NBER WORKING PAPER SERIES

HOUSING BOOMS AND CITY CENTERS

Edward L. Glaeser
Joshua D. Gottlieb
Kristina Tobio

Working Paper 17914
http://www.nber.org/papers/w17914

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2012

The authors thank the Taubman Center for State and Local Government for research support. This paper is an expanded version of a shorter essay that will appear in the AEA: Papers and Proceedings. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Housing Booms and City Centers
Edward L. Glaeser, Joshua D. Gottlieb, and Kristina Tobio
NBER Working Paper No. 17914
March 2012
JEL No. D0,R3

ABSTRACT

Popular discussions often treat the great housing boom of the 1996-2006 period as if it were a national phenomenon with similar impacts across locales, but across metropolitan areas, price growth was dramatically higher in warmer, less educated cities with less initial density and higher initial housing values. Within metropolitan areas, price growth was faster in neighborhoods closer to the city center. The centralization of price growth during the boom was particularly dramatic in those metropolitan areas where income is higher away from the city center. We consider four different explanations for why city centers grew more quickly when wealth was more suburbanized: (1) gentrification, which brings rapid price growth, is more common in areas with centralized poverty; (2) areas with centralized poverty had more employment concentration which led to faster centralized price growth; (3) areas with centralized poverty had the weakest supply response to the boom in prices in the city center; and (4) poverty is centralized in cities with assets, like public transit, at the city center that became more valuable over the boom. We find some support for several of these hypotheses, but taken together they explain less than half of the overall connection between centralized poverty and centralized price growth.

Edward L. Glaeser
Department of Economics
315A Littauer Center
Harvard University
Cambridge, MA 02138
and NBER
eglaeser@harvard.edu

Joshua D. Gottlieb
Department of Economics
Littauer Center 200
Harvard University
Cambridge, MA 02138
jdgottl@fas.harvard.edu

Kristina Tobio
Kennedy School of Government
79 JFK St- T347
Cambridge, MA 02138
kristina_tobio@ksg.harvard.edu

I. Introduction

The great housing boom that ran from 1996 to 2006 may seem like a vast nation-wide phenomenon, but there was enormous spatial heterogeneity in housing price growth over this time period. Some areas, like New York and Las Vegas, experienced enormous growth, while others, like Cleveland and Dallas, did not. The post-2006 bust then overwhelming struck the formerly booming areas. There has been considerable heterogeneity within metropolitan areas as well, and price increases were higher in central locations, especially in those metropolitan areas where poverty is also disproportionately centralized.

The heterogeneity of price growth over the boom can potentially help us sort the myriad causes of great housing explosion. By showing that there was more price growth in areas that had limited access to credit before 2000, Mian and Sufi (2009) make a compelling case that more generous lending standards helped boost prices. Across 300 metropolitan areas, four variables can explain together over 70 percent of the considerable variation in price growth between 1996 and 2006: price growth was highest in initially expensive areas, with warm winters, less density and less educated citizens. The first two effects, at least, are compatible with a model that suggests that buyers during the boom overestimated the long-run value of positive local attributes.

Within metropolitan areas, centralization appears to be a primary predictor of price growth during the boom. While there was more new construction on the urban periphery, price growth was stronger at the metropolitan center. Indeed, the relative ease of building at the edge is one explanation for why prices typically rose faster in the urban center. A second explanation is that price growth at the center also reflected an optimistic projection of the value of a real urban amenity, in this case access to the central city.

The tendency of prices to grow faster in the metropolitan core was not, however, universal and price growth was far more centralized in some metropolitan areas than in others. In particular, price growth was more centralized in those metropolitan areas where income was more suburbanized, and this tendency does not reflect a general tendency of poorer places to experience faster growth when they are closer to the city center. Figure 1 shows this curious correlation. The horizontal axis shows the coefficient from a metropolitan area-specific regression of the logarithm of Census tract median incomes on the logarithm of distance to the central business district (CBD) as of 1990. The vertical axis shows the difference in price growth in the urban center and in the periphery of the metropolitan area between 1996 and 2006. When poverty was centralized, so was price growth. If poorer people typically live where there are fewer amenities, then this fact cuts against the hypothesis that the great boom represented an overestimation of the long run value of locational assets.

After presenting these stylized facts about price growth during the boom, we present a short model that motivates our attempt to interpret this core fact. In the model, neighborhoods differ with respect to distance to the central business district, other exogenous amenities and neighborhood composition. As in Schelling (1971) and Guerrieri, Hartley, and Hurst (2011), neighborhoods are either filled with richer or poorer people, and over time, areas may change their character. The possibility of tipping greatly increases the impact that over-optimistic beliefs about fundamentals may have on poorer areas, because those incorrect beliefs lead buyers to think that tipping is imminent.

The model presents one interpretation of the core fact: price growth was faster in poor urban centers because those central, poor areas were seen as being more likely to gentrify, which would lead to faster price growth than either centralized rich areas or outlying poorer areas. In Section IV, we test this idea along with three other interpretations of the phenomenon. One possibility is that areas with centralized poverty had the lowest supply response to the housing boom at the city center, perhaps because they are older or more restricted by supply, or perhaps because even during the boom, prices remained too low to justify teardowns and new construction. Another hypothesis is that the centralization of poverty is a proxy for the centralization of employment and that areas with more centralized employment had more centralized price growth during the boom.

A final hypothesis is that centralized poverty actually reflects urban assets, not a lack of urban amenities, and that buyers valued those assets more highly during the boom. In particular, centralized poverty might reflect the presence of a stronger public transit network, which attracts the poor to the city center (Glaeser, Kahn and Rappaport, 2008). As poverty is more centralized in cities that have more public transit dependence in the city center, those cities may be more difficult to reach by automobile from outlying suburbs which might make also make those prone to gentrification.

In Section IV, we test for these possibilities with a variety of tests. We test the gentrification hypothesis in two ways. First, we follow Guerrieri, Hartley, and Hurst (2011) and focus on poorer zip codes. We control both for zip code proximity to richer areas and for the heterogeneity of income across nearby zip codes within the MSA. We find general support for the gentrification hypothesis, but controlling for these variables explains only about one-quarter of the tendency of centralized zip codes to have faster price growth in areas with more centralized poverty.

We find little support for the supply elasticity hypothesis. Supply proxies, such as the share of new housing built in the area during the 1990s, strongly negatively predict price growth, but controlling for the interaction between supply elasticity and distance to central business district and our zip code measure of new supply does little to the coefficient on the interaction between centralized poverty and distance to the central business district.

We do, however, find somewhat more evidence that areas with public transit dependence in the city center, relative to the suburbs, had faster price growth in the city center and controlling for this tendency does some of our core puzzle during the 1990s. We interpret this as suggesting that cities where car access within the city center is more difficult experienced more price growth in the urban core because buyers saw suburban development as a worse substitute for inner city properties. This interpretation is therefore, compatible with the view that the boom reflected an overestimate of the value of urban assets, because central city poverty reflects central city strength not weakness, but these questions need substantially more research.

II. Heterogeneity in Housing Booms

Heterogeneity in the impact of increased subprime lending helps explain some of this rise, for price growth was typically greater in areas where more borrowers were shut out of credit markets before 2000 (Mian and Sufi, 2009). An inelastic housing supply also matters, as

Glaeser, Gyourko and Saiz (2008) find substantially faster price growth in areas with more topographic and regulatory restrictions on building. Yet Las Vegas and Phoenix saw price explosions despite having few visible barriers to building and the outsized price growth in high income neighborhoods suggests that subprime lending was unlikely to be the only factor driving housing price appreciation.

We start with a brief overview of the cross-sectional heterogeneity in price growth during two recent booms. We use Case-Shiller repeat sales price indices for up to 300 metropolitan areas and for zip codes within those areas. These data go some distance to control for the changing composition of the houses that are being sold at different points of the housing cycle. We have defined the two most recent housing boom periods as 1982-1989 and 1996-2006 to reflect the periods during which prices were rising nationwide. We will examine how much of the variation in cross-metropolitan area and within-metropolitan area price growth over these periods can be explained with a simple set of controls.

Table 1 shows our cross-metropolitan area results for the two boom periods. Regressions (1) and (3) show results controlling only for January temperature, initial housing values and initial density levels. Regressions (2) and (4) include controls for median income and the share of the adult population with a college degree. All variables, except for January temperature come from U.S. Census data, and we used data from the last census before the beginning of each time period (1980 and 1990 respectively).

Regression (1) shows that the three core variables can explain 27 percent of the variation in house price growth from 1982 to 1989. Warmer places had less price growth during this earlier boom, perhaps because of more elastic housing supply in the Sunbelt (Glaeser and Tobio, 2008); as January temperature increases by 10 degrees price growth drops by approximately 0.05 log points, or roughly five percent. Price growth was faster in denser metropolitan areas during the 1980s. As density doubles (log density increases by 0.69), prices rose by an extra 0.08 log points. There was also a tendency of places with higher initial housing values to see faster growth. As housing values in 1980 doubled, growth from 1982 to 1989 increased by 0.1 log points.

