i,j}$: the percentage increase in neighborhood income between t and $t+k$, the percentage point change in the poverty rate between t and $t+k$, and the percentage point change in the fraction of residents in the neighborhood who had a college degree or more. Aside from the change in the dependent variable, (21) is analogous to (20) estimated above for neighborhood housing price growth. Moreover, the sample and definition of $\widehat{IncShock}_{t,t+k}^j$ are exactly the same as the specifications used to estimate (20) in Table 4.³⁵

The results from estimating (21) are shown in Table 7. In response to a city-wide housing demand shock, it is the poor census tracts that are in close proximity to the rich census tracts that are much more likely to experience rising incomes, declines in the poverty rate, and rising educational attainment of residents relative to poor census tracts that are farther from the rich census tracts. Specifically, in response to a one-standard deviation instrumented income shock, poor census tracts that were 1 mile from rich neighborhoods experienced income growth that was 1.7 percentage points higher than poor neighborhoods that were 4 miles away. Given that the average census tract in our sample experienced income growth of 14.9 percent during the decade, this represents an increase in income of 11 percent for poor neighborhoods that are close to rich neighborhoods in response to a one standard deviation instrumented income shock. Likewise, poor neighborhoods that are 1 mile from the rich neighborhoods experienced 23 percent lower increases in the poverty rate and 25 percent higher increases in the fraction of residents with a college degree or more relative to otherwise similar poor neighborhoods that are 4 miles from the rich neighborhoods.

7 Conclusion

In this paper, we explore the response of housing price dynamics across neighborhoods to a city-wide housing demand shock. The main empirical fact that we document is that poor neighborhoods on the border of richer neighborhoods experience the largest increase in house price appreciation in response to a city-wide housing demand shock. In particular, we find that during the 1980s poor neighborhoods that bordered richer neighborhoods had house prices that appreciated by 7.0% more than otherwise similar poor neighborhoods which were farther away from rich neighborhoods in response to a one standard deviation Bartik shock. Moreover, these neighborhoods simultaneously experienced a more dramatic rise in resident income and education and a more dramatic decline in the resident poverty rate.

We propose a model where positive city-wide housing demand shocks endogenously result in neighborhood gentrification that is consistent with the facts. The key assumption is that all individuals prefer neighborhoods populated by richer households as opposed to neighborhoods populated by poorer households. While we do not take a stand on the exact source of the externality, we have in mind that richer neighborhoods have lower levels of crime, higher provisions of local public

³⁵Additionally, (21) does not include the Z vector of controls which were designed to capture other reasons why house prices differed across neighborhoods.

goods, better peer effects, and a more extensive provision of service industries (like restaurants and entertainment options).

Our work adds to that of Brueckner (1977) and Kolko (2007) by showing that there is a large spatial component to neighborhood gentrification. Aside from our results showing that it is the poor neighborhoods next to the rich neighborhoods that respond the most to city-wide housing demand shocks, we also show that proximity to rich neighborhoods is a defining feature of neighborhood gentrification. Choosing neighborhoods that have ex-post gentrified, we find that the probability of gentrification is 64 percent higher for neighborhoods that were within 0.5 miles of an existing rich neighborhood than otherwise similar neighborhoods that were farther away.

We also present a series of new facts about within city house price movements during city-wide housing booms. As far as we know, we are the first paper to systematically analyze the differential housing price dynamics across neighborhoods within cities during city-wide housing price booms. We show that the within-city variation in house prices is almost as large as the well documented cross-city variation in house prices during the last three decades. Also, we document that poor neighborhoods experience larger housing price increases and a greater variation in housing price increases relative to richer neighborhoods during city-wide housing price booms. The larger the city-wide housing price boom, the more poor neighborhoods appreciate relative to rich neighborhoods. Although our gentrification results only exploit the variation in house price appreciation among poor neighborhoods during city-wide housing price booms, the facts we document suggest there are many other interesting patterns in the data worth exploring in future work.

