Superstar Cities

By JOSEPH GYOURKO, CHRISTOPHER MAYER, AND TODD SINAI*

We document large long-run differences in average house price appreciation across metropolitan areas over the past 50 years and show they can be explained by an inelastic supply of land in some unique locations combined with an increasing number of high-income households nationally. The resulting high house prices and price-to-rent ratios in those "superstar" areas crowd out lower-income households. The same forces generate a similar pattern among municipalities within a metropolitan area. These facts suggest that disparate local house price and income trends can be driven by aggregate demand, not just changes in local factors such as productivity or amenities.

Keywords: house prices, income inequality, housing supply.

JEL codes: R0, J0, D4, N9

A striking feature of urban housing markets after World War II is the considerable dispersion across U.S. metropolitan areas and towns in long-run real house price appreciation rates. In Figure 1, which plots the kernel density of average annual real house price growth between 1950 and 2000 for 280 U.S. metropolitan areas, average real house price appreciation ranged from about 0.2 percent to over 3.8 percent per year, with an especially thick right tail of growth rates above 2.6 percent.[1] This distribution is not an artifact of a few small areas that grew very rapidly. For example, Table 1, which reports the annualized house price growth rates for the top and bottom ten Metropolitan Statistical Areas (MSAs) with populations above 500,000 in 1950, shows that San Francisco enjoyed an average annualized real house price appreciation rate of more than 3.5 percent. By contrast, Buffalo realized barely 0.5 percent average annual real price growth.

These differences in long-run rates of appreciation led to an ever-widening gap in the price of housing between the most expensive metropolitan areas and the average ones. Figure 2 plots the distribution of log mean real house values across metropolitan areas in 1950 and 2000. In 1950, house prices in the most expensive cities were twice the national average. By 2000, the gap had risen to four times the national average. A similar evolution occurred between 1970 and 2000 among U.S. municipalities.

Why house price dispersion has increased so much over such a long time span is not well understood. Standard compensating differential models in urban economics attribute differences in house prices across markets to differences in the economic value to a household from living in one MSA versus another, with that value driven by factors such as inherent local productivity (and thus wages),

---

[1] The Census data underlying these figures is described more fully below. All monetary amounts are in constant 2000 dollars throughout the paper. The 280 metropolitan areas in Figure 1 had populations of at least 50,000 in 1950.

amenities, or fiscal policies.[2] Differences across markets in the elasticity of housing supply also could lead to differences in capitalization into land prices. However, in order to extrapolate this cross-sectional logic to explain differences in house price growth, long-run changes in local productivity, amenities, or housing supply elasticities would have to match the pattern of dispersion in long-run house price appreciation rates. There is little empirical evidence on whether that is the case.[3]

In this paper, we propose a simple mechanism that generates dispersion in long-run house price growth rates without relying on persistent changes over time in local fundamentals or people's tastes over where to live. Instead, we show that when households have constant preferences over location – perhaps due to cross-sectional differences in local amenities, productivity, or fiscal policies, or to heterogeneity in household tastes for local features – and the supply of places to live is not perfectly elastic everywhere, a change in *aggregate* housing demand is manifested in different local house price growth rates and yields a changing composition of local resident populations.

Locations that experience persistently high house price growth relative to housing unit growth are called "superstars." Two traits are critical to a location being a superstar. There must be some inelasticity to its supply of housing. And, it must be preferred by a large enough share of the population that it has excess demand.

This simple mechanism yields several powerful implications. First, incremental differentiation across locations can yield outsized differences in

house price growth. [4] This means that the gap in house prices between cities or towns can keep increasing even when the inherent value of any particular location is constant, the housing supply has not become more inelastic, and the willingness-to-pay for each location by any individual family is unchanged. When the number of high income families nationally grows, the number of households who would like to reside in any given community increases and the aggregate willingness-to-pay for expensive locations rises (presuming the distribution of households' preferences over where to live is constant). If the growth in latent housing demand for a particular location exceeds the growth in local housing supply due to supply inelasticity, housing rents must rise to clear the market, with lower-income households crowded out by higher-income households.

Second, our mechanism implies that a change in the house price induces a change in the local income distribution. This is in contrast to prior research which assumes that local productivity growth yields higher wages that are then capitalized into house prices. In our model, land prices act as a clearing mechanism by which higher-income households crowd-out lower income households from a scarce location. As the number of high income families grows nationally, existing residents of supply-constrained areas are outbid by even higher-income families, raising the price of land yet further. This process induces a shift to the right in the local income distribution.

Third, the dispersion in expected house price growth rates should yield differences in the price-to-rent ratio for houses. If home buyers in superstar cities expect their houses to appreciate over the long run, they should be willing to pay

---

[4] Our "superstars" terminology is a nod to Rosen's (1981) classic paper on the economics of superstars in the sense that, as the market becomes larger, people pay ever more for slight differences across locations. However, in Rosen's seminal work, a superstar reaps outsized rewards because an incremental improvement in quality yields large increases in market share. In our model, incremental differences among locations yield an increasing house price premium for superstar cities despite declining market share. As the population grows, the market for a differentiated city becomes ever more rarified. In that way, our model is more in keeping with scarce luxury goods than Rosen's concept of a superstar.

more (relative to the rental service flow), today. Of course, in asset market equilibrium, superstar locations do not necessarily have higher risk-adjusted returns. Rather, they are like growth stocks in the sense that higher expected capital gains come at the expense of lower dividend (implicit rent) yields.

Although these implications hold qualitatively as long as housing supply is not perfectly elastic, the elasticity of housing supply in superstar cities – or the difficulty of constructing substitute locations – is a key determinant of their magnitudes. If housing is easily built either in a locality or in a close substitute, there should be little superstar effect. As it becomes more difficult to build new housing or replicate expensive cities, the excess price growth and subsequent shifts in the income distribution can be large.

We use U.S. Census data from 1950-2000 at the metropolitan area and municipality levels to test the implications of our model. Our main empirical result exploits the fact that our model predicts that a single national demand shock should have differential effects on metropolitan areas depending on their superstar status. We report evidence confirming each of the implications just discussed.

For example, when the aggregate number of high-income households in the U.S. grows, house prices in superstar MSAs increase by more than in non-superstar MSAs, and both the average and right tail of the income distributions in superstar cities increase relatively more than in non-superstar cities. We also allow for time-varying superstar status. When a metropolitan area transitions to being a superstar, we see an acceleration in house price growth and in the right-shift of the income distribution. And, in the cross-section, superstar metropolitan areas or municipalities have higher house prices and a higher-income population. Overall, the superstar mechanism explains a substantial portion of the increase in house price dispersion over the last 50 years.

We find that these empirical patterns also hold within a metropolitan area. That is, an increase in the high-income population in a metropolitan area yields

increasing price and income dispersion across municipalities within that area, with price growth, income growth, and income skewness in 'superstar suburbs' outpacing those in non-superstar localities.

We also document that house prices in superstar MSAs and municipalities (within an MSA) are a higher multiple of current rents. When aggregate housing demand increases, those multiples expand more for superstar MSAs than for non-superstars.

This superstar cities mechanism is an important addition to the standard theories of urban housing demand growth, which include productivity shocks or growth in agglomerations. Such theories typically depend on local shocks to housing demand, and it would be unusual for those shocks to match the propagation from higher to lower geographies that we find. In addition, to the extent that productivity growth within a labor market area is more uniform than across metropolitan areas, productivity growth-based theories are less consistent with the house price growth dispersion across communities within a metropolitan area that we observe. However, standard determinants of urban housing demand are likely to be what differentiates locations in the cross section, and our empirical evidence does not preclude them from being additional components of urban housing demand growth.