Regression (2) includes controls for income and human capital. Interpreting the separate impact of these two variables is always somewhat difficult, both because they are highly correlated with each other and both are measured with considerable error. During the 1980s, both variables are associated with less price growth. Controlling for them causes the coefficient on initial housing values to rise substantially. One interpretation of these results is that the 1980s price boom reflected higher demand for area amenities, other than the weather. Another interpretation is that people expected there to be faster income growth in lower income areas.

Regression (3) shows our results for the 1996-2006 period. Perhaps the most striking difference from the earlier boom is that the three core variables are far better at predicting price variation during this later period. Those three variables alone can explain almost 70 percent of the variation in price growth across metropolitan areas. Moreover, as opposed to the earlier time period, January temperature was an extremely powerful, positive predictor of price growth during this time period. An extra ten degrees of January temperature is associated with 0.06 log points more price growth.

The effect of initial housing prices is even stronger. As initial housing prices double, price growth increases by 0.36 log points. Finally, metropolitan area level population density was negatively associated with price growth from 1996-2006, although the effect is smaller in magnitude than the positive effect found during the 1980s. As population density doubles, price growth drops by 0.036 log points.

Regression (4) includes our controls for income and share of adults with college degrees. In this case, the education variable is strongly negative, and the income variable is insignificant. As the share of adults in the metropolitan with a college degree increases by ten percentage points, price growth from 1996-2006 falls by 0.09 log points.

The effect of January temperature and initial housing values during the 1996-2006 boom can be interpreted as suggesting that the boom represented an overestimation of the value of area assets. After all, January temperature is a strong predictor of area growth, and higher housing values do typically reflect area amenities or labor demand (Roback, 1982). But the negative effect of college education during the boom is somewhat harder to interpret, especially since skills are such a strong predictor of long run area success (Glaeser and Saiz, 2004).

Our second table turns to within-metropolitan area evidence. In this case, we control for metropolitan area fixed effects and cluster our standard errors at the MSA level. Regressions (1) and (2) show results for the 1982-1989 period; regressions (3) and (4) show results for the 1996-2006 period. The first regression illustrates the core finding that motivates the rest of the paper: the connection between price growth and proximity to the central business district. We include two main controls: the logarithm of zip code distance to the central business district and the interaction between that variable and the income-distance gradient at the metropolitan area level.

The gradient is calculated using census tract data from the 1990 census. For each metropolitan area, we separately estimate a regression of the form

(1) *Log(Income)= Intercept + Slope • Log(Distance to Central Business District)*.

Across MSAs, the slopes range from -0.11 to 0.51 and the standard deviation is 0.11. We subtract the mean slope so that the raw effect of distance to central business district can be interpreted as the effect of centralization for an average metropolitan area.

Regression (1) indicates that during the 1982-1989 boom, prices rose by about 0.015 log points less as the distance to the Central Business District doubles. This effect gets substantially stronger in those areas where income rises more quickly with distance to the central business district. We do not mean to suggest that the income gradient interaction is causal, but the magnitude of the effect is remarkable and seems well worth understanding.

This regression shows that even controlling for the interaction between area income and distance to the central business district has no significant effect. Regression (2) shows results for 1982-1989 where we have controlled for other area attributes using data from the 2000 Census. We would have preferred to use 1980 zip code data, and use the later year because of data availability. Including a bevy of local controls has almost no impact on our core effects.

The only controls that appear to have reliably significant effects on the zip code-level housing price boom are population density, which has a positive and significant effect on price growth, and the share of the housing stock that is single family-owner occupied, which has a negative effect. In both cases, these variables corroborate the metropolitan area level regressions, which also show a positive connection between density and price growth during this time period. The 1980s boom was centered in denser areas that were closer to the central business district.

The third regression shows our results for the 1996-2006 time period. In this case, both of our key coefficients are significantly larger in magnitude. As distance to the central business district doubles, price growth drops by 0.028 log points. That effect more than doubles in areas with a steeper income-distance to central business district gradient. Figure 1 shows this cross-effect graphically.

Regression (4) includes our other controls. Again, there is no meaningful general interaction between income and distance to the central business district. In this case, the controls cause our core coefficients to drop slightly, but they remain quite large in magnitude. In the 1996-2006 boom, income and percent with college degrees both have strong effects on zip code price growth, but the effects go in opposite directions. Richer areas had less price growth, but areas with more educated inhabitants had faster price growth. If we don't control for income levels, the coefficient on adult education flips sign. One interpretation for the price growth in low income areas is that these areas increasingly had access to subprime lending (Mian and Sufi, 2009).

Still, the results leave us with two core puzzles: why did both booms push prices up more at the city center and why was this effect more pronounced in areas that had more poverty within the urban core?   Price growth at the city center may seem quite reasonable. Central real estate may be more desirable and increased demand for central locations may be harder to satiate with new supply. But why was this effect more pronounced in areas where incomes are higher on the urban periphery? We turn to that puzzle now.

III. A Model of Gentrification and Housing Booms

In this section, we present a model that is meant to illustrate one reason why the centralization of poverty might be associated with faster price growth during a bubble. The key idea is that centralized poverty might be particularly susceptible to changes in neighborhood composition that can have a dramatic effect on future prices. In this section, we begin with a static model of neighborhood choice and housing prices with heterogeneous consumers, which builds heavily on Schelling (1971) and Guerrieri, Hartley, and Hurst (2011). We then embed that static model into a dynamic setting appropriate for discussing housing price dynamics. In the dynamic version of the model, variables will typically acquire time subscripts.

The Static Model

We assume that the city has *N* housing units, each of which is identical in everything other than location, and each of which houses exactly one resident. The city is divided into *K* neighborhoods, each of which has *N*/*K* housing units. The population of the city is split between high and low human capital individuals, and the proportion of low human capital individuals in each neighborhood is denoted π_k. The income of the high and low human capital individuals is denoted y_H and y_L respectively.

Each neighborhood has a different amenity level denoted $A_{city} + a_k - \varphi(\pi_k)$, where A_{city} is a city-wide amenity level, a_k is a neighborhood specific amenity level and $\varphi(\cdot)$ is an increasing function of the share of the residents of the neighborhood who have low human capital. We let r_k denote the per-period housing costs of living in neighborhood *k*, which could represent either the rental cost or the per-period expected cost of owner-occupied housing. We also assume that individuals pay transport costs of $c(y, d_k)$, where the function depends on y to reflect heterogeneous costs in the opportunity cost of time.

We assume linear utility so an individual of type *i* living in neighborhood *k* has utility of:

(2) $y_i - r_k - c(y_i, d_k) + \gamma_i\big(A_{city} + a_k - \varphi(\pi_k)\big)$,

where γ_i is a type-specific constant that multiplies the amenity level. We assume that human capital complements neighborhood amenities so that $\gamma_H > \gamma_L$. This assumption will ensure sorting across neighborhoods, and it is one of many ways of achieving that result. If higher human capital individuals had a lower marginal utility of income, that would also produce sorting, but concavity of utility with respect to consumption would considerably complicate the dynamic model. Sorting will also result if high human capital types particularly enjoy the company of their own kind.

High and low human capital individuals have options outside of the city which implies that their reservation utility levels are $\underline{U}_H$ and $\underline{U}_L$ respectively, and these utility levels will pin down wages. We will focus on equilibria that are stable, where instability means that a slight increase in the number of less skilled individuals causes less skilled people to be willing to pay more than skilled people to live in the neighborhood. In the appendix, we prove:

Proposition 1: In a stable equilibrium, all neighborhoods will include only high human capital or low human capital individuals. For any given level of distance, there is a maximum local amenity level a_k, denoted $\overline{a}(d_k)$, for low human capital communities and a minimum local amenity level, denoted $\underline{a}(d_k)$, which equals $\overline{a}(d_k) - (\varphi(1) - \varphi(0))$, for high human capital communities. Both $\overline{a}(d_k)$ and $\underline{a}(d_k)$ are increasing with distance if and only if $\frac{\partial c(y_H, d_k)}{\partial d_k} > \frac{\partial c(y_L, d_k)}{\partial d_k}$.

The proposition shows that there are a range of amenity values for which both high and low human capital communities can exist and the size of that range depends on the size of $\varphi(1) - \varphi(0)$, the increase in amenity value associated with being a high human capital community rather than a low human capital community. When that gap gets small, then exogenous amenities completely determine whether the community is high or low human capital.

If $\overline{a}(d_k)$ and $\underline{a}(d_k)$ are rising with distance to the city center, this suggests that the overall level of human capital is also rising with distance from the city center, at least if the distribution of exogenous amenities is essentially independent of distance. The condition $\frac{\partial c(y_H, d_k)}{\partial d_k} > \frac{\partial c(y_L, d_k)}{\partial d_k}$ suggests that central cities will tend to have higher poverty rates if wealthier people have a lower marginal cost of commuting, which might occur if wealthier people invest in technologies, like cars, that have a higher fixed cost but lower variable cost of commuting (see Glaeser, Kahn and Rappaport, 2008).