References

- Alonso, William**, *Location and Land Use*, Harvard University Press, 1964.
- Autor, David H., Christopher J. Palmer, and Parag A. Pathak**, “Housing Market Spillovers: Evidence from the End of Rent Control in Cambridge Massachusetts,” 2010.
- Bartik, T.J.**, “Who Benefits from State and Local Economic Development Policies?,” *W.E. Upjohn Institute for Employment Research: Kalamazoo, Mich.*, 1991.
- Bayer, Patrick, Robert McMillan, and Fernando Ferreira**, “A Unified Framework for Measuring Preferences for Schools and Neighborhoods,” *The Journal of Political Economy*, 2007, 115 (4), 588–638.
- Becker, Gary and Kevin Murphy**, *Social Economics: Market Behavior in a Social Environment*, Harvard University Press, 2003.
- Benabou, R.**, “Workings of a city: location, education, and production,” *The Quarterly Journal of Economics*, 1993, 108 (3), 619.
- Blanchard, O.J. and L.F. Katz**, “Regional Evolutions,” *Brookings Papers on Economic Activity*, 1992, pp. 1–75.
- Brueckner, Jan K.**, “The Determinants of Residential Succession,” *Journal of Urban Economics*, 1977, 4 (1), 45–59.
- and **Stuart S. Rosenthal**, “Gentrification and Neighborhood Housing Cycles: Will America’s Future Downtowns be Rich?,” 2008.
- Busso, Matias and Patrick Kline**, “Do Local Economic Development Programs Work? Evidence from the Federal Empowerment Zone Program,” Unpublished 2007.
- Case, Karl E. and Christopher J. Mayer**, “Housing Price Dynamics within a Metropolitan Area,” *Regional Science and Urban Economics*, 1996, 26 (3-4), 387–407.
- and **Maryna Marynchenko**, “Home Price Appreciation in Low-and Moderate-Income Markets,” *Low-income Homeownership: Examining the Unexamined Goal*, 2002, pp. 239–256.
- and **Robert J. Shiller**, “A Decade of Boom and Bust in the Prices of Single-Family Homes: Boston and Los Angeles, 1983 to 1993,” *New England Economic Review*, 1994, pp. 40–40.
- Charles, K.K., E. Hurst, and M.J. Notowidigdo**, “Manufacturing Busts, Housing Booms, and Declining Employment: A Structural Explanation.”
- Davis, Morris A. and Francois Ortalo-Magné**, “Household Expenditures, Wages, Rents,” *Review of Economic Dynamics*, 2010.
- , — , and **Peter Rupert**, “What’s Really Going on in Housing Markets?,” Federal Reserve Bank of Cleveland, Economic Commentary 2007.
- Ferreira, Fernando and Joseph Gyourko**, “Anatomy of the Housing Boom in U.S. Neighborhoods and Metropolitan Areas, 1993-2009,” Unpublished Manuscript 2011.

- Glaeser, Edward L., Jed Kolko, and Albert Saiz**, “Consumer city,” *Journal of Economic Geography*, 2001, 1 (1), 27.
- , **Joseph Gyourko, and Charles Nathanson**, “Housing Dynamics,” Updated Version of NBER Working Paper 12787 2008.
- Glaeser, E.L., J.D. Gottlieb, and K. Tobio**, “Housing Booms and City Centers,” *American Economic Review, Papers and Proceedings*, 2012, 102 (3), 127–33.
- Guerrieri, Veronica, Daniel Hartley, and Erik Hurst**, “Endogenous Gentrification and Housing Price Dynamics,” Updated Version of NBER Working Paper 16237 2010.
- , —, and —, “Within-City Variation in Urban Decline: The Case of Detroit,” *American Economic Review, Papers and Proceedings*, 2012, 102 (3), 120–26.
- Kahn, Matthew E., Ryan Vaughn, and Jonathan Zasloff**, “The Housing Market Effects of Discrete Land Use Regulations: Evidence from the California Coastal Boundary Zone,” 2009.
- Kolko, Jed**, “The Determinants of Gentrification,” Unpublished manuscript 2007.
- Mayer, Christopher J.**, “Taxes, Income Distribution, and the Real Estate Cycle: Why all Houses do not Appreciate at the Same Rate,” *New England Economic Review*, 1993, pp. 39–39.
- Mian, Atif and Amir Sufi**, “The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis,” *Quarterly Journal of Economics*, 2009, 124 (4).
- Mills, Edwin S.**, “An Aggregative Model of Resource Allocation in a Metropolitan Area,” *The American Economic Review*, 1967, pp. 197–210.
- Muth, Richard F.**, *Cities and Housing*, Univ. of Chicago Press, 1969.
- Nieuwerburgh, S. Van and P.O. Weill**, “Why Has House Price Dispersion Gone Up?,” *The Review of Economic Studies*, 2010.
- Notowidigdo, M.J.**, “The Incidence of Local Labor Demand Shocks,” 2010.
- Poterba, James M.**, “House Price Dynamics: The Role of Tax Policy and Demography,” *Brookings Papers on Economic Activity*, 1991, pp. 143–203.
- Roback, Jennifer**, “Wages, Rents, and the Quality of Life,” *The Journal of Political Economy*, 1982, 90 (6), 1257.
- Rosen, Sherwin**, “Wage-based Indexes of Urban Quality of Life,” *Current Issues in Urban Economics*, 1979, 3.
- Rossi-Hansberg, Esteban, Pierre-Daniel Sartre, and Raymond Owens III**, “Housing Externalities,” *Journal of Political Economy*, 2010, 118 (3), 485–535.
- Saiz, A.**, “The Geographic Determinants of Housing Supply,” *Quarterly Journal of Economics*, 2010.
- Zheng, S. and M.E. Kahn**, “Does Government Investment in Local Public Goods Spur Gentrification? Evidence from Beijing,” NBER Working Paper 2011.