Finally, we emphasize that our superstar mechanism is intended to explain differences in long-run trend growth rates of house prices, not short-run boom/bust cycles. By definition, superstar MSAs have higher long-run trend house price growth rates than others, but that does not mean their house prices increase every year. Superstar MSAs can and do experience considerable short-run house price volatility, with prices that cycle around strongly positive appreciation trends.

The plan of the paper is as follows. The next section outlines a simple two location model that shows how heterogeneity in location preferences and supply

elasticity combined with growing aggregate demand can combine to generate the patterns in the data described above. Section II then discusses the data used in our analysis. Section III reports the results. There is a brief summary and conclusion.

I. Superstar Cities: A Simple Model

In this section, we derive five implications of the superstar cities mechanism that we will then take to the data. In doing so, we sketch the underlying framework and provide intuition, leaving formal proofs to the appendix.

To focus on the economic forces central to our hypothesis, we simplify the model to a few key elements. One is that we consider only two locations.[5] They differ in their elasticity of land supply. Location A, the always-Available location, has perfectly elastic supply and housing is always available at a normalized rent of zero. Location B has Barriers to development, and thus has a capacity of $K(r)$ households. That capacity can be increased if rents, r, are high enough to make new construction worthwhile, so $K'(r) \geq 0$, with B having perfectly inelastic supply if this holds as an equality.[6] The new capacity could be in location B, or in new locations that are perfect substitutes for B. Thus, $K'(r)$ is finite when newly developed locations are not perfect substitutes for B or are not perfectly elastically supplied.

There are N households in the economy, and each has a constant preference for A or B, denoted by $c_i \sim H(\cdot)$ on [0,1]. A higher c_i corresponds to a greater taste for B. It is not necessary that one location is universally "better" than the

---

[5] These locations could be metropolitan areas or towns within a given area. If the former, the aggregate growth discussed below pertains to the nation; if the latter, it reflects that of the relevant metropolitan area.
[6] Epple and Platt (1998) present a more formal and extensive treatment of a similar model with multiple locations, but assume that land supply is perfectly inelastic in all jurisdictions. By limiting the model to two locations and allowing for elastic supply, we emphasize the testable empirical implications of differences in the elasticity of supply.

other, only that enough households prefer B (for whatever reason) at zero rent to fully occupy it.[7]

Households also vary by their inherent productivity, with type $w_i \sim F(\cdot)$ on $[0, \infty)$, and are paid a productivity wage. We also allow for the possibility that A and B are differently productive in the sense that the same worker would be more productive in one city than the other. A worker of type w_i produces w_i if she works in A, but $\beta w_i + \alpha$ if she works in B, where $\beta \geq 0$. This exogenous difference in location productivity could be due to a variety of factors ranging from a production agglomeration to simple natural advantage. Obviously, the special case of $\alpha = 0$ and $\beta = 1$ reduces to all households being equally productive in either city. We assume that $w_i \perp c_i$, so that households of all abilities have the same distribution of preferences over the two cities.

The utility for household i is denoted by V_i and is a function of being in the preferred location and of non-housing consumption. Household utility in A is given by $V_i^A = (1 - c_i)w_i$, and in B, by $V_i^B = c_i(\beta w_i + \alpha - r)$. Thus, if $c_i = 1$, the household would prefer location B to the exclusion of all else. We will make the common simplifying assumption that there are no costs of moving, so the household chooses whichever location gives it the most utility: $V_i = max(V_i^A, V_i^B)$. This framework further assumes that the marginal rate of substitution between housing and non-housing consumption is zero and that housing can be consumed only in a fixed quantity. These serve to emphasize the households' choice of location and are common assumptions (e.g. Sinai and Souleles (2005)).

In choosing to live in the location where their utilities are highest, households trade off rental costs, their preferences for that location, their incomes, and any

---

[7] The taste parameter is intended to reflect heterogeneity in household preferences for local traits. Examples include the type of amenities, weather, consumption possibilities, and location.

productivity difference. We focus on the case where latent demand to live in B exceeds the space that would be available if rent were zero. If B was not fully occupied, it would be free and households would sort between the locations based only on their tastes and any productivity differences.

Virtually all of the important implications of the superstar mechanism follow from the first lemma which deals with how sorting across locations occurs.

Lemma 1: Conditional on an agent's taste, c_i, and the productivity differences between the two locations, the agent chooses to live in B if her wage, w_i, is greater than a cutoff value $\underline{w}(c_i)$.

A utility maximizing agent prefers B to A when $V_i^B \geq V_i^A$:

(1)
$$c_i(\beta w_i + \alpha - r) \geq (1 - c_i)w_i.$$

Solving this expression in terms of w_i, the cutoff as a function of the agent's wage is given by the following expression:

$$\underline{w}(c_i) \equiv \frac{(r - \alpha)c_i}{(1 + \beta)c_i - 1}.$$

(2)

This cutoff is binding only if $\underline{w}(c_i) \geq \dfrac{r - \alpha}{\beta}$, since for $w_i < \dfrac{r - \alpha}{\beta}$ the agent obtains negative utility from living in B and would always choose to live in A instead, where she is guaranteed non-negative utility. Moreover, unless $r \geq \alpha$, the cutoff trivially binds given $w_i \in [0, \infty)$.

Two relevant corollaries are as follows.

Corollary 1: Conditional on her wage, w_i, an agent with taste greater than a cutoff value $\underline{c}(w_i)$ chooses location B.

Inverting (2), the cutoff value is given by:

(3)
$$\underline{c}(w_i) = \frac{w_i}{(1+\beta)w_i + \alpha - r}$$

Since the wages for agents living in B are constrained to $w_i \in [\frac{r-\alpha}{\beta}, \infty)$, the

taste threshold is bounded in the following way: $\frac{1}{1+\beta} \leq \underline{c}(w_i) \leq 1$.

Corollary 2: If $r > \alpha$, both the wage cutoff $\underline{w}(c_i)$ and the taste cutoff $\underline{c}(w_i)$

are decreasing in their arguments c_i and w_i, respectively, and increasing in the

rental price r.

This can be shown by taking the derivatives of the cutoffs with respect to their

arguments.

(4)
$$\frac{d\underline{w}(c_i)}{dc_i} = \frac{-(r-\alpha)}{[(1+\beta)c_i - 1]^2} < 0$$

(5)
$$\frac{d\underline{c}(w_i)}{dw_i} = \frac{-(r-\alpha)}{[(1+\beta)w_i + \alpha - r]^2} < 0$$

Due to the tradeoff in utility between non-housing consumption and location,

agents with higher tastes for B will sort into B at a lower wage threshold. Due to

the curvature of wages in the utility function, agents with higher wages will sort

into B for lower levels of intrinsic taste for the location. Corollary 1 showed that

$r \geq \alpha$ is a necessary condition for the wage cutoff to bind. That, plus the results

in Lemma 1, and Corollary 1, make it evident that both $\underline{w}(c_i)$ and $\underline{c}(w_i)$ are

increasing in the rent, r.

From these basic results flow the implications of the propositions outlined

next.

Proposition 1: Rent and the average wage are higher in B than in A.

Rent is higher in B than A as long as $K'(r) < \infty$ since rent is zero in A and greater than zero in B due to the assumption of excess demand for B.[8] The difference in rent between B and A increases with the inelasticity of supply in B or of close substitutes to B. Since only households with wages in excess of their cutoffs $\underline{w}(c_i)$ are willing to pay the rent premium to live in B, and tastes are independent of wages, the wage distribution in B is shifted to the right relative to A.