We now turn to the impact of a shift in city-wide demand, which we capture with an increase in A_{city}. We assume that a marginal A_{city} does not immediately cause neighborhoods which had been below $\underline{a}(d_k)$ to shift from low to high human capital, because the low human capital communities are still stable at the higher city-wide amenity level. However, low human capital communities with amenity levels at $\overline{a}(d_k)$ will immediately switch to high human capital communities, because it is no longer an equilibrium for them to remain low human capital. If the distribution of amenity areas for any given distance level is described by a density function $f_{d_k}(a)$ between $a_{min} < \underline{a}(d_k)$ and $a_{max} > \underline{a}(d_k)$, and if we let $\theta_{d_k}(a)$ denote the share of neighborhoods at each amenity level that are high human capital, then the impact of a city-wide increase in amenities on average prices at distance d_k equals:

$$(3) \int \left(\theta_{d_k}(a)\gamma_H + \left(1-\theta_{d_k}(a)\right)\gamma_L\right) f_{d_k}(a)da + f_{d_k}(\overline{a}(d_k))(1-\theta_{d_k}(\overline{a}(d_k))\gamma_H(\varphi(1)-\varphi(0))$$

The increase in the city level amenity has two effects on the price level. First, there is the direct effect of increasing amenities, which is valued more by higher human capital communities. Second, there is the discrete switch of low human capital neighborhoods to high human capital neighborhoods at the border. This effect will bias any attempt to use standard hedonics, since higher amenity values also increase the neighborhood human capital composition which will have an independent effect on average prices.

This comparative static suggests that an abundance of low human capital neighborhoods that have high innate amenity levels will magnify the price impact of any upward shift in city demands. This is unlikely to matter in areas where human capital is already quite high, but in areas where there is substantially more mixing, the composition shift created by the demand shift is likely to be more significant.

The impact of an increase in wages for the high skilled in the city can be even more extreme. In this case, the rental cost increase due to an increase in y_H equals

$$(4) \left(1-c_1(y_H,d_k)\right)\left(\int \theta_{d_k}(a) f_{d_k}(a)da + \frac{\gamma_H(\varphi(1)-\varphi(0))}{\gamma_H-\gamma_L} f_{d_k}\left(\overline{a}(d_k)\right)(1-\theta_{d_k}(\overline{a}(d_k))\right)$$

Again, there is overall increase in willingness to pay that is concentrated in the initially wealthier districts. However, there is also a secondary effect coming from the impact that higher skilled wages have on the skill composition of the neighborhoods. Again, the preponderance of less skilled neighborhoods at relatively high amenity levels is critical.

The Dynamic Model

We now move to a dynamic setting where we can consider the impact of an optimistic belief about price growth in different parts of the city. We assume that individuals are perfectly mobile, and continue to maximize the same one period welfare condition, but the cost of housing is now determined by borrowing costs and housing prices. Specifically, the interest rate is denoted ρ, and the one period expected cost of housing services, denoted *r* above, is equal to $\rho p(t) - E(\dot{p}(t))$, where $p(t)$ is the price at time *t* and $E(\dot{p}(t))$, is the expected price at time *t+1*.

Since our focus is on the heterogeneity within the city, we will assume that there are exogenous trends in the key variables, and allow for the possibility that consumers significantly overestimate those trends. The first variable that we allow to change over time is the city level amenity A_{city}, which we now denote A_t, and we assume that this equals $A_t = e^{g_A t} A_0$.

Second, we allow the income of the wealthy, now denoted $y_{H,t}$, to trend upward so that it equals $y_{H,t} = e^{g_Y t} y_{H,0}$. The assumption that the wealthy residents' income grows, but the income of the poor does not, is motivated by the general trend towards rising inequality across U.S. households (e.g. Goldin and Katz, 2008), but notably we are not assuming that the reservation utility of high human capital individuals is changing. As such, the trend is best understood as capturing the particular increase in the wages of skilled people in some, especially high skilled cities, such as Boston and San Francisco (Glaeser and Saiz, 2004).

We also assume that $c(y_H, d_k) = k_H(d_k) + c_H(d_k) y_H d_k$, and $c(y_L, d_k) = k_L(d_k) + c_L(d_k) y_H d_k$where $k_H(d_k)$ and $k_L(d_k)$ represents the fixed cost of the transportation technologies used by rich and poor at this particularly distance and $c_H(d_k)$ represents the variable costs, which scale up with distance and income. This functional form makes it possible for rich and poor to use different transportation technologies at different distances, but it does not allow the rich to change their transportation technology as they grow wealthier. This assumption is made to improve tractability. We assume that $c_i(d_k) + c_i{}'(d_k) d_k > 0$, so that transport costs always rise with distance.

If amenities and income are to growth at rates $\hat{g}_A$ and $\hat{g}_Y$ respectively, then high human capital neighborhoods will always stay high human capital and the pricing equation will satisfy:

$$(5)\ y_{H,t}(1 - c_H(d_k)d_k) - \underline{U}_H - k_H(d_k) + \gamma_H\big(A_t + \ a_k - \varphi(0)\big) = \rho p(t) - E(\dot{p}(t)).$$

Assuming that individuals believe that they know the growth rates of income and amenities with certainly and imposing the transversality condition implies that

$$(6)\ \frac{y_{H,t}(1 - c_H(d_k)d_k)}{\rho - \hat{g}_Y} + \frac{\gamma_H A_t}{\rho - \hat{g}_A} + \frac{1}{\rho}\big(\gamma_H\big(\ a_k - \varphi(0)\big) - \underline{U}_H - k_H(d_k)\big) = p(t)$$

We will treat a bubble as a change in beliefs about amenity or income growth and the price impact of the bubble is just the increase in prices associated with these distorted beliefs.

Proposition 2: Greater optimism about the growth of either amenities or income for high human capital individuals will increase the logarithm of price and both effects will get stronger when the interest rate is lower, as long as expected growth rates are sufficiently close to zero. Optimism about amenity growth has a larger impact on the logarithm of prices if γ_H is high, and if A_t is large, as long as $(t) > \frac{\gamma_H A_t}{\rho - \hat{g}_A}$. Optimism about the growth of income of high human capital individuals will have a larger impact on the logarithm of prices if $y_{H,t}$ is larger or d_k is smaller as long as $p(t) > \frac{1}{\rho - \hat{g}_Y}(1 - c_H(d_k)d_k)y_{H,t}$.

The proposition delivers the unsurprising result that an increase in optimism about the growth of either income or amenities will increase the logarithm of housing prices. Lower interest rates will tend to make the impact of optimism stronger, which suggests a complementarity between easy credit and optimistic beliefs that is similar to the more rational effect emphasized by Himmelberg, Mayer and Sinai (2005). The impact of amenities is higher when the taste for amenities is stronger, or when the initial amenity level is higher, as long as the discounted value of the amenity flow is less than the housing price itself. The impact of higher expected income growth on the logarithm of prices will be greater if housing prices are higher than the expected flow of future earnings (net of commuting costs). If that condition does not hold, higher income levels push up the level of prices more than the derivative of the level with respect to income growth, and this means that the derivative of the logarithm of prices is lower.

The same condition ensures that there will be greater price growth in the center of the city, because less income is lost in commutes.

For neighborhoods that begin as low human capital areas, the price impact of growth and of a bubble can be more impressive. We consider first the time path of prices given correct expectations about prices and then discuss the impact of incorrect price expectations. For simplicity, we just treat the case of rising amenity levels, although the results with rising income levels are quite similar. We assume that the low human capital neighborhood does not tip until the point where the low human capital equilibrium is no longer sustainable.

Proposition 3: If city-wide amenities are growing at a continuous rate g_A, then at some point, low human capital communities will flip to high human capital communities. Before that point, low human capital communities have lower price levels than high human capital communities with equal amenity levels, so there is a discrete decline in price growth at the tipping point. Low human capital communities experience faster price growth near the tipping point than equal amenity high human capital communities. The price growth of low human capital communities increases over time and rate of price growth ($\frac{\dot{p}(t)}{p(t)}$) also increases over time as long as $(1 - cLdkdk - kLdk + \gamma Lak - \varphi 1 > UL$.

This third proposition describes the rational expectations equilibrium of the model with flipping. In neighborhoods with moderate amenity levels, there will be both low and high human capital communities. As city-wide amenities or wages for high human capital individuals increase, the low human capital communities will eventually flip to high human capital communities. Since this transition is correctly anticipated, prices will adjust and there will be no discrete jumps in prices even at the tipping point. However, the rate of price growth will slow down. The decline in housing price appreciation is an offset for the increased quality of the community at that point.

Typically, the price growth of the low human capital community will be faster than comparable high human capital communities, because of the anticipated flipping, even though price growth would actually be lower for the low human capital communities if owners didn't expect neighborhood compositional change.

Finally, we turn to the impact of a bubble. As before, we model a bubble as an increase in the expected growth rate of either the city-wide amenity level or the wages of high skilled workers. Individuals hold their beliefs with certainty. We do not address the dynamics of this belief, but just ask what impact this belief change will have on the price. Of course, the impact of the belief on the price is identical to the impact of an actual shift in the growth rate. We consider the case where there is a trend in amenities and a trend in high human capital wages separately, so when we consider the impact of mistaken beliefs about one of the variable's growth rates, we assume that the other variable is time invariant.