Housing Price Growth versus Initial Housing Price 2000-2006: New York MSA

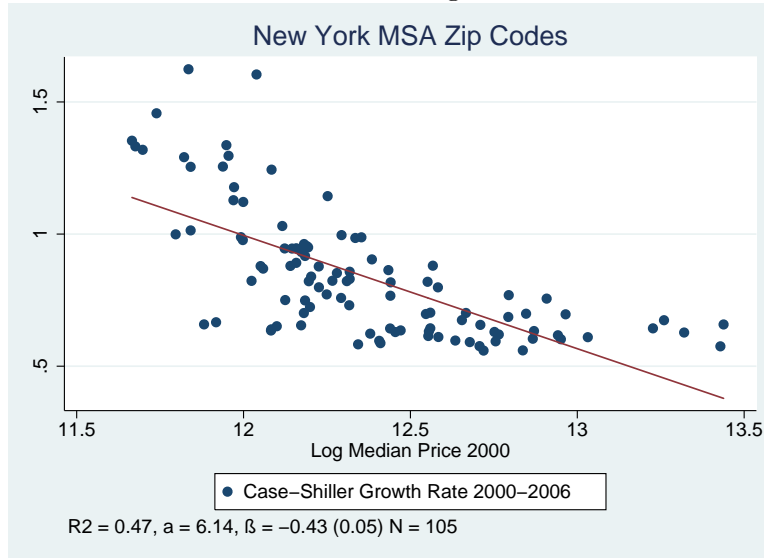


Figure 1: Figure shows the initial house price in a zip code in 2000 versus the subsequent house price growth in that zip code between 2000 and 2006. The sample includes all zip codes within the New York metro area for which a Case-Shiller house price index exists. We measure the initial house price using median home value from the 2000 Census. We use the Case-Shiller index to compute the growth rate in house price between 2000 and 2006.