Figure 3 provides the intuition behind this result. It plots $\underline{w}(c_i)$ as a solid line in (c_i, w_i) space. Households with (c_i, w_i) to the southwest of $\underline{w}(c_i)$ will choose to live in A and those to the northeast will pick B. The households in A include those that would prefer A even if B was free ($c_i < 0.5$), and those that would prefer B if rent were lower but will choose A at the clearing rent.[9]

Proposition 2: The share of households that are high income is higher in B than in A.

The intuition follows directly from Proposition 1 and Figure 3. Low-wage households are more likely than high-wage households to defect from B if living there requires paying rent. That implies that the households remaining in B are

---

[8] This rent premium is due to the interaction of the underlying scarcity of land with heterogenous tastes for location rather than the cost of housing structures, which is similar across markets. This is consistent with the literature, which shows that house price differences across markets are greater than construction cost differences (e.g., Gyourko and Saiz (2006)).

[9] Figure 3 assumes the wage distribution is lognormal(0,1), c_i is distributed uniform[0,1], $\alpha = 0$, and $\beta = 1$. The proofs in the appendix show that our results hold for any distribution of wages or tastes, as long as they are independent, and for any productivity differences. Because the B region always occupies the upper right-hand section of Figure 3, as long as the wage distribution is constant across the taste dimension, relatively high income households always will disproportionately live in B. However, if the taste distribution is evolving, it is not necessarily distinct from the income distribution drivers that we consider below in our empirical work. In that section, we report evidence that changes in the income distribution generate a large fraction of the empirical patterns we observe, but that does not preclude changing tastes from being an additional factor. However, we consider it unlikely that changing tastes would be spuriously correlated with the effect of the income distribution so as to confound our estimates.

disproportionately high-income. As Figure 3 illustrates, households in the region choosing B must be weighted toward the high end of the wage distribution.

Proposition 3: Aggregate population growth causes rent growth in B and the effect is increasing in the inelasticity of housing supply.

Corollary 3.1: Aggregate population growth results in an increase in the average wage of agents choosing to live in B.

As the population grows, the absolute number of households that prefer to live in B will increase because the fraction that prefers B is unchanged. Some of the new households have a higher willingness-to-pay for B than some of the old households and, since supply in B is not perfectly elastic, rents increase to clear the market. The less elastic is the supply of B, or of locations that are close substitutes to B, the less location B can accommodate the increased demand, and the more rents must rise. A higher rent requires that a household have a higher wage to be indifferent between A and B, so the average wage of the households that are still willing to live in B is higher.

Supply inelasticity plays an important role in determining the magnitude of the differences between A and B. The proofs in the appendix show that when supply is nearly perfectly elastic the differences between A and B are infinitesimal, and as it becomes harder to bring new supply to market the magnitude of the superstar effect increases. Thus, in practice, supply in B would have to be noticeably inelastic for there to be sizeable differences between A and B in rent growth and consequently their income distributions. Similarly, since we can think of substitute cities for B as akin to additional supply of B, only if perfect substitutes for B are in completely elastic supply would we expect to see no superstar mechanism whatsoever. However, large differences between B and A would arise only if replicating a city is difficult or there are barriers to construction in substitute cities.

The case of perfectly inelastic supply of B is depicted in Figure 3 where the wage cutoff function $\underline{w}(c_i)'$, plotted with a dashed line, corresponds to a doubling of the population relative to $\underline{w}(c_i)$ and an unchanged capacity of B and distribution of wages. Rent nearly triples, from 62 to 164, to clear the market. A smaller *share* of the population now chooses B, and the households that do so on average have a higher taste for B and more income.

Proposition 4: A more skewed aggregate wage distribution with a thicker right tail leads to higher wages and rents in B.

If the wage distribution shifts to the right, more households will have wages in excess of $\underline{w}(c_i)$ and would choose B. This increase in demand for B will be met with higher rent in B as long as new construction in B is not perfectly elastic. A rise in rent increases $\underline{w}(c_i)$ at every c_i, but more so for low-wage households than for high-wage households. The higher indifference wage crowds out relatively low-wage and low-taste-for-B households who will instead live in A. The remaining households in B will be even more relatively high-wage than before, increasing the average wage of residents of B. To the extent that B does not have perfectly inelastic supply, new construction will accommodate some of the new demand, but not enough to fully undo the increase in wages and rent.

The dotted line in Figure 3 shows what would happen in our example if the wage distribution shifted *and* the population doubled without a commensurate increase in the capacity of B. The clearing rent rises to 277, and most lower-wage households are crowded out of B by the new high-wage, high-taste households.

Proposition 5: An anticipated growth in aggregate population or the number of high income individuals in the aggregate population results in a higher price-to-rent ratio in B than in A.

We follow the tradition in the housing literature in presuming that in asset market equilibrium, house price in city m, which we denote P_0^m, equals the expected present value of future rents $r_{t,m}$ plus a risk adjustment or

$$P_0^m = \int_0^\infty r_{t,m} e^{-\delta t} dt + \pi_m$$

(6)

where δ is the discount rate and π_m is the MSA-constant risk premium (e.g., Meese and Wallace (1994), Sinai and Souleles (2005), Ortalo-Magné and Prat (2011)). If rents grow at a constant rate so that $r_{t,m} = r_{0,m} e^{g_m t}$, we obtain the continuous-time Gordon Growth Model with an additional adjustment for risk:

$$\frac{P_{0,m}}{r_{0,m}} = \int_0^\infty e^{-(\delta - g_m t)} dt + \pi_m = \frac{1}{\delta - g_m} + \pi_m.$$

(7)

From equation (7), it follows that the price-to-income ratio is increasing in the growth rate:

$$\frac{d}{dg_m}\left(\frac{P_{0,m}}{r_{0,m}}\right) = \frac{1}{(\delta - g_m)^2} > 0.$$

(8)

By the chain rule of differentiation,

$$g_m = \frac{dr_m}{dt} = \frac{dr_m}{dN}\frac{dN}{dt}$$

(9)

In Proposition 3, we showed that $\dfrac{dr_m}{dN}$ is higher for location B, hence the growth rate of rent in B is higher than in A. Aggregate population growth, $\dfrac{dN}{dt}$, enters identically into equation (9) for locations B and A, so $g_B > g_A$. Since the price-to-rent ratio is increasing in the rental growth rate, there is a higher price-to-rent ratio in B than in A. The second part of this proposition follows similar logic. If agents anticipate a right-ward shift in the wage distribution, then this leads to higher future rents per Proposition 3. Higher future rents precipitate a higher

growth rate of rents by definition. It then follows from the previous results that a higher growth rate in rents leads to a higher price-to-rent ratio in B than in A.

II. Data description

Our primary data source is the six decennial United States Censuses taken between 1950 and 2000. We obtained information on the distributions of house values, rents, family incomes, population, and the number of housing units at two levels of geographical aggregation: metropolitan areas, and Census-designated places, which are municipalities such as cities and towns. All dollar values are converted into constant 2000 dollars using the CPI-U price index.