Proposition 4: If $\hat{g}_Y = 0$, then an increase in $\hat{g}_A$ will always increase prices in the low human capital community and the effect of this optimism on price will always be stronger if ρ is smaller. If $2\frac{\varphi(1)}{A_{t*}} > Log\left(\frac{\varphi(1)}{A_t}\right)$, then $\hat{g}_A$ will raise prices more for the low human capital community than for the high human capital community as long as $\rho/\hat{g}_Y$ is sufficiently close to one.

If $\hat{g}_A = 0$, then an increase in $\hat{g}_Y$ will always increase prices in the low human capital community and the effect of this optimism on price will always be stronger if ρ is smaller. If $2\left(\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t*}(1-c_H(d_k)d_k)}\right) > Log\left(\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t}(1-c_H(d_k)d_k)}\right)$, then $\hat{g}_Y$ will raise prices more for the low human capital community than for the high human capital community as long as $\rho/\hat{g}_Y$ is sufficiently close to one.

Proposition 4 notes that optimism will also increase prices in the low human capital community, even if that optimism only relates to the income of high human capital individuals. The spillover between high human capital incomes and low human capital area prices occurs because low human capital areas are expected to someday switch and become high human capital areas.

As before, low interest rates will tend to make the impact of the optimism more pronounced. Again, as in Himmelberg, Mayer and Sinai (2005), this effect occurs because lower interest rates mean that future price appreciation becomes more valuable.

Perhaps the most surprising result is that the impact of optimism, even about earnings of high human capital individuals, can be higher in low human capital areas than in high human capital areas. There are two critical ingredients for this unusual result. The difference between the interest rate and the expected growth rate must be sufficiently small, and the expected time before the flipping point must also be modest. The extra price appreciation for the low human capital communities reflects the expected shift to high human capital areas, and this is more valuable when the future isn't discounted too highly (at least relative to the growth rate) and when the expected time to flipping is small.

This model focused on gentrification, but there are other reasons why cities with more centralized poverty could have experienced faster price growth at the city center. One possible explanation is that areas with centralized poverty could have less elastic housing supply. Central city poverty could mean that the area is older and more regulated. Alternatively, in poor central cities prices could be too low to justify serious redevelopment, even after the boom. Less supply response could mean more price response.

Alternatively, areas with centralized poverty might have centralized urban assets that attract the poor. One such asset might be a public transit system that provides mobility for the poor (Glaeser, Kahn and Rappaport, 2008). Alternatively, the poor might locate in the central city to be close to jobs, and it may be that areas with more centralized poverty have more centralized employment. While the model does not precisely speak to these issues, they can be thought of as higher local amenity levels, which can be associated with faster price growth during a boom.

IV. Examining the Hypotheses

We first focus on the gentrification hypothesis that is the core idea expounded in the model; we then turn to our other hypotheses about the link between decentralized income and centralized price growth.

To examine the gentrification hypothesis, in Table 3 we focus only on those zip codes that have incomes below the metropolitan area median. This enables us to focus on those areas that could conceivably switch from less skilled to more skilled inhabitants, as in the model.

Regressions (1) and (3) essentially reproduce our core specifications (regressions (2) and (4) in Table 2) for this subset of zip codes. As before, we find that centralized areas had faster price growth. We also find that centralized price growth is stronger in those metropolitan areas where income is more decentralized.

Regressions (2) and (4) add in three controls for the potential for gentrification. We follow Guerrieri, Hartley, and Hurst (2011) and control for distance to nearest high-income zip code (defined as those in the top quartile of the MSA's zip code income distribution). We also interact this variable with a dummy variable indicating that the zip code is in the second quartile of the metropolitan area income distribution. These areas, with relatively higher incomes within this sample, may have more potentially for gentrification.

We also include a control for the heterogeneity of incomes within the metropolitan area. At the metropolitan area level, we estimate the average square of the difference between zip code income and average income in zip codes within one mile. This measure will be low when zip codes typically have neighbors with similar incomes. The measure will be high when wealthy areas abut poor areas.

Regressions (3) and (4) show that controlling for these measures causes the interaction between distance to central business district and income decentralization to drop by about 21 percent in the later period. The new controls have no effect on this interaction during the earlier growth period. We estimate a statistically significant effect both for distance to high income zip code during the earlier boom but not during the later boom. The more important effect is that the income mixing variable seems to flatten the price growth distance gradient. In places that have more mixing of incomes throughout the metropolitan area, the tendency of price growth to occur more in the city center is attenuated and that causes the core interaction to become less powerful. Overall, these results suggest that gentrification may explain some, but probably not all of the core interaction.

Tables 4 and 5 show the results for our other controls during the two time periods. The first regressions in each table reproduce regressions (2) and (4) in Table 2. Our new measures are unfortunately only available for a subset of the metropolitan areas. To avoid losing sample size, we impute the sample median variable of the missing variable for the relevant interaction with distance to the central business district. We then also include a dummy variable that equals one when the relevant variable is missing and interact that variable with distance to the central business district.

Regression (2) in each table includes our measures of housing supply elasticity. We include the Saiz (2010) measure of supply elasticity at the metropolitan area level and interact it with distance to the central business district. This measure is meant to capture the possibility that less restrictive areas might find it easier to build in the city center. We also include a zip code-level measure of the share of the housing stock that was built in the 1990s and we interact that variable with distance to the central business district. This second measure is, of course, an endogenous outcome variable, and omitted demand side factors should lead to a positive relationship between this variable and price growth that complicates interpretation.

The zip code measure of new housing built is negatively associated with price growth in both tables. There was less price growth in areas where more housing was built. Given that the endogeneity of this variable should generate a positive coefficient, the strong negative coefficients reinforce the powerful role that supply had in dampening price growth within the boom. During the 1980s, this supply measure is less negatively associated with price growth in areas that are further from the city center. The interaction between area level supply elasticity and distance to the central business district has no significant effect in either decade. Including these supply measures lead the interaction between centralized poverty and distance to the central business district to drop by about 20 percent in the 1990s.

In the third regression of each table, we include measures of public transit access. We include a zip code measure of the share of adults taking public transit. We also include the interaction of a metropolitan area measure—the share of people living more than five miles from the city center who take public transit to work minus the share of people taking public transit within five miles of the city center—and the distance to the central business district. This metropolitan area-level measure should capture the extent to which the inner city has a comparative advantage in access to public transit.

During the 1980s, the share taking public transit to work has a positive association with price growth. During the 1990s, the interaction between decentralized transit use and distance to the central business district is weakly positive, meaning that price growth was faster on the edge when there was relatively more transit use on the edge. While the coefficient is modest, it has a large effect on the estimated interaction between centralized poverty and centralized price growth, causing it to drop by almost 40 percent.

We do not exactly why know why there was a link between centralized transit systems and centralized price growth, but there are several plausible interpretations of this effect. If the boom represented a temporary over-estimation of the valuation of some urban assets, like January temperature, then it is not unreasonable that buyers during the boom estimated that the value of public transit access is large. It is also possible that these cities that are highly dependent on public transit, like New York, are also places where suburban access is more difficult. Since more widely available suburban space is a poorer substitute for central locations in these cities, it is reasonable that central city land would have been seen as scarcer and more valuable.

The fourth regression in each table controls for the centralization of employment within the metropolitan area. We use the measure used by Glaeser and Kahn (2001) of the share of employment within three miles of the central business district (updated to more recent data by Kneebone [2009]). During both time periods, more centralized employment was associated with a flatter price growth gradient, which substantially diminishes the core interaction during the 1990s.

The fifth regression in each table includes all of these controls. Taken together, our hypotheses explain almost half of the centralized poverty-centralized price growth effect during the 1990s and almost nothing during the 1980s. Unfortunately, when we control for everything the controls become so imprecisely measured that it is essentially impossible to assess their relative importance with much precision.

IV. Conclusion

The two national housing booms of the past three decades did not impact everywhere equally. Metropolitan areas that had initially higher prices experienced faster price growth during both the 1980s boom and during the more recent price episode. January temperature was very strongly associated with price growth between 1996 and 2006, but not during the earlier time period.

There were also striking differences in price growth within metropolitan areas. In both booms, price growth was faster in city centers than on the periphery of urban areas. The tendency of price growth to occur in more centralized locations was more pronounced within those metropolitan areas where richer people were more likely to live farther away from the urban core.

We explore several explanations for this phenomenon. We find some evidence supporting the view that poorer inner city areas experienced faster price growth because these were natural places for changes in neighborhood composition and gentrification. However, this phenomenon does not explain why price growth was stronger in poorer central cities. We find support for the view that price growth was faster in areas where supply was more constrained, but we find little support for the view that inelastic housing supply explains the connection between centralized poverty and centralized price growth.

We find modest evidence supporting the view that urban form underpins the phenomenon. Place with more centralized poverty seem to have more centralized public transit systems. In these areas, price growth was also faster in the urban core. We also find that places with more centralized employment, had marginally more price growth on the urban edge.