Model Generated Externality and House Prices across Neighborhoods

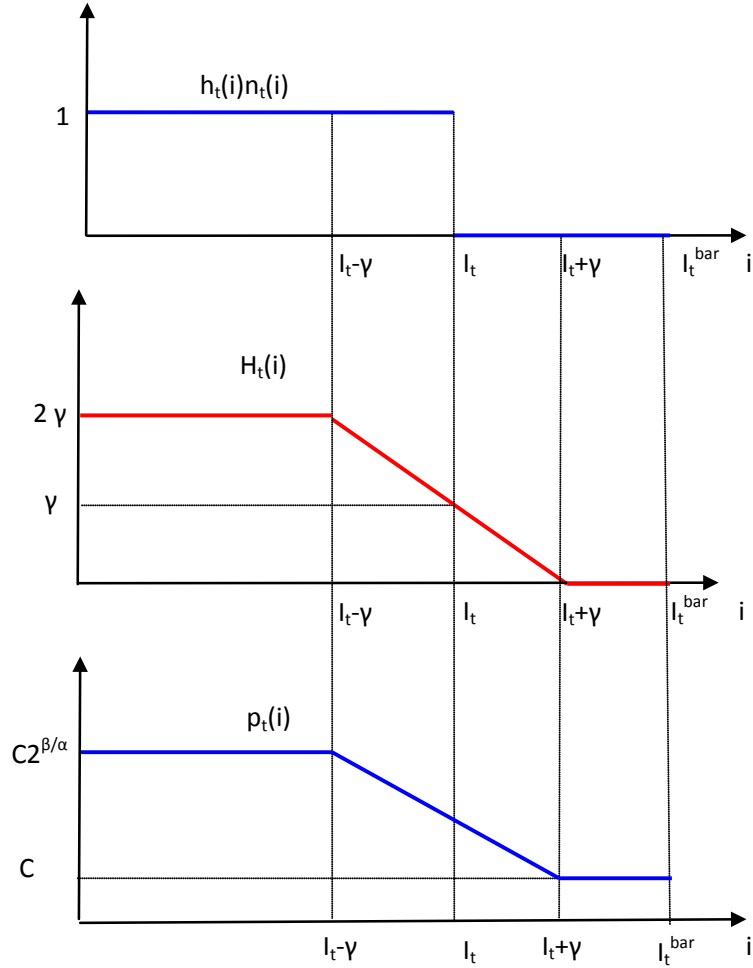


Figure 2: Figure shows the model predicted relationship between the size of the rich neighborhood (top panel), the value of the neighborhood externality (middle panel), and the house price in the neighborhood (bottom panel). For this figure, we assume $C^P = C^R$. This is done for illustrative purposes.

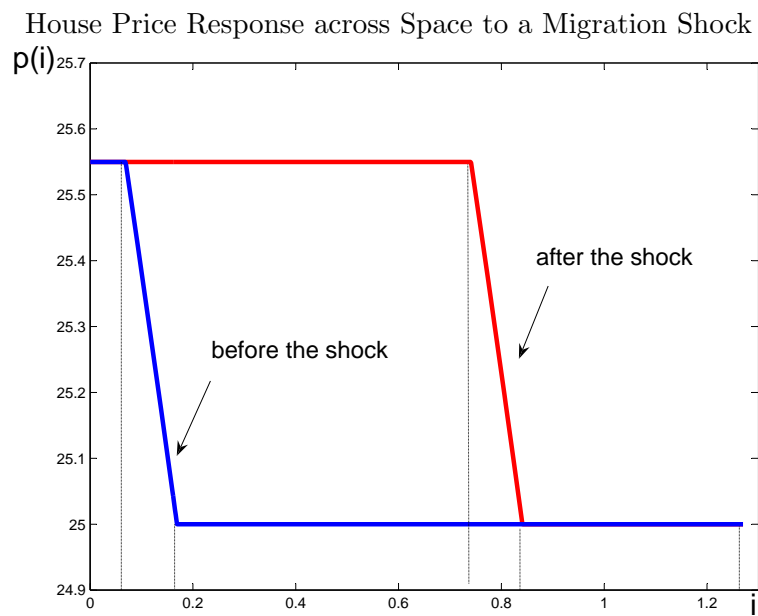


Figure 3: We set $\alpha = .8, \beta = .8, \delta^R = .2, \delta^P = 0, A = 1, \gamma = .1, r = .03, y^R = 1, y^P = .5, C^R = C^P = 25, N^R = N^P = .5$. The shock is an unexpected and permanent increase in ϕ from $\phi = 1$ to $\phi = 5$.