The MSA data used in our empirical analysis consists of a panel of 279 areas that had populations of at least 50,000 in 1950 and are in the continental United States.[10] Our metro area definition is based on 1990 county boundaries to project consistent metro area boundaries forward and backward through time.[11] Data were collected at the county level and aggregated to the metropolitan statistical area (MSA) or primary metropolitan statistical area (PMSA) level in the case of consolidated metropolitan statistical areas.[12] Data for the 1970-2000 period were obtained from GeoLytics, which compiles long-form data from the decennial *Censuses of Housing and Population*. We hand-collected data spanning 1950 and 1960 from hard copy volumes of the *Census of Population and Housing*. Both

---

[10] Thirty-six areas with populations under 50,000 in 1950 were excluded from our analysis because of concerns about abnormal house quality changes in markets with so few units at the start of our period of analysis. None of our key results are materially affected by this paring of the sample. Similar concerns account for our not using data from the first *Census of Housing* in 1940. (All individual housing trait data from the 1940 census were lost, so we cannot track any trait changes over time from that year.) We did repeat our MSA-level analysis over the 1940-2000 time period. While the point estimates naturally differ from those reported below, the magnitudes, signs, and statistical significance are essentially unchanged. Finally, the New York PMSA is excluded from the analysis reported below because it is missing house price data for 1960.
[11] We use definitions provided by the Office of Management and Budget, available at http://www.census.gov/population/estimates/metro-city/90mfips.txt.
[12] All our conclusions are robust to aggregating to the CMSA level.

sources are based on 100 percent population counts. At the Census place level, we extracted data for 1970-2000 from the GeoLytics CD-ROMs.[13]

The primary strength of using house price data from the decennial censuses is that it is available on a consistent basis over the half century-long period needed for our analysis. The weakness is that the underlying observations are both self-reported and not quality adjusted. However, correlations between constant quality and unadjusted house price series are high over decadal-length periods. For example, the correlation across house price appreciation rates for a large set of consistently defined MSAs in our Census data and the Federal Housing Finance Administration (FHFA) constant quality house price index is 0.94 in the 1980s and 0.87 in the 1990s.

Income also is central to our analysis. To categorize the distribution of income, we divide real family incomes into five categories that are consistent over time. The income categories in the original Census data change in each decade, so we set the category boundaries equal to 25, 50, 75, and 100 percent of the 1960 family income top code, and then populate the resulting five bins using a weighted average of the actual categories in $2000 (assuming a uniform distribution of families within the bins). Since 1960 had amongst the lowest top code in real terms, using it as an upper bound reduces miscategorization of families into income bins. This results in the following bins: We call a family "poor" if its income is less than $36,384 in $2000. "Middle-poor" are those families with incomes between $36,384 and $72,769, "middle" income families have incomes between $72,769 and $109,153, "middle-rich" families lie between $109,153 and $145,537, and "rich" families have incomes in excess of the 1960

---

[13]While states differ in the extent to which local jurisdictions control new construction or even whether they can change their boundaries, census-designated places provide a useful comparable sample. The 1970 data include only 6,963 out of 20,768 places. (Conversations with the Census Bureau suggest that the micro data on the remaining places has been lost or is not readily available.) Fortunately, these places account for more than 95 percent of U.S. population in 1970. In 2000, 161 million people lived in these 6,963 places, 206 million people in all places, and 281 million people in the entire U.S. We further limit the sample to places within a MSA.

real topcode of $145,537. It is important to recognize that the quartiles of the 1960 income top code do not correspond to quartiles of the income distribution; there are far more families in the "poor" category than in the "rich" category. Thus, our choice of income bin boundaries provides more detail in the right tail of the income distribution.

III. Empirical evidence

The underlying conditions necessary for the superstar cities hypothesis to be true have been present in the post-WWII era. Between 1950 and 2000, the number of families in U.S. metropolitan areas doubled, with the number making more than $140,000 in constant $2000 dollars increasing more than eight-fold according to the U.S. Census. And, some metropolitan areas and local communities have more inelastic housing supply than others, either because of local regulation or geographic restrictions (e.g., Gyourko, Saiz, and Summers and (2008); Saiz (2010); Paciorek (2011)). Given that, we now take the model to the data and test the five Propositions from Section I.

A. Defining Superstar Markets

Our first step is to define empirical proxies for the theoretical characteristics of superstar markets: preference for the location and inelastic supply. We use the fact that demand growth has to be manifested either in price growth or housing unit growth to construct these variables. We measure growth in mean real prices and housing units at the MSA level over 20-year periods. This window size gives us growth rates defined over four time periods: 1950-70, 1960-80, 1970-90, and 1980-2000. We identify high demand MSAs by applying a simple cutoff of whether the sum of the price and quantity growth rates for the market is above the sample median. This definition captures the idea that both inelastically-supplied

markets with very limited new construction but high price appreciation and elastic markets with minimal price growth but lots of construction should be categorized as in high demand. We allow the high-demand cutoff to vary over time in order to account for changes in the aggregate economy. For example, a MSA is defined as being in high demand in 1970 if the sum of its price and housing unit growth rates from 1950-70 exceeds the sample median for that period. We proxy for the inverse supply elasticity with the ratio of the price growth rate to the housing unit growth rate. In a city with inelastic supply, demand growth is manifested more in price growth than in housing unit growth, so this ratio should be high.

We define an indicator variable for superstar status (*Superstar*) based on whether a MSA is in the "high demand" category and in the top decile of the ratio of price growth rate-to-housing unit growth rate based on growth rates over the prior two decades.[14] Due to the two-decade lag for computing growth rates, 1970 is the first year for which we can define a superstar. The sample of superstars, broken down by decade in Appendix Table B, includes major metropolitan areas that are superstars in multiple years as well as smaller MSAs that enter and exit superstar status. To purge those noisy MSAs from our sample, we define the MSA-constant $Superstar_i$ as one if the time-varying $Superstar_{it}$ equals one in any two decades between 1970 and 2000, and redefine $Superstar_{it}$ to equal one only for MSAs that are superstars for at least two decades. However, we obtain very similar results when we define $Superstar_i$ as those MSAs that are in the superstar category for one or more census years.

To illustrate where MSAs fall along the dimensions that make up a superstar city, Figures 4 and 5 plot average real annual house price growth against housing

---

[14]As shown in Section I, there can be different degrees of "superstarness," so if a location is more preferred relative to its capacity and/or has less elastic housing supply, it will exhibit more pronounced superstar characteristics. In our empirical analysis, we refer to only the most prominent examples as "superstars" since that lines up well with the empirical distribution, which has a skewed right tail, and makes the exposition more concise. However, all our key conclusions are robust to estimation with a continuous measure, defined as the ratio of the price growth rate to the housing unit growth rate for high demand MSAs or places.

unit growth during the 1960-1980 and 1980-2000 periods, respectively. Three regions are outlined in each figure. Region C, below the negatively-sloped dashed line, corresponds to low demand as defined above (i.e., they have sums of price and housing unit growth rates below the sample median for the corresponding time period). Among the high-demand MSAs, regions B and A are, respectively, above and below the 90th percentile price growth rate to quantity growth rate threshold (about 1.7) that we use for our binary definition of a superstar city. Thus, the markets in Region B are superstars because they are both in the high growth region and in the top decile of inelasticity of supply. By contrast, MSAs in the A range also have high demand, but they have more elastic supply since they are closer to the X-axis and have built more new units relative to the real house price appreciation they experienced.

Beyond providing snapshots of superstar status at two points in time, these two figures also illustrate some of the time series variation that we will exploit in our regression analysis. In the face of geographic constraints or politically-imposed restrictions on development, it seems natural that at least some high-demand metropolitan areas would become more inelastically supplied over time as demand for their scarce housing units becomes larger and they begin to "fill up". This process would appear as a market moving counter-clockwise around the origin over time. We do observe such evolutions. For example, Figure 4 shows that by 1980, San Francisco and Los Angeles qualified as superstars. In 1970 (which is based on data from 1950-1970), both markets were in the A range of the plot. Figure 5 then shows that by the end of our sample period in 2000, 20 more high demand MSAs filled up, also becoming superstars.