Yet taken as a whole, these forces can explain less than half of our core puzzle during the 1996-2006 period and far less during the 1980s. As such, the core puzzle remains fully explained. We believe that this merits more investigation, as part of a broader research agenda investigating neighborhood and metropolitan area differences in price growth during booms and busts.

Appendix: Proofs of Propositions

Proof of Proposition 1:

The willingness of type *i* to pay for neighborhood *k* equals

$$y_i - c(y_i, d_k) + \gamma_i\left(A_{city} + a_k - \varphi(\pi_k)\right) - \underline{U_i}.$$

The willingness to pay differential between high and low human capital individuals equals

$$\left(y_H - c(y_H, d_k) - \underline{U_H}\right) - \left(y_L - c(y_L, d_k) - \underline{U_L}\right) + (\gamma_H - \gamma_L)\left(A_{city} + a_k - \varphi(\pi_k)\right).$$

This must equal zero in any mixed neighborhood, but any increase in π_k from that point will cause the difference to become negative as $-\varphi'(\pi_k)\,(\gamma_H - \gamma_L) < 0$.

For any low human capital neighborhood, $y_L - r_k - c(y_L, d_k) + \gamma_L\left(A_{city} + a_k - \varphi(1)\right) = \underline{U_L}$ and $y_H - r_k - c(y_H, d_k) + \gamma_H\left(A_{city} + a_k - \varphi(1)\right) < \underline{U_H}$. Using the equality to substitute in for price in the inequality yields $a_k < \frac{\left(\underline{U_H} - y_H + c(y_H, d_k)\right) - \left(\underline{U_L} - y_L + c(y_L, d_k)\right)}{\gamma_H - \gamma_L} - A_{city} + \varphi(1)$, so $\frac{\left(\underline{U_H} - y_H + c(y_H, d_k)\right) - \left(\underline{U_L} - y_L + c(y_L, d_k)\right)}{\gamma_H - \gamma_L} - A_{city} + \varphi(1)$ defines $\overline{a}(d_k)$. Differentiation gives us that $\overline{a}(d_k)$ is increasing with d_k if and only if $\frac{\partial c(y_H, d_k)}{\partial d_k} > \frac{\partial c(y_L, d_k)}{\partial d_k}$.

For any high human capital neighborhood, $y_H - r_k - c(y_H, d_k) + \gamma_H\big(A_{city} + a_k -$ *φ(0)=UH* and *yL−rk−cyL,dk+γLAcity+ ak−φ(0)<UL.* Combining these conditions gives us that $a_k > \frac{\left(\underline{U_H} - y_H + c(y_H, d_k)\right) - \left(\underline{U_L} - y_L + c(y_L, d_k)\right)}{\gamma_H - \gamma_L} - A_{city} + \varphi(0)$, and that defines $\underline{a}(d_k)$which equals $\overline{a}(d_k) - (\varphi(1) - \varphi(0))$. The comparative static on $\underline{a}(d_k)$ follows.

Proof of Proposition 2:

The derivative of the log price with respect to $\hat{g}_A$ is $\frac{1}{p(t)(\rho - \hat{g}_A)^2}\gamma_H A_t$ or

$$\frac{\gamma_H A_t}{\left(\frac{(\rho - \hat{g}_A)^2}{\rho}\left(\gamma_H(a_k - \varphi(0)) - \underline{U_H} - k_H(d_k)\right) + (\rho - \hat{g}_A)\gamma_H A_t + \frac{(\rho - \hat{g}_A)^2 y_{H,t}(1 - c_H(d_k) d_k)}{\rho - \hat{g}_Y}.\right)}$$

which is positive. The derivative of this derivative with respect to ρ is

$$\frac{-\gamma_H A_t\left(\frac{(\rho^2-\hat{g}_A^2)}{\rho^2}(\gamma_H(a_k-\varphi(0))-\underline{U}_H-k_H(d_k))+\gamma_H A_t+\frac{(\rho-\hat{g}_A)(\rho-2\hat{g}_Y+\hat{g}_A)y_{H,t}(1-c_H(d_k)d_k)}{(\rho-\hat{g}_Y)^2}.\right)}{\left(\frac{(\rho-\hat{g}_A)^2}{\rho}(\gamma_H(a_k-\varphi(0))-\underline{U}_H-k_H(d_k))+(\rho-\hat{g}_A)\gamma_H A_t+\frac{(\rho-\hat{g}_A)^2 y_{H,t}(1-c_H(d_k)d_k)}{\rho-\hat{g}_Y}.\right)^2}.$$

The denominator of this is positive and the numerator can be written as

$$\frac{(\rho^2-\hat{g}_A^2)}{\rho}p(t)-\frac{\hat{g}_A}{\rho}\gamma_H A_t-\frac{(\rho-\hat{g}_A)^2\hat{g}_Y}{\rho(\rho-\hat{g}_Y)^2}(1-c_H(d_k)d_k)y_{H,t}$$

Which is strictly positive when $\hat{g}_A=\hat{g}_Y=0$ and which will therefore remain positive as long as the growth rates are small. The derivative of this derivative with respect to A_t is

$$\frac{(\rho-\hat{g}_A)^2\gamma_H\left(p(t)-\frac{\gamma_H A_t}{\rho-\hat{g}_A}\right)}{\left(\frac{(\rho-\hat{g}_A)^2}{\rho}\left(\gamma_H(a_k-\varphi(0))-\underline{U}_H-k_H(d_k)\right)+(\rho-\hat{g}_A)\gamma_H A_t+\frac{(\rho-\hat{g}_A)^2 y_{H,t}(1-c_H(d_k)d_k)}{\rho-\hat{g}_Y}.\right)^2}$$

which is positive if $p(t)>\frac{\gamma_H A_t}{\rho-\hat{g}_A}$. The derivative of the derivative with respect to γ_H is

$$\frac{\frac{(\rho-\hat{g}_A)^2 y_{H,t}(1-c_H(d_k)d_k)}{\rho-\hat{g}_Y}A_t}{\left(\frac{(\rho-\hat{g}_A)^2}{r}\left(\gamma_H(a_k-\varphi(0))-\underline{U}_H-k_H(d_k)\right)+(\rho-\hat{g}_A)\gamma_H A_t+\frac{(r-\hat{g}_A)^2 y_{H,t}(1-c_H(d_k)d_k)}{r-\hat{g}_Y}.\right)^2}$$

which is positive.

The derivative of the log price with respect to $\hat{g}_Y$ is $\frac{1}{p(t)(r-\hat{g}_Y)^2}y_{H,t}(1-c_H(d_k)d_k)$,or

$\frac{y_{H,t}(1-c_H(d_k)d_k)}{\frac{(r-\hat{g}_Y)^2}{r}(\gamma_H(a_k-\varphi(0))-\underline{U}_H-k_H(d_k))+\frac{(r-\hat{g}_Y)^2}{r-\hat{g}_A}\gamma_H A_t+(r-\hat{g}_Y)y_{H,t}(1-c_H(d_k)d_k}$. The derivative of this derivative with respect to r is

$$\frac{-y_{H,t}(1-c_H(d_k)d_k)\left(\frac{(r^2-\hat{g}_Y^2)}{r^2}(\gamma_H(a_k-\varphi(0))-\underline{U}_H-k_H(d_k))+\frac{(r-2\hat{g}_A+\hat{g}_Y)(r-\hat{g}_Y)}{(r-\hat{g}_A)^2}\gamma_H A_t+y_{H,t}(1-c_H(d_k)d_k)\right)}{\left(\frac{(r-\hat{g}_Y)^2}{r}(\gamma_H(a_k-\varphi(0))-\underline{U}_H-k_H(d_k))+\frac{(r-\hat{g}_Y)^2}{r-\hat{g}_A}\gamma_H A_t+(r-\hat{g}_Y)y_{H,t}(1-c_H(d_k)d_k)\right)^2}$$

The numerator can be rewritten as $-y_{H,t}(1-c_H(d_k)d_k)$ times

$$\frac{(r^2-\hat{g}_Y^2)}{r}p(t)-\frac{\hat{g}_Y}{r}y_{H,t}(1-c_H(d_k)d_k)-\frac{\hat{g}_A(r-\hat{g}_Y)^2}{r(r-\hat{g}_A)^2}\gamma_H A_t$$

which is positive as long as the growth rate of *Y* and *A* are sufficiently close to zero. The derivative of the derivative with respect to $y_{H,t}$

$$\frac{(1-c_H(d_k)d_k)(r-\hat{g}_Y)^2\left(p(t)-\frac{1}{r-\hat{g}_Y}(1-c_H(d_k)d_k)y_{H,t}\right)}{\left(\frac{(r-\hat{g}_Y)^2}{r}(\gamma_H(\,a_k-\varphi(0))-\underline{U}_H-k_H(d_k))+\frac{(r-\hat{g}_Y)^2}{r-\hat{g}_A}\gamma_H A_t+(r-\hat{g}_Y)y_{H,t}(1-c_H(d_k)d_k)\right)^2},$$

Which is positive if and only if $p(t) > \frac{(1-c_H(d_k)d_k)y_{H,t}}{\rho-\hat{g}_Y}$. Finally, the derivative with respect to d_k equals

$$\frac{(-c_H(d_k)-c_H{'}(d_k)d_k)y_{H,t}(r-\hat{g}_Y)^2\left(p(t)-\frac{1}{r-\hat{g}_Y}(1-c_H(d_k)d_k)y_{H,t}\right)}{\left(\frac{(r-\hat{g}_Y)^2}{r}(\gamma_H(\,a_k-\varphi(0))-\underline{U}_H-k_H(d_k))+\frac{(r-\hat{g}_Y)^2}{r-\hat{g}_A}\gamma_H A_t+(r-\hat{g}_Y)y_{H,t}(1-c_H(d_k)d_k)\right)^2},$$ which is negative as long as $p(t) > \frac{1}{r-\hat{g}_Y}(1-c_H(d_k)d_k)y_{H,t}$.