Table 1: House Price Growth Variation Across MSAs and Across Neighborhoods Within an MSA or City

Time Period	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2000 - 2006 (observations)	0.33 (384)	0.42 (20)	0.18 (1,693)	0.18 (497)	-	-	-
1990 - 2000 (observations)	0.17 (348)	0.21 (17)	0.16 (1,498)	0.17 (496)	0.15 (496)	0.40 (4,021)	0.54 (16,161)
Variation	Between MSA	Between MSA	Within MSA	Within City	Within City	Within City	Within City
Neighborhood Definition	-	-	Zip Code	Zip Code	Zip Code	Census Tract	Census Tract
House Price Data	FHFA	Case Shiller	Case Shiller	Case Shiller	Census	Census	Census

Notes: Table shows the between MSA standard deviation of house price growth (columns 1 and 2) and the within-city/MSA standard deviation across neighborhoods (remaining columns) for different house price measures, different time periods, and different definitions of neighborhoods. The Case-Shiller data is available for the 1990s and the 2000s. The Census house price growth measure is available only during the 1990s. For column 1, we use all 384 MSAs available in the FHFA data. For column 2, we use the 20 MSAs for which Case-Shiller reports an index. For columns 3 - 5, we use all available zip codes for which a reliable Case-Shiller index exists. See the text for details. In column 6, we use all census tracts that overlap the zip codes used in column 5. In column 7, we use all census tracts in all cities where there at least 30 census tracts within the city. See text for additional details. For the Census data, the top and bottom 1 percent of neighborhoods with respect to median home price growth are dropped.

Table 2: Housing Price Growth by Initial Price Quartile, Case-Shiller Data

	(1) Quartile 4	(2) Quartile 3	(3) Quartile 2	(4) Quartile 1	(5) p-val of Quartile 4 = Quartile 1
<u>2000 - 2006: Housing Booms</u>					
Chicago, City Level	0.53	0.66	0.72	0.88	0.00
Chicago, MSA Level	0.47	0.50	0.49	0.69	0.00
New York City, MSA Level	0.64	0.76	0.86	1.11	0.00
Boston, MSA Level	0.40	0.47	0.54	0.61	0.00
Los Angeles, MSA Level	1.21	1.41	1.58	1.76	0.00
San Francisco, MSA Level	0.35	0.41	0.49	0.61	0.00
Washington D.C., MSA Level	1.29	1.37	1.49	1.61	0.00
<u>1990 - 1997: Housing Booms</u>					
Denver, MSA Level	0.51	0.50	0.52	0.89	0.00
Portland, MSA Level	0.41	0.52	0.49	0.69	0.00
<u>1984 - 1989: Housing Booms</u>					
Boston, MSA Level	0.65	0.70	0.75	0.84	0.00
<u>1990 - 1997: Housing Busts</u>					
San Francisco, MSA Level	-0.08	-0.07	-0.11	-0.14	0.00
Boston, MSA Level	-0.05	-0.08	-0.13	-0.11	0.00

Notes: This table shows the mean Case-Shiller house price appreciation rates for neighborhoods grouped by quartile of initial housing prices, during different time periods. Quartile 4 has the highest initial price zip codes within the city while quartile 1 has the lowest initial price zip codes within the city. Each row labels a city or metro area for a given time period.

Table 3: Regression of Neighborhood House Price Appreciation on Distance to Nearest High-Price Neighborhood and Other Controls, Across Different Samples With Different House Price Measures

Time Period	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Distance to Nearest High-Price Neighborhood	-0.061 (0.021)	-0.062 (0.019)	-0.044 (0.029)	-0.067 (0.037)	-0.231 (0.042)	-0.136 (0.032)	-0.140 (0.029)
Log Distance to Nearest High-Price Neighborhood * City Wide Bust Indicator	-	-	-	0.070 (0.044)	0.068 (0.051)	0.077 (0.029)	0.083 (0.032)
House Price Measure/ Neighborhood Aggregation	C-S Zip Code	C-S Zip Code	C-S Zip Code	C-S Zip Code	Census Census Tract	Census Census Tract	Census Census Tract
Time Period	00-06	00-06	90-00	90-00	90-00	90-00	80-90
Vector of Z Controls Included	No	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	236	236	223	223	3,099	7,955	4,253
Mean Log Distance to Nearest High-Price Neighborhood	1.23	1.23	1.22	1.22	0.401	0.499	0.322
Std. Dev. Log Distance to Nearest High-Price Neighborhood	0.524	0.524	0.488	0.488	0.778	0.719	0.716

Note: Table shows regression of neighborhood house price appreciation between period t and $t+k$ on log distance to nearest high price neighborhood within the neighborhood's city, city fixed effects, and a vector of neighborhood controls. High price neighborhoods are those neighborhoods that are within the top quartile of average neighborhood house prices in year t . We restrict our analysis in this table to those neighborhoods within the city which were in the bottom half of the house price distribution in period t . See text for additional sample descriptions and discussion of the controls included. Robust standard errors, clustered by city, are shown in parentheses. All regressions are weighted by the number of owner occupied housing units in the neighborhood in the initial year.