At the Census place level, we categorize a place as a superstar if it is both high demand and in the top quartile of the ratio of price growth to unit growth. The methodology for determining which communities are "high demand" and for

computing their supply elasticities is comparable to our MSA-level procedure.[15]

However, the place data are available only from 1970 to 2000 so, after accounting

for the two decades of lags required to compute these variables, our useable place-

level sample covers only 1990 and 2000.

B. Results

Propositions 1 and 2: Do Superstar cities or suburbs have different prices or incomes?—

Propositions 1 and 2 state that superstar MSAs or towns should have higher

house prices and higher average incomes. In addition, the income distributions

should be shifted more to the right – superstars should have relatively larger

shares of their populations in the high-income bins and smaller shares in the low-

income bins. We will first see if these predictions hold in the cross-section for

MSAs, then for Census places, and then within MSAs as they change Superstar

status. Appendix Tables 1A and 1B report summary statistics for all variables

used in this section.

We estimate the following bivariate regression using our panel of 279 MSAs

over four two-decade periods (for a total of 1,116 MSA × year observations):

$$Y_{it} = \beta_1 \text{Superstar}_i + \delta_t + \varepsilon_{it}$$

for MSA i in year t.[16] The dependent variable, Y_{it}, takes a variety of outcomes,

including the log house value, log income, and the share of families in each of the

income categories. Year dummies also are included. Thus, the estimated

---

[15] A place is considered to be high-demand if its sum of price and housing unit growth rates exceeds that period's median across all places in all MSAs. The 75th percentile ratio of price growth rate to unit growth rate for places (2.0) is close to the 90th percentile for MSAs (1.7) because the distribution for places has thicker tails than for MSAs.

[16] To be faithful to the model in section I, we compare the outcomes for superstar cities to all other metropolitan areas. In the model, cities with less demand than capacity are perfectly elastically supplied, whereas cities with excess demand can have varying supply elasticities. We compare the high-demand, inelastic cities to both low-demand cities and high-demand, elastic cities. In practice, not all low-demand cities appear to have elastic supply by our measures. This may be due to measurement error, or house prices being below construction cost as in Glaeser and Gyourko (2005). However, we have obtained comparable results even when including a separate control for low-demand areas.

coefficient β_1 measures the average difference between MSAs that ever are superstars and other MSAs.

The results are reported in panel (A) of Table 2. Superstar status is associated with higher average and 10th percentile house values, higher average income, a greater share of the MSA's residents in the high-income categories and a lower share in the low-income categories. Moreover, the point estimates are economically large as well as statistically significant. For example, in the first column, where the dependent variable is the log of the MSA's average house value, the estimated coefficient of 0.6053 for log house value (0.0729 standard error) implies that superstar MSAs have 60 percent higher average house values. The second column uses the log of the MSA's 10th percentile house value as the dependent variable in an attempt to more closely reflect the changing minimum entry price for an MSA due to rising land values, as well as better control for differences in spending on the structure component of housing. The estimated coefficient is even larger, 0.7844 (0.0855 standard error).[17]

In the cross-section, superstar MSAs also have income distributions that are shifted to the right relative to other MSAs. In column 3, average incomes are shown to be nearly 24 percent higher (standard error of about 3 percent) in superstars. We look at other points in the income distribution in columns 4 through 8. The outcome variables in these columns correspond to each of the five income bins (y) in each MSA i in year t: $Y_{it} = \frac{\# \ in \ income \ bin_{yit}}{\# \ of \ households_{yit}}$. For example, in column 4, we find that the mean share of households in the 'rich' group in superstar MSAs is 3.4 percentage points higher than in other MSAs. Since the income distribution outcome variables in columns 4 through 8 are not in logs, in those columns we report the estimated elasticity (in square brackets) in addition to

_____

[17] While our empirical results, in keeping with the prior literature, focus on house values, the model in section I is expressed in terms of rents. We have obtained comparable results throughout this paper when using log average rent as the dependent variable.

the usual point estimate. For example, since the share of the income distribution that is in the "rich" category averages just 3.3 percent, the estimated effect of 0.0339 amounts to a 101.7 percent increase in the "rich" share relative to the average. In addition, we find that superstars have an 80 percent higher share middle-rich, 41 percent higher share middle-income, and 26 percent lower "poor" households.

Panel B of Table 2 reports the analogous results using Place-level data. In this case, our research design can control nonparametrically for any confounding unobservable factor that might vary by MSAs or across MSAs over time, using only variation across towns within an MSA in a given year to identify our estimates.[18] The place-level version of our regression is:

$$Y_{kit} = \beta_1 \text{superstar}_k + \delta_{it} + \varepsilon_{kit}$$

The unit of observation is now Census place k in MSA i in year t. Superstar status is determined at the place level. MSA × year fixed effects also are included.

The Place-level results in panel (B) of Table 2 exhibit very similar patterns to those found across MSAs. Superstar towns in an MSA have higher house values and average incomes than other towns in the same MSA and year. They also have income distributions that are shifted to the right. In particular, we find that average house values in superstar towns are 37 percent higher (column 1) and incomes are 24 percent higher (column 3). The share of the population in the top two income categories is substantially greater, and the fraction of the population in the bottom two income categories experiences an offsetting decline.

Another approach to controlling for unobservable MSA-level characteristics is to see what happens when an MSA becomes a superstar. By using the time-

---

[18] For example, to the extent that productivity growth is more constant within MSAs than between them, our Place-level estimates, by controlling nonparametrically for MSA x year unobservables, will better control for it. However, Moretti's (2011) and Baum-Snow and Pavan's (2011) evidence that households sort among cities differentially based on growth in the returns to skill underscore how difficult it is to completely reject productivity-based explanations.

varying definition of superstar, we can include MSA-level fixed effects, thus controlling for unobserved differences across MSAs that could confound the relationship between demand, supply elasticity, and the outcome variables. Instead, the identification strategy measures how much the outcome variables change when the MSA is a superstar versus when it is not. In the bottom panel of Table 2, we report estimates from the following regression equation:

$$Y_{it} = \beta_1 \text{Superstar}_{it} + \delta_t + \gamma_i + \varepsilon_{it}$$

Relative to panel (A), we have added a MSA fixed effect, γ_i, and allowed Superstar_{it} to vary over time, as defined in the Data section.

We find the same pattern in within-MSA differences over time that we observed across MSAs. MSAs experience increases in house prices and average incomes, and become more rich and less poor, when they are in the Superstar region. The effect on house values and average income are smaller than in panel (A), but still economically large and statistically significant. This pattern indicates that MSAs that become superstars had higher house prices and incomes than other MSAs prior to becoming superstars, and experienced an additional jump in house prices and average incomes after becoming superstars. The coefficients on the income bins are also smaller in panel (C) than in panel (A), though the fundamental pattern and significance is maintained. The share of an MSA's population in the top two income categories increases by at least 90 percent when the MSA enters the superstar region, the middle-income category grows by about 6 percent, and the share in the bottom two income categories falls between 4 and 14 percent.

Propositions 3 and 4: Are superstars differently affected when the aggregate income distribution changes?—

At the national/MSA levels, the model implies that when either the U.S. population or share of the population that is high-income increases, land prices

should rise fastest and the local income distributions should shift to the right the most in superstar MSAs. We do not try to distinguish between the effects of population and income share in our empirical analysis, instead combining the two factors into one measure: the number of high-income families. We then look for empirical evidence at the national/MSA and then MSA/place levels.

The top panel of Table 3 reports results from our national/MSA regression specification that relates a time-varying MSA outcome to changes in the national income distribution and time-invariant differences across MSAs in their superstar status. Specifically, the regression equation takes the following form:

$$Y_{it} = \beta_1 \text{Superstar}_i \times \ln(\# \text{Rich}_t) + \gamma_i + \delta_t + \varepsilon_{it}$$

for MSA i in year t. The dependent variable, Y_{it}, takes the usual set of outcomes.