Proof of Proposition 3:

We let Δ_H denote $\gamma_H\big(\,a_k-\varphi(0)\big)-\underline{U}_H-k_H(d_k)$,the invariant elements in the high human capital individual's valuation of a neighborhood, so prices in the high human capital community, denoted $p_H(t)$ can be written as $\frac{y_{H,t}(1-c_H(d_k)d_k)}{\rho-\hat{g}_Y}+\frac{\gamma_H A_t}{\rho-\hat{g}_A}+\frac{\Delta_H}{\rho}$. We let Δ_L denote $(1-c_L(d_k)d_k)y_L-\underline{U}_L-k_L(d_k)+\gamma_L\big(a_k-\varphi(1)\big)$ the time invariant elements of the low human capital individual's valuation of the neighborhood. If there was no possibility of flipping from low to high human capital communities, the price of housing in the low human capital community must equal $\frac{\gamma_L A_t}{\rho-\hat{g}_A}+\frac{\Delta_L}{\rho}$.

Following the proof of proposition 1, a low human capital community can only persist if

$(\gamma_H-\gamma_L)A_t+y_{H,t}(1-c_H(d_k)d_k<\Delta_L-\Delta_H+(\gamma_H-\gamma_L)\varphi(1)$. This means that there must exist some time period, t^*, at which every low human capital community flips to becoming a high human capital community.

At t^*, the price will equal $\frac{y_{H,t}(1-c_H(d_k)d_k)}{\rho-\hat{g}_Y}+\frac{\gamma_H A_t}{\rho-\hat{g}_A}+\frac{\Delta_H}{\rho}$. Before that point, the area will be occupied only by low human capital individuals and so the housing price equation must satisfy:

(A1) $\gamma_L A_t+\Delta_L=\rho p(t)-\dot{p}(t)$.

The price that satisfies (A1) and that equals $p_H(t\,*)$ at $t=t^*$, is $\frac{\gamma_L A_t}{\rho-g_A}+\frac{\Delta_L}{\rho}+e^{\rho(t-t*)}(p_H(t^*)-p_L(t^*))$, where the housing costs for paying for the interest on extra term $p_H(t^*)-p_L(t^*))$, are just offset by the capital gains associated with this portion of the price, which is rising just as fast as the interest rate. Using the condition for t^*, the value of $p_H(t^*)-p_L(t^*)$ equals $\frac{g_Y y_{H,t*}(1-c_H(d_k)d_k)}{(\rho-g_Y)\rho}+\frac{g_A A_{t*}(\gamma_H-\gamma_L)}{(\rho-g_A)\rho}+\frac{(\gamma_H-\gamma_L)\varphi(1)}{\rho}$ or $\frac{y_{H,t*}(1-c_H(d_k)d_k)}{\rho-g_Y}+\frac{(\gamma_H-\gamma_L)A_{t*}}{\rho-g_A}+\frac{\Delta_H-\Delta_L}{\rho}$. The

The gap between the price in a high human capital community and the price in a low human capital community with equal amenities equals:

$$\frac{y_{H,t}(1-c_H(d_k)d_k)}{\rho-\hat{g}_Y}+\frac{(\gamma_H-\gamma_L)A_t}{\rho-\hat{g}_A}+\frac{\Delta_H-\Delta_L}{\rho} - e^{\rho(t-t*)}\left(\frac{y_{H,t*}(1-c_H(d_k)d_k)}{\rho-\hat{g}_Y}+\frac{(\gamma_H-\gamma_L)A_{t*}}{\rho-\hat{g}_A}+\frac{\Delta_H-\Delta_L}{\rho}\right),$$

which is strictly positive since $\rho > g_Y$

and $\rho > g_A$. Price growth in the low human capital community equals $\frac{g_A\gamma_L A_t}{\rho-g_A}+e^{\rho(t-t*)}\left(\frac{g_Y y_{H,t*}(1-c_H(d_k)d_k)}{(\rho-g_Y)}+\frac{g_A A_{t*}(\gamma_H-\gamma_L)}{(\rho-g_A)}+(\gamma_H-\gamma_L)\varphi(1)\right)$ and that is increasing with *t*. Price growth is a high human capital community is $\frac{g_Y y_{H,t}(1-c_H(d_k)d_k)}{\rho-g_Y}+\frac{g_A\gamma_H A_t}{\rho-g_A}$. At *t*=*t**, price growth is strictly higher for the low human capital community so there must be a discrete jump downward in the growth of prices when the community flips.

The growth rate ($\frac{\dot{p}(t)}{p(t)}$) in low human capital communities equals $\frac{g_A\frac{\gamma_L A_t}{\rho-\hat{g}_A}+\rho e^{\rho(t-t*)}(p_H(t^*)-p_L(t^*))}{p_L(t)+e^{\rho(t-t*)}(p_H(t^*)-p_L(t^*))}$ and the derivative of this with respect to *t* equals: $\frac{\frac{g_A^2\gamma_L A_t\Delta_L}{(\rho-g_A)\rho}+(\rho\Delta_L+(\rho-g_A)\gamma_L A_t)e^{\rho(t-t*)}(p_H(t^*)-p_L(t^*))}{\left(p_L(t)+e^{\rho(t-t*)}(p_H(t^*)-p_L(t^*))\right)^2}$, which is positive as long as $\Delta_L > 0$.

Proof of Proposition 4:

The impact of a change in the anticipated growth rate for either amenities or high wage human capital on prices in high human capital communities has already been worked out in the proof of Proposition 2.

The impact of a shift in the growth rate on prices in low human capital communities is somewhat more complex, because the anticipated tipping time also changes with the growth rate. Let *J* denote the expected number of periods before flipping, so as of time *t*, *J* solves: $(\gamma_H - \gamma L A t e g A J + y H, t e g Y J(1-c H d k d k) = \Delta L - \Delta H + \gamma H - \gamma L \varphi(1)$. If $gY=0,$ then the derivative of *J* with respect to $\hat{g}_A$ equals $\frac{-J}{\hat{g}_A}$ and if $\hat{g}_A = 0$ then the derivative of *J* with respect to $\hat{g}_Y is\ \frac{-J}{\hat{g}_Y}$. In both cases, there is no change in the value of A_{t+J} or $y_{H,t+J}$ at the tipping point.

Given these assumptions, the price in the low human capital community will equal $\frac{\gamma_L A_t}{\rho-\hat{g}_A}+\frac{\Delta_L}{\rho}+e^{-\rho J}\left(\frac{(\gamma_H-\gamma_L)A_{t*}}{\rho-\hat{g}_A}+\frac{y_{H,t*}(1-c_H(d_k)d_k)}{\rho-\hat{g}_Y}+\frac{\Delta_H-\Delta_L}{\rho}\right)$. The derivative of prices in the low human capital community with respect to $\hat{g}_A$ (holding $\hat{g}_Y = 0$) equals

$\frac{\gamma_L A_t}{(\rho-\hat{g}_A)^2}+\frac{\rho J}{\hat{g}_A}e^{-\rho J}\left(\frac{(\gamma_H-\gamma_L)A_{t*}}{\rho-\hat{g}_A}+\frac{\Delta_{+H}-\Delta_L}{\rho}\right)+e^{-\rho J}\frac{(\gamma_H-\gamma_L)A_{t*}}{(\rho-\hat{g}_A)^2}$, which is strictly positive and where $\Delta_{+H}=\Delta_H+y_{H,t}(1-c_H(d_k)d_k)$.

The derivative of this derivative with respect to ρ equals $-\frac{2\gamma_L A_t}{(\rho-\hat{g}_A)^2}-\frac{\rho J^2}{\hat{g}_A}e^{-\rho J}\left(\frac{(\gamma_H-\gamma_L)A_{t*}}{\rho-\hat{g}_A}+\frac{\Delta_{+H}-\Delta_L}{\rho}\right)-(J+\rho)e^{-\rho J}\frac{(\gamma_H-\gamma_L)A_{t*}}{(\rho-\hat{g}_A)^2}-2e^{-\rho J}\frac{(\gamma_H-\gamma_L)A_{t*}}{(\rho-\hat{g}_A)^3}$, which is strictly negative. The impact of the optimistic expectations becomes weaker as the interest rate rises.