Table 4: Regression of House Price Appreciation on Proximity to Nearest High-Price Neighborhood, Census Data 1980 - 1990

	(1)	(2)
Log Dist. to Nearest High-Price Neighborhood	0.35 (0.09)	-
Log Dist. to Nearest High-Price Neighborhood * Bartik Predicted City-Wide Income Shock	-2.33 (0.49)	-
Dummy: High-Price Neighborhood Within 0 - 1 Miles	-	-0.64 (0.22)
Dummy: High-Price Neighborhood Within 1 - 3 Miles	-	-0.15 (0.10)
Dummy: High-Price Neighborhood Within 0 - 1 Miles * Bartik Predicted City-Wide Income Shock	-	4.09 (1.17)
Dummy: High-Price Neighborhood Within 1 - 3 Miles * Bartik Predicted City-Wide Income Shock	-	0.99 (0.51)

Note: Table reports the results from the regression specified by Equation (20) from the text. The sample is the same as that used in Table 3, column 7. Standard errors bootstrapped. First stage is re-sampled from IPUMS, second stage from census tract tabulations, and stratified by city (2,500 repetitions). P-value on test of whether last two coefficients in column 2 are equal is 0.020.

Table 5: Regression of Neighborhood Income Growth on Neighborhood House Price Appreciation, Across Different Samples With Different House Price Measures

	(1)	(2)	(3)	(4)	(5)
Neighborhood House Price Growth	0.145 (0.053)	0.406 (0.069)	0.076 (0.022)	0.088 (0.021)	0.076 (0.024)
House Price Measure	C-S	Census	Census	Census	Census
Neighborhood Aggregation	Zip Code	Zip Code	Census Tract	Census Tract	Census Tract
Time Period	90-00	90-00	90-00	90-00	80-90
Number of Observations	223	223	3,099	7,955	4,253
Mean Neighborhood House Price Growth	0.331	0.146	0.586	0.512	0.240
Std. Dev. Neighborhood House Price Growth	0.479	0.348	0.967	0.747	0.788

Note: Table shows regression of neighborhood income growth between period t and $t+k$ on house price appreciation, city fixed effects, and a vector of neighborhood controls. The vector of neighborhood controls is the same as in Table 3. Also, as in Table 3 and 4, we restrict our analysis to those neighborhoods in the bottom half of the neighborhood house price distribution in period t . The specifications in columns 1 and 2 use low price zip codes from Case-Shiller cities where a Case-Shiller price index exists in 1990 and 2000. The specification in column 3 uses all census tracts in Case-Shiller cities. The specifications in column 4 uses all census tracts from all cities in the U.S. which have at least 30 consistently defined census tracts between 1990 and 2000. The specification in column 5 uses all census tracts from all cities in the U.S. which have at least 30 consistently defined census tracts between 1980 and 1990. For the specifications in columns 3 - 5, we also trim the top and bottom 1 percent of the house price growth and the income growth distributions. See text for additional details. Robust standard errors clustered at the city level are in parentheses.