The first regressor interacts the time-invariant Superstar indicator with the log of the number of households at the national level that are in the "rich" income bin ($\ln(\#\text{Rich}_t)$). The Superstar indicator varies across MSAs and the number of rich households varies over time, so the interaction varies over time within MSA. Thus, the estimated coefficient β_1 measures how changes in the number of rich families at the aggregate level differentially affect superstar cities relative to all other cities. The MSA fixed effects (γ_i) control for MSA-level unobserved heterogeneity and the year dummies (δ_t) absorb influences that vary only over time, such as aggregate macroeconomic factors. These fixed effects also subsume the uninteracted effects of supply elasticity or the aggregate number of rich families.[19]

The results support Propositions 3 and 4. In the first column, where the dependent variable is the log of the MSA-average house value, the estimated coefficient of 0.3943 (0.0356 standard error) indicates that house values rise by more in more inelastic, high demand MSAs when the national number of rich

---

[19] We obtain similar results by taking first-differences within MSAs.

families increases. We observe a smaller, though still statistically significant, effect on the 10th percentile house value.

To get a sense of the magnitudes, consider that between 1970 and 2000 the number of rich families in the U.S. grew by 160 percent. In the first column, the average house values in superstar MSAs are estimated to rise by 39 percentage points more than in other MSAs when the number of rich families nationally doubles. In actuality, mean house prices in superstar MSAs grew 75 percentage points more than in other MSAs, so the pressure of the growing national income distribution can account for more than 80 percent of the excess growth in house prices in Superstar cities in that specification.[20]

The remaining columns of panel (A) of Table 3 address the implications of Propositions 2 and 3 that the rise in house prices in superstar cities should also affect the distribution of local incomes. Column 3 uses the log of the mean income in the MSA as the dependent variable. The estimated coefficient of 0.1292 (0.0143) in the first row implies that doubling of the number of rich families in the country is associated with an 12.9 percentage point higher growth rate in average income in a superstar MSA. This represents all of the actual difference in the growth in average income between superstar MSAs and other MSAs over the 1970 to 2000 period.

Columns 4 through 8 report the estimated effects of growth in the national number of rich families at the various points in an MSA's income distribution. These results show that when the national number of rich families increases, the income distribution shifts to the right more in superstar MSAs. Relative to other MSAs, superstars experience a larger increase in their share of households that are in the highest-income categories and a bigger decline in their middle-low-income households. For example, the estimated coefficient of 0.0407 (0.0022) in the first row of column 4 implies that a doubling of the number of rich families nationally

_____

[20] $160 \times 0.3943 = 63.1$, which is 84.1 percent of 75.

would increase the share of households in the "rich" category for superstar cities by 4 percentage points more than in other MSAs. A similar, but smaller, effect is found among the "middle-rich" households in column 5, and no effect is found for middle-income households.

At the other end of the income spectrum, a doubling of the number of national rich families would yield more than a 6 percentage point excess decline in the share of households in the "middle-poor" category, consistent with the higher-income households crowding out the lower-income ones. We discern little differential change between superstar MSAs and other MSAs in the share of households in the "poor" category.[21]

These results also help distinguish the superstar cities mechanism from other potential sources of local housing demand. It seems unlikely that local growth (for example, changes in the α or β parameters from the model in Section I) would match the geographic pattern, timing, and linkage to the national income distribution of MSA price growth that is predicted by our framework. In addition, potential confounding effects due to defining superstar cities based in part on average house price growth over the entire sample period are mitigated by directly controlling for the superstar nature of an MSA, thereby identifying the effect from the interaction of those variables with changes in the national income distribution.

The 'superstar suburbs' logic implies that the number of rich families at the MSA level should be positively correlated with house price growth, income growth, and the rich share of families at the Census place level. The place-level version of our regression is:

$$Y_{kit} = \beta_1 \text{superstar}_k \times \ln(\# \text{ Rich}_{it}) + \gamma_k + \delta_{it} + \varepsilon_{kit}.$$

The unit of observation is now Census place k in MSA i in year t. Superstar status is determined at the place level, and Superstar$_k$ is set equal to one if the Census place is in the superstar region in either 1990 or 2000. The aggregate growth in the number rich is measured at the MSA × year level. Place and MSA × year fixed effects also are included.

The results reported in the bottom panel of Table 3 reveal a similar pattern to the MSA results in the top panel. The magnitudes on the estimated coefficients are attenuated, but remain statistically significant. In sum, there is substantial evidence among towns within a given metropolitan area that aggregate, MSA-level changes have disproportionate impacts on prices, wages, and the share rich in superstar communities that have inelastic supplies and are in strong demand.[22]

Proposition 5: Price-to-Rent Ratios in Superstar Markets —

Proposition 5 stated that prices would be a greater multiple of rents in superstar markets if growth in rents was anticipated. The first column of Table 4 uses the cross sectional, MSA-level specification from panel (A) of Table 2, but with the log of the MSA-average price-to-rent ratio as the dependent variable. The estimated coefficient of 0.3145 from the first row indicates that, on average, prices are a 31 percent larger multiple of rents in superstar MSAs. Column 2 repeats the cross-section analysis at the Census place level, with MSA × year fixed effects. We find that superstar suburbs have a 26 percent higher price-to-rent ratio than other towns. This pattern persists when MSAs transition to superstar status (Column 3). In the years that MSAs are superstars, their price-to-

---

[22] One drawback of this level of geography is that our place-level data date only to 1970, which makes it more difficult to assess within-town changes over time. Because $(\Delta P / \Delta Q)_{ki}$ requires two lagged decades to construct, we observe only one change per Census place – between 1990 and 2000. Essentially, we are estimating whether the change in the left hand side variable between 1990 and 2000 is related to the growth in the number of rich families in the "parent" MSA over that time period. Because each of the 279 "parent" MSAs experienced different rates of growth in the number of rich families between 1990 and 2000, we have plenty of variation to identify the effects on the Census places within those MSAs. We also have applied the measure of superstar status defined over the 1990-2000 period to the entire 1970 to 2000 sample, with consistent results.

rent ratios are 27 percent greater than in the years when they are not superstars. The last two columns of Table 4 relate changes in the price-to-rent ratio at the MSA or Place levels to changes in the number of "rich" households at an aggregate geography. In both cases, when the number of "rich" households increases, the price-to-rent ratio goes up by more in superstar MSAs or Places.

It is worth underscoring that this result is consistent with standard asset market equilibrium. Homeowners in superstar markets do not necessarily obtain a higher return; instead, they receive a higher expected capital gain at the expense of a lower current yield. In that way, superstar markets are like growth stocks in the equity investment universe.

C. Ex ante *versus* ex post *definitions of Superstar status*

As a robustness check, we redefined our proxy for superstar status using *ex ante* MSA characteristics rather than *ex post* realizations of price and quantity growth. The model in Section I implies that it is a combination of supply inelasticity and high demand for the location that defines superstar status. Our *ex ante* definition of an inelastically supplied MSA is one that is in the top decile of Saiz's (2010) topography-based measure of the difficulty of building. We have two approaches to defining "high-demand" based on *ex ante* data. For one, we denote the top third of MSAs ranked by the sum of their price and housing unit growth in the pre-sample period of 1950-1970 as high-demand. In the other, we denote the top third of MSAs ranked by their mean January temperature as high-demand. Including our baseline definitions, we had two proxies for the elasticity of supply and three proxies for high demand. We replicated all the MSA-level specifications using each of the six combinations of these definitions, with the exception of Panel C of Table 2. Since neither the Saiz elasticity measure nor the

ex ante high-demand proxies are time-varying, we could not use combinations of these variables to estimate the effect of changing superstar status.