The derivative of prices in the low human capital community with respect to $\hat{g}_Y$ (holding $\hat{g}_A=0$) equals $\frac{\rho J}{\hat{g}_Y}e^{-\rho J}\left(\frac{y_{H,t*}(1-c_H(d_k)d_k)}{\rho-\hat{g}_Y}+\frac{\Delta_{++H}-\Delta_{++L}}{\rho}\right)+e^{-\rho J}\frac{y_{H,t*}(1-c_H(d_k)d_k)}{(\rho-\hat{g}_Y)^2}$ which is also strictly positive where $\Delta_{++H}=\Delta_H+\gamma_H A$ and $\Delta_{++L}=\Delta_L+\gamma_L A$. The derivative of this derivative with respect to ρ is $\frac{-\rho J^2}{\hat{g}_Y}e^{-\rho J}\left(\frac{y_{H,t*}(1-c_H(d_k)d_k)}{\rho-\hat{g}_Y}+\frac{\Delta_{++H}-\Delta_{++L}}{\rho}\right)-Je^{-\rho J}\frac{y_{H,t*}(1-c_H(d_k)d_k)}{(\rho-\hat{g}_Y)^2}-Je^{-\rho J}\frac{y_{H,t*}(1-c_H(d_k)d_k)}{(\rho-\hat{g}_Y)^2}-2e^{-\rho J}\frac{y_{H,t*}(1-c_H(d_k)d_k)}{(\rho-\hat{g}_Y)^3}$ which is also negative.

By comparison, the derivative of price in a high human capital community, with respect to $\hat{g}_A$ holding $\hat{g}_Y=0$, equals $\frac{\gamma_H A_t}{(\rho-g_A)^2}$ and the derivative of price in a high human capital community with respect to $\hat{g}_Y$ holding $\hat{g}_A=0$ is $\frac{y_{H,t}(1-c_H(d_k)d_k}{(\rho-\hat{g}_Y)^2}$.

When $\hat{g}_Y=0$, the derivative of the low human capital community price will be higher than the derivative of the high human capital price with respect to $\hat{g}_A$ if $q\left(z-1+(z-1)^2\frac{\varphi(1)}{A_{t*}}\right)+1-e^{(z-1)q}$ is positive where $\hat{g}_A J=q$, which must be constant with changes in either ρ or $\hat{g}_A$, and *z* denotes $\frac{\rho}{\hat{g}_A}$. The derivative of that difference with respect to *z* is $q(1+2(z-1)\frac{\varphi(1)}{A_{t*}}-e^{(z-1)q})$ and the derivative of that with respect to *z* is $q(2\frac{\varphi(1)}{A_{t*}}-qe^{(z-1)q})$. As long as $q<2\left(\frac{\varphi(1)}{A_{t*}}\right)$, there will be some values of *z*, near one, for which $q(1+2(z-1)\frac{\varphi(1)}{A_{t*}}-e^{(z-1)q})$ is positive and in that range $q(1+2(z-1)\frac{\varphi(1)}{A_{t*}}-e^{(z-1)q})$ and $qq\left(z-1+(z-1)^2\frac{\varphi(1)}{A_{t*}}\right)+1-e^{(z-1)q}$ are both positive. Hence, the impact of optimism on the prices must be stronger for low human capital communities than for high human capital communities. The condition that $q<2$ is that $2\frac{\varphi(1)}{A_{t*}}>Log\left(\frac{\varphi(1)}{A_t}\right)$.

When $\hat{g}_A=0$, the derivative of the low human capital community price will be higher than the derivative of the high human capital price with respect to $\hat{g}_Y$ if $q\left(z-1+(z-1)^2\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t}*(1-c_H(d_k)d_k)}\right)+1-e^{(z-1)q}$ is positive where $\hat{g}_Y J=q$, which must be constant with changes in either ρ or $\hat{g}_Y$, and *z* denotes $\frac{\rho}{\hat{g}_Y}$. The derivative of that difference with respect to *z* is $q(1+2(z-1)\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t*}(1-c_H(d_k)d_k)}-e^{(z-1)q})$ and the derivative of that with respect to *z* is $q(2\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t*}(1-c_H(d_k)d_k)}-qe^{(z-1)q})$. As long as $q < 2\left(\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t*}(1-c_H(d_k)d_k)}\right)$, there will be some values of *z*, near one, for which $q(1+2(z-1)\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t^*}(1-c_H(d_k)d_k)}-e^{(z-1)q})$ is positive and in that range $q(1+2(z-1)\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t^*}(1-c_H(d_k)d_k)}-e^{(z-1)q})$ and $qq\left(z-1+(z-1)^2\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t*}(1-c_H(d_k)d_k)}\right)+1-e^{(z-1)q}$ are both positive. Hence, the impact of optimism on the prices must be stronger for low human capital communities than for high human capital communities. The condition that $q < 2\left(\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t*}(1-c_H(d_k)d_k)}\right)$ is that $2\left(\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t*}(1-c_H(d_k)d_k)}\right) > Log\left(\frac{(\gamma_H-\gamma_L)\varphi(1)}{y_{H,t}(1-c_H(d_k)d_k)}\right)$.

References

Glaeser, Edward L., Joseph Gyourko, and Albert Saiz (2008), "Housing Supply and Housing Bubbles." *Journal of Urban Economics* 64(2):198-217, September.

Glaeser, Edward L. and Matthew E. Kahn (2001), "Decentralized Employment and the Transformation of the American City." *Brookings/Wharton Papers on Urban Affairs* 2:1-63.

Glaeser, Edward L., Matthew E. Kahn, and Jordan Rappaport (2008), "Why Do the Poor Live in Cities? The Role of Public Transportation." *Journal of Urban Economics* 63(1):1-24, January.

Glaeser, Edward L. and Albert Saiz (2004), "The Rise of the Skilled City." *Brookings-Wharton Papers on Urban Affairs* 5:47-105.

Glaeser, Edward L. and Kristina Tobio (2008), "The Rise of the Sunbelt." *Southern Economic Journal* 74(3): 610-643, January.

Goldin, Claudia and Lawrence F. Katz (2008). *The Race Between Education and Technology.* Cambridge, Mass.: Harvard University Press.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst (2011), "Endogenous Gentrification and Housing Price Dynamics." University of Chicago, mimeo. Available online at http://faculty.chicagobooth.edu/erik.hurst/research/gentrification_final_august2011_submit.pdf.

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

Kneebone, Elizabeth (2009), "Job Sprawl Revisited: The Changing Geography of Metropolitan Employment." Brookings Institution Research Report, April. Available online at http://www.brookings.edu/~/media/Files/rc/reports/2009/0406_job_sprawl_kneebone/20090406_jobsprawl_kneebone.pdf.

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

Roback, Jennifer (1982), "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90(6):1257-1278, December.

Saiz, Albert (2010), "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics* 125(3): 1253-1296, August.

Schelling, Thomas C. (1971), "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1(2):143-186.

Table 1
MSA Level Regressions of Housing Price Changes

	(1)	(2)	(3)	(4)
	Price Growth, 1982-89		***Price Growth, 1996-2006***	
January Temperature	-0.00490***	-0.00915***	0.00606***	0.00562***
	(0.00103)	(0.00121)	(0.000693)	(0.000842)
Log Population Density, Previous Census	0.121***	0.161***	-0.0519***	-0.0554***
	(0.0140)	(0.0150)	(0.0103)	(0.0112)
Log Median Housing Value, Previous Census	0.140***	0.434***	0.524***	0.582***
	(0.0486)	(0.0682)	(0.0245)	(0.0363)
Percent with College Degree, Previous Census		-0.539*		-0.897***
		(0.286)		(0.171)
Log Median Income, Previous Census		-0.816***		0.0683
		(0.144)		(0.103)
Constant	-1.552***	3.407***	-5.254***	-6.400***
	(0.515)	(1.148)	(0.255)	(0.832)
Observations	300	300	300	300
R^2	0.270	0.352	0.688	0.717

*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.

Table 2
Zip Code Level Regressions of Housing Price Changes

	(1)	(2)	(3)	(4)
	Price Growth, 1982-1989		***Price Growth, 1996-2006***	
Log Dist CBD (zip code) * Income-Distance Gradient 1990 (MSA) - Demeaned	-0.106***	-0.104***	-0.201***	-0.191***
	(0.0385)	(0.0348)	(0.0455)	(0.0457)
Log Med Inc (zip code) * Log Distance CBD (zip code) - Demeaned	0.000699	0.00491	0.00124	-0.00240
	(0.00826)	(0.00752)	(0.0104)	(0.0107)
Log Distance to CBD (zip code)	-0.0216***	-0.0132**	-0.0400***	-0.0380***
	(0.00507)	(0.00526)	(0.00687)	(0.00804)
Log Median Income, 2000 Census (zip code)	0.00732	0.0777***	-0.102***	-0.136***
	(0.0146)	(0.0227)	(0.0118)	(0.0237)
Percent BA or Higher, 2000 Census (zip code)		-0.0254		0.0839**
		(0.0354)		(0.0352)
Owner-Occupied Single-Family Share of Housing Units, 2000 Census (zip code)		-0.138***		-0.0152
		(0.0273)		(0.0317)
Log Population Density, 2000 Census (zip code)		0.00539*		-0.00417
		(0.00289)		(0.00410)
Constant	0.305*	-0.428*	1.910***	2.286***
	(0.156)	(0.230)	(0.126)	(0.247)
Observations	3,342	3,342	3,342	3,342
R-squared	0.947	0.950	0.937	0.938

*** p<0.01, ** p<0.05, * p<0.1
All regressions include metropolitan area fixed effects. Standard errors are clustered at the metropolitan area level.