Table 6: Measures of Gentrification and Proximity to High Income Neighborhoods: Descriptive Analysis 1980 - 1990 and 1990 - 2000

	Gentrification Measure:			
	Neighborhood Income Growth During Time Period			
	Greater than 50%		Greater than 25%	
	(1)	(2)	(3)	(4)
	1980 - 1990	1990 - 2000	1980 - 1990	1990 - 2000
Dummy: High-Price Neighborhood Within 0 - 0.5 Miles	0.070 (0.017)	0.057 (0.028)	0.083 (0.036)	0.110 (0.041)
Dummy: High-Price Neighborhood Within 0.5 - 1 Miles	0.014 (0.007)	0.016 (0.009)	0.094 (0.020)	0.062 (0.020)
Dummy: High-Price Neighborhood Within 1 - 2 Miles	0.006 (0.008)	0.018 (0.007)	0.076 (0.020)	0.028 (0.014)
Dummy: High-Price Neighborhood Within 2 - 3 Miles	-0.005 (0.007)	0.002 (0.005)	0.023 (0.019)	0.018 (0.014)
Fraction of Neighborhoods that Gentrified	11.0%	5.9%	30.2%	19.8%
Sample Size	4,253	7,955	4,253	7,955

Note: Table shows the results from a linear probability regression of a dummy variable indication of whether a neighborhood gentrified between t and $t+k$ on the proximity of that neighborhood to an existing rich neighborhood. We define rich neighborhoods as those census tracts within a city that are in the top quartile of the period t house price distribution. The samples in columns 1 and 3 (columns 2 and 4) are the same as column 7 (column 6) of Table 3. All regressions include city fixed effects. Robust standard errors, clustered at the city level, are in parentheses.

Table 7: Estimation of Measures of Gentrification and Proximity to High Income Neighborhoods, Census Data 1980 - 1990

	Dependent Variable: Measure of Neighborhood Gentrification		
	(1)	(2)	(3)
	Growth in Median Income	Change in Poverty Rate	Change in Fraction of Residents with Bachelor's Degree
Log Dist. to Nearest High-Price Neighborhood	0.09 (0.05)	-0.04 (0.02)	0.03 (0.01)
Log Dist. to Nearest High-Price Neighborhood * Bartik Predicted City-Wide Income Shock	-0.55 (0.26)	0.22 (0.11)	-0.23 (0.08)
Mean: Dependent Variable	0.149	0.029	0.028
Sample Size	4,253	4,253	4,253

Note: Table reports the results from the regression specified by Equation (21) from the text. The sample is the same as that used in Table 3, column 7. Standard errors bootstrapped. First stage is re-sampled from IPUMS, second stage from census tract tabulations, and stratified by city (2,500 repetitions).

Table A1: MSA Level House Price Appreciation for Case-Shiller Covered MSAs

	In MSA Sample, City Sample, or Both	MSA Price Appreciation 2000 - 2006	MSA Price Appreciation 1990 - 2000
Akron	Both	3.6%	22.8%
Atlanta	Both	13.8%	14.6%
Boston	MSA Only	49.0%	12.5%
Charlotte	Both	9.2%	12.7%
Chicago	Both	36.8%	14.4%
Cincinnati	Both	7.4%	15.4%
Columbus (OH)	Both	7.4%	16.8%
Denver	Both	10.6%	65.9%
Fresno	Both	124.1%	-8.5%
Jacksonville	Both	69.4%	11.2%
Las Vegas	Both	88.8%	-0.3%
Los Angeles	Both	121.7%	-20.9%
Memphis	Both	6.0%	8.7%
Miami	Both	125.6%	15.3%
New York	Both	72.4%	0.0%
Oakland	Both	76.7%	9.0%
Philadelphia	Both	59.1%	-10.4%
Phoenix	Both	82.8%	20.8%
Portland (OR)	Both	47.0%	53.4%
Raleigh	Both	8.1%	14.5%
Sacramento	Both	96.0%	-11.7%
San Francisco	MSA Only	52.1%	18.9%
San Diego	Both	93.5%	-0.7%
San Jose	Both	44.0%	30.6%
Seattle	Both	46.7%	22.9%
St. Paul	Both	38.2%	29.9%
Tampa	Both	88.6%	7.0%
Toledo	Both	4.7%	19.7%
Washington DC	MSA Only	98.3%	-8.5%

Notes: Table shows the MSA level house price appreciation rates for each MSA for the 2000-2006 period (column 2) and the 1990-2000 period (column 3) using the FHFA MSA level house price indices. Column 1 is an indicator whether the data from these cities or MSAs are included in our MSA samples or in our main city only samples. The data from Boston, San Francisco, and Washington DC are not included in our main city sample because there are not 10 zip codes within the city that have a reliably computed price index.