In Table 5, we report the estimated coefficients corresponding to the two combinations that used only the three new *ex ante* definitions. The alternative definitions yield lower estimates than in our baseline results, but they remain economically and statistically meaningful. The estimates from the specifications corresponding to panel (A) of Table 2 are reported in the top panel of Table 5. Each row corresponds to a different construction of Superstar status. Superstar MSAs exhibit higher house prices – about the same magnitudes as in Table 2 – and a right-shift in the income distribution that is about half the magnitude of that reported in Table 2. The bottom panel of Table 5 estimates the effect of growth in the national number of rich households on the newly-defined Superstar MSAs, akin to panel (A) of Table 3. These estimates are typically 40 to 50 percent lower in magnitude than in our baseline estimates, but they are still economically and statistically significant. For example, in the first row of the bottom panel, the estimated coefficient in the regression of log house value on the interaction of Superstar status with the log number of rich households nationally is 0.2685 (0.0361). This estimated coefficient implies that the 160 percent increase in the national number rich between 1970 and 2000 would yield a 43 percent increase in house prices in Superstar cities, or about 57 percent of the actual growth. In the third column, the estimated coefficient of 0.0836 (0.0144) corresponds to a 13 percent greater increase in log mean income for Superstar MSAs over the same time period.

IV. Conclusion

This paper argues that much of the growing dispersion in house prices in the post-WWII era is the consequence of aggregate population growth and the

skewing of incomes nationally interacting with preferences for location and differences in local supply conditions. This combination of conditions has generated an important economic and social phenomenon. Because high house prices disproportionately crowd out lower-income potential residents, the evolution of entire metropolitan areas into superstars influences the way we spatially organize our society. Mere population growth forces residency in preferred cities and towns effectively to be auctioned off to the highest bidder, with existing landowners in those places benefitting from the rise in prices. In contrast to the standard urban growth analysis, the house price growth in superstar cities that we describe is not due to an increasing service flow or greater productivity.

Although our analysis does not rule out a role for other factors such as persistent differences in local productivity, we provide empirical evidence at the MSA and Census Place geographies that is consistent with the superstar cities mechanism described above being one of the key forces in effect. Our data suggests that as much as two-thirds of the growth in dispersion in house prices, and almost all of the growth in dispersion in average incomes, between superstar MSAs and others over the 1970-2000 period can be explained by the increase in high-income households at the national level.

Our framework also helps us understand the conditions under which widening dispersion in house prices may or may not continue. For the superstars mechanism to be operative, metropolitan areas must be differentiated and in limited supply, and there must be growth in aggregate housing demand. If cities that are close substitutes to a superstar city can be created, the superstar location effectively has a higher supply elasticity and the superstar effect would be smaller. Similarly, an increase in the elasticity of supply in a superstar city itself also would attenuate its excess price growth. Our results imply that despite increasing prices over the last 50 years, close substitutes to superstar cities have

either failed to arise or have not grown fast enough to fully offset the superstar effect.

Our model does suggest two other factors that could affect the superstar mechanism. First, household preferences could shift significantly away from a superstar location. Second, Superstar cities are sensitive to changes in aggregate demand. When housing demand increases, superstar cities and suburbs achieve disproportionate growth in house prices and changes in their income distributions. When housing demand contracts, the opposite should be true. The Great Recession reminds us that aggregate growth can falter substantially, both in terms of income increases and household formation. It is the waxing and waning of these factors which seem most likely to determine whether superstar cities maintain the same high long-run house price growth over the next 50 years as they did over the previous five decades.

APPENDIX: Proofs of Propositions 1-4:

Proposition 1: Rent and the average wage are higher in B than in A.

Proof. Let $E(w^B)$ denote the average wage in B. By the law of iterated expectations, the average wage in B is given by:

(1)
$$E(w^B) = E[E[w_i \mid w_i \geq \underline{w}(c_i)]]$$

The innermost calculation $E[w_i \mid w_i \geq \underline{w}(c_i)]$ gives the expected wages for all agents whose wages exceed the threshold $\underline{w}(c_i)$.

(2)
$$E[w_i \mid w_i \geq \underline{w}(c_i)] = \int_0^\infty f(w_i \mid w_i \geq \underline{w}(c_i))w_i dw_i$$
$$= \int_{\underline{w}(c_i)}^\infty \frac{w_i f(w_i)}{1 - F(\underline{w}(c_i))} dw_i$$

Taking the expectation over all possible thresholds $\underline{w}(c_i)$ gives the mean wages of all agents whose wages and decision rule results in them optimally sorting into B:

(3)
$$E(w^B) = E[E[w_i \mid w_i \geq \underline{w}(c_i)]] = \int_0^1 \left[\int_{\underline{w}(c_i)}^\infty \frac{w_i f(w_i)}{1 - F(\underline{w}(c_i))} dw_i \right] h(c_i) dc_i$$

We now make use of the fact that rents are higher in B than in A. In particular rents in B are *r* and rents in A are set to zero. To show that mean wages are higher in B, it suffices to show that $E[w^B]$ is an increasing function of rents, $\dfrac{dE[w^B]}{dr} > 0$. Before continuing with the next step of the proof, introduce the following simplification in our notation $\underline{w}(c_i) \equiv \underline{w}$. By the Fundamental Theorem of Calculus, we obtain the following result:

$$\frac{\partial E[w^B]}{\partial r} = \int_0^1 \left[\frac{f(\underline{w})}{[1-F(\underline{w})]^2} \left(\frac{\partial \underline{w}}{\partial r} \right) \int_{\underline{w}}^{\infty} f(w_i) w_i dw_i \right] h(c_i) dc_i - \int_0^1 \left[\frac{f(\underline{w})\underline{w}}{1-F(\underline{w})} \left(\frac{\partial \underline{w}}{\partial r} \right) \right] h(c_i) dc_i$$

$$\geq \int_0^1 \left[\frac{f(\underline{w})}{[1-F(\underline{w})]^2} \left(\frac{\partial \underline{w}}{\partial r} \right) \int_{\underline{w}}^{\infty} f(w_i)\underline{w} dw_i \right] h(c_i) dc_i - \int_0^1 \left[\frac{f(\underline{w})\underline{w}}{1-F(\underline{w})} \left(\frac{\partial \underline{w}}{\partial r} \right) \right] h(c_i) dc_i$$

(4)

$$\geq \int_0^1 \left[\frac{f(\underline{w})\underline{w}}{[1-F(\underline{w})]^2} \left(\frac{\partial \underline{w}}{\partial r} \right) \int_{\underline{w}}^{\infty} f(w_i) dw_i \right] h(c_i) dc_i - \int_0^1 \left[\frac{f(\underline{w})\underline{w}}{1-F(\underline{w})} \left(\frac{\partial \underline{w}}{\partial r} \right) \right] h(c_i) dc_i$$

$$\geq 0.$$

The inequality introduced in the second step of this computation comes from the fact that $w_i \in [\underline{w}, \infty)$ is bounded below by $\underline{w} \geq 0$ and the wage cutoff is increasing in the rental price $\left(\frac{\partial \underline{w}}{\partial r} \right) > 0$ (Corollary 2).

Proposition 2: The share of individuals that are high income is higher in B than in A.