Table 3
Zip Code Level Regressions of Housing Price Changes, Zips Below Median Income for MSA Sample

	(1)	(2)	(3)	(4)
	Price Growth 1982-1989		***Price Growth 1996-2006***	
Log Distance to CBD (zip code) * Income-Distance Gradient 1990 (MSA) - Demeaned	-0.115***	-0.120***	-0.182***	-0.144***
	(0.0383)	(0.0417)	(0.0492)	(0.0466)
Log Distance to CBD (zip code)	-0.0117**	-0.0117**	-0.0390***	-0.0387***
	(0.00580)	(0.00571)	(0.00913)	(0.00862)
Quartile 2 of MSA Median Income (zip code)	0.00474	-0.000567	-0.00497	-0.0305**
	(0.00494)	(0.0112)	(0.00767)	(0.0151)
Log Median Income, 2000 Census (zip code)	0.0525*	0.0453	-0.138***	-0.142***
	(0.0303)	(0.0308)	(0.0363)	(0.0368)
Percent BA or Higher, 2000 Census (zip code)	-0.0238	-0.0305	0.0722	0.0736
	(0.0344)	(0.0339)	(0.0508)	(0.0501)
Owner-Occupied Single-Family Share of Housing Units, 2000 Census (zip code)	-0.148***	-0.142***	-0.0196	-0.0205
	(0.0332)	(0.0332)	(0.0368)	(0.0367)
Log Pop Dens, 2000 Census (zip code)	0.00614*	0.00324	-0.00205	-0.00186
	(0.00344)	(0.00369)	(0.00569)	(0.00594)
Log Distance to CBD (zip code) * Average (zip code income – average income within 1 mile)^2 (MSA-level)		0.000855		0.00842**
		(0.00261)		(0.00419)
Log Distance to Nearest High Price Zip (zip code) * Quartile 2 of MSA Median Income (zip code)		0.00385		0.0206**
		(0.00779)		(0.0102)
Log Distance to Nearest High Price Zip (zip code)		-0.0166**		-0.0107
		(0.00734)		(0.0116)
Constant	-0.160	-0.0402	2.294***	2.345***
	(0.313)	(0.322)	(0.375)	(0.385)
Observations	1,531	1,531	1,531	1,531
R^2	0.957	0.957	0.936	0.937

*** p<0.01, ** p<0.05, * p<0.1. All regressions include metropolitan area fixed effects. Standard errors are clustered at the metropolitan area level.

Table 4
Zip Code Level Regressions of Housing Price Changes, 1982-1989, Full Sample

	(1)	(2)	(3)	(4)	(5)
	Price Growth, 1982-1989				
Log Dist CBD (zip code) * Income-Distance Gradient 1990 (MSA)	-0.104***	-0.113**	-0.0857*	-0.112**	-0.0925*
	(0.0351)	(0.0436)	(0.0441)	(0.0474)	(0.0477)
Log Distance to CBD (zip code)	-0.0140***	-0.00637	-0.0120**	-0.00272	-0.00249
	(0.00500)	(0.00577)	(0.00510)	(0.0169)	(0.0175)
Log Median Income, 2000 Census (zip code)	0.0757***	0.0860***	0.0854***	0.0738***	0.101***
	(0.0222)	(0.0221)	(0.0224)	(0.0214)	(0.0205)
Percent BA or Higher, 2000 Census (zip code)	-0.0228	-0.0273	-0.0355	-0.0206	-0.0460
	(0.0355)	(0.0329)	(0.0387)	(0.0350)	(0.0350)
Percent Own Occ Single Family, 2000 Census (zip code)	-0.137***	-0.140***	-0.131***	-0.135***	-0.137***
	(0.0267)	(0.0257)	(0.0283)	(0.0252)	(0.0260)
Log Pop Dens, 2000 Census (zip code)	0.00521*	0.00469	0.00396	0.00508*	0.00270
	(0.00286)	(0.00287)	(0.00264)	(0.00290)	(0.00252)
Log dist CBD * Saiz supply elasticity missing indicator		-0.00799			-0.0375***
		(0.0104)			(0.0129)
Log dist CBD * Saiz supply elasticity		0.000954			-0.00185
		(0.00462)			(0.00431)
Share of housing stock built in last 10 yrs (zip code)		-0.210***			-0.192***
		(0.0531)			(0.0564)
Log dist CBD * Share of housing stock built in last 10 yrs (zip code)		0.0103***			0.00871**
		(0.00340)			(0.00357)
Share of workers taking public transportation to work, 2000 Census (zip code)			0.165**		0.178***
			(0.0640)		(0.0677)
Log dist CBD * Pub trans share beyond 5 mi minus w/in 5 mi fr CBD			-0.000401		0.00179
			(0.00644)		(0.00729)
Log dist CBD * Emp share w/in 3 mi of CBD missing indicator				-0.00229	0.0313*
				(0.0113)	(0.0158)
Log dist CBD * Emp share w/in 3 mi of CBD, 2006				-0.000492	-1.42e-05
				(0.000686)	(0.000699)
Constant	-0.404*	-0.491**	-0.513**	-0.385*	-0.651***
	(0.223)	(0.222)	(0.225)	(0.217)	(0.208)
Observations	3,342	3,342	3,342	3,342	3,342
R-squared	0.950	0.951	0.951	0.950	0.951

All regressions include metropolitian area fixed effects. Standard errors are clustered at the metropolitan area level.
Missing supply elasticity and centralization data are replaced by the mean of these variables and a dummy variable indicator is used.

Table 5
Zip Code Level Regressions of Housing Price Changes, 1996-2006, Full Sample

	(1)	(2)	(3)	(4)	(5)
			Price Growth, 1996-2006		
Log Dist CBD (zip code) * Income-Distance Gradient 1990 (MSA)	-0.191***	-0.152***	-0.119**	-0.124**	-0.105**
	(0.0457)	(0.0444)	(0.0519)	(0.0505)	(0.0485)
Log Distance to CBD (zip code)	-0.0376***	-0.0478***	-0.0401***	-0.0876***	-0.0762***
	(0.00761)	(0.00878)	(0.00911)	(0.0234)	(0.0216)
Log Median Income, 2000 Census (zip code)	-0.135***	-0.0894***	-0.131***	-0.116***	-0.0781***
	(0.0233)	(0.0263)	(0.0238)	(0.0244)	(0.0274)
Percent BA or Higher, 2000 Census (zip code)	0.0827**	0.0389	0.0732**	0.0578*	0.0248
	(0.0351)	(0.0329)	(0.0357)	(0.0346)	(0.0342)
Percent Own Occ Single Family, 2000 Census (zip code)	-0.0159	-0.0443	-0.0125	-0.0335	-0.0478
	(0.0320)	(0.0338)	(0.0299)	(0.0309)	(0.0319)
Log Pop Dens, 2000 Census (zip code)	-0.00408	-0.00834**	-0.00467	-0.00459	-0.00877**
	(0.00405)	(0.00411)	(0.00407)	(0.00408)	(0.00431)
Log dist CBD * Saiz supply elasticity missing indicator		0.0199			0.00136
		(0.0178)			(0.0171)
Log dist CBD * Saiz supply elasticity		0.00370			0.000440
		(0.00973)			(0.00957)
Share of housing stock built in last 10 yrs (zip code)		-0.192**			-0.198**
		(0.0756)			(0.0795)
Log dist CBD * Share of housing stock built in last 10 yrs (zip code)		0.00441			0.00457
		(0.00443)			(0.00466)
Share of workers taking public transportation to work, 2000 Census (zip code)			0.0299		0.0677
			(0.128)		(0.130)
Log dist CBD * Pub trans share beyond 5 mi minus w/in 5 mi fr CBD			0.0170		0.00733
			(0.0120)		(0.0132)
Log dist CBD * Emp share w/in 3 mi of CBD missing indicator				0.0312**	0.0226
				(0.0120)	(0.0157)
Log dist CBD * Emp share w/in 3 mi of CBD, 2006				0.00169*	0.00141
				(0.000910)	(0.000858)
Constant	2.274***	1.875***	2.239***	2.108***	1.760***
	(0.240)	(0.260)	(0.250)	(0.253)	(0.276)
Observations	3,342	3,342	3,342	3,342	3,342
R-squared	0.938	0.940	0.939	0.939	0.941

All regressions include metropolitian area fixed effects. Standard errors are clustered at the metropolitan area level.
Missing supply elasticity and centralization data are replaced by the mean of these variables and a dummy variable indicator is used.

Figure 1
Price Growth in Periphery Minus Core vs. Income-Distance Gradients