Proof. Without loss of generality we fix a level of wealth w_H such that agents with earnings $w_i \geq w_H$ are considered to be high income individuals. The share of high income individuals in B, $S_{H,B}$ is given by:

(5) $S_{H,B} = E[Prob(w_i \geq w_H \mid w_i \geq \underline{w}(c_i))] = \int_0^1 Prob(w_i \geq w_H \mid w_i \geq \underline{w}(c_i)) h(c_i) dc_i$

The conditional probability in equation (5) is given by:

$$Prob(w_i \geq w_H \mid w_i \geq \underline{w}(c_i)) = \begin{cases} \dfrac{Prob(w_i \geq max(w_H, \underline{w}))}{1-F(\underline{w})}, & \text{if } w_i \geq max(\underline{w}, w_H) \\ 0, & \text{else} \end{cases}$$

First we define $c_H \equiv \underline{c}(w_H)$. By Corollary 2, $c_i \leq c_H \Leftrightarrow \underline{w}(c_i) \geq \underline{w}(c_H)$. By Corollary 1, $\underline{w}(c_H) = \underline{w}(\underline{c}(w_H)) = w_H$. We use these derived taste cutoffs to rewrite the conditional probability as:

$$Prob(w_i \geq w_H \mid w_i \geq \underline{w}(c_i)) = \begin{cases} \dfrac{1-F(w_H)}{1-F(\underline{w})}, & \text{if } c_i \geq c_H \\ 1, & \text{if } c_i < c_H \end{cases}$$

We use this result to simplify the expression in equation (5) for $S_{H,B}$:

(6)
$$S_{H,B} = \int_0^{c_H} h(c_i)dc_i + \int_{c_H}^1 \frac{1-F(w_H)}{1-F(\underline{w})} h(c_i)dc_i$$

$$\geq H(c_H)$$

In the second to the last step we use the fact that $c_H \leq 1 \Rightarrow \underline{w}(1) \leq \underline{w}(c_H)$. In the

final step, we use $\underline{w}(1) = \dfrac{r-\alpha}{\beta}$.

For the second part of the proof, we compute $S_{H,A}$, the share of high income

individuals in city A, which is given by the following expression:

(7) $\quad S_{H,A} = E[Prob(w_i \geq w_H \mid w_i \leq \underline{w}(c_i))] = \int_0^1 Prob(w_i \geq w_H \mid w_i \leq \underline{w}(c_i))h(c_i)dc_i$

The conditional probability in equation (7) is given by:

$$Prob(w_i \geq w_H \mid w_i < \underline{w}(c_i)) = \begin{cases} \dfrac{Prob(w_H \leq w_i < \underline{w})}{F(\underline{w})}, & \text{if } w_H \leq w_i \leq \underline{w} \\ 0, & \text{else} \end{cases}$$

By Corollary 2, $w_H \leq w_i \leq \underline{w}(c_i) \Leftrightarrow c_H \geq \underline{c}(w_i) \geq c_i$, which reduces to $c_i \leq c_H$.
Rewriting the conditional probability we obtain:

$$Prob(w_i \geq w_H \mid w_i < \underline{w}(c_i)) = \begin{cases} 1 - \dfrac{F(w_H)}{F(\underline{w})}, & \text{if } c_i \leq c_H \\ 0, & \text{else} \end{cases}$$

$$S_{H,A} = \int_0^{c_H} \left[1 - \frac{F(w_H)}{F(\underline{w})}\right] h(c_i)dc_i$$

(8)
$$= H(c_H) - \int_0^{c_H} \left[\frac{F(w_H)}{F(\underline{w})}\right] h(c_i)dc_i$$

$$\leq H(c_H)$$

Comparing $S_{H,B}$, the share of high income individuals in B from equation (6)

with $S_{H,A}$, the share of high income individuals in A from equation (8), we find

that there is a weakly greater share of high income individuals in B than in A:

(9) $$S_{H,A} \leq H(c_H) \leq S_{H,B}$$

When $c_H > 0$, or equally $w_H > \dfrac{r-\alpha}{1+\beta}$ this inequality in is strict and $S_{H,A} < S_{H,B}$.

If $w_H < \dfrac{r-\alpha}{1+\beta}$ then by Lemma 1 everyone in B is classified as high income since

$w_H < \dfrac{r-\alpha}{\beta}$. Consequently $S_{H,B} = 1$ and the inequality is again strict.

Proposition 3: Aggregate population growth causes rent growth in B and the effect is increasing in the inelasticity of housing supply.

Proof. First let N be the aggregate population and N_B be the number of agents

living in B. We then define a binary variable D_i such that $D_i = 1$ if individual i

lives in B and $D_i = 0$ if individual i lives in A. The number of agents living in B

is then given by $N_B = N \times E[D_i]$. The proof proceeds in two steps. First we solve

for $E[D]$ using the law of iterated expectations, i.e. $E[D] = E[E[D|c]]$ and set

that equal to the capacity in B at the equilibrium rental rate r. Second, we

differentiate this expression with respect to r and solve out for $\dfrac{\partial r}{\partial N}$ in order to

show that it is positive.

(10) $$E[D|c_i] = Prob[w_i \geq \underline{w}(c_i)] = 1 - F(\underline{w}(c_i))$$

Now by the law of iterated expectations:

(11) $$E[D] = \int_0^1 E[D|c_i]h(c_i)dc_i = \int_0^1 [1 - F(\underline{w})]h(c_i)dc_i = 1 - \int_0^1 F(\underline{w})h(c_i)dc_i$$

Hence the number of individuals living in B is given by:

(12) $$N\left(1-\int_0^1 F(\underline{w})h(c_i)dc_i\right) = K(r)$$

Differentiating equation (12) with respect to r yields:

(13) $$\frac{dN}{dr}\left(1-\int_0^1 F(\underline{w})h(c_i)dc_i\right) - N\int_0^1\left[f(\underline{w})\frac{\partial w}{\partial r}\right]h(c_i)dc_i = \frac{dK}{dr}$$

Rearranging this equation into the desired form:

(14) $$\frac{dr/r}{dN/N} = \frac{1}{\dfrac{dK/K(r)}{dr/r} + \dfrac{r\int_0^1\left[f(\underline{w})\frac{\partial w}{\partial r}\right]h(c_i)dc_i}{1-\int_0^1 F(\underline{w})h(c_i)dc_i}}$$

The manipulation in the second line above comes inserting the definition of equation (12). The change of integration limits in the third line come from the assumption that $f(w)$ has support on $w \in [0,\infty)$. Setting $\underline{w}(c_i) \geq 0 \Rightarrow c_i \geq \frac{1}{1+\beta}$.

Since $\frac{dK}{dr} > 0$ by assumption (i.e. higher capacity is offered for a higher market clearing rent, and $\frac{\partial w}{\partial r} > 0$ by Corollary 2, we obtain $\frac{dr/r}{dN/N} > 0$. In particular,

$\frac{dr/r}{dN/N}$ is decreasing in the elasticity of supply $\frac{dK/K}{dr/r}$.

Corollary 3.1: Aggregate population growth results in an increase in the average wage of agents choosing to live in B.

Proof. The wage cutoff for living in B is increasing with aggregate population growth, i.e. $\frac{d\underline{w}(c_i)}{dN} = \frac{c_i}{(1+\beta c_i)-1}\frac{dr}{dN} \geq 0$. This follows from the fact that $\frac{dr}{dN} > 0$

and the minimum taste cutoff for agents living in B is $\underline{c}_{min} = \dfrac{1}{1+\beta}$, both of which

are shown in Proposition 3. The proof that $\dfrac{\partial E[w^B]}{\partial N}$ follows via a similar

computation to the one in equation (4), with N taking the place of B.

Proposition 4: A more skewed aggregate wage distribution with a thicker right tail leads to higher wages and rents in B.

Proof. Let F(w) and G(w) denote two non-identical wage cdfs with common

support $w \in [0,\infty)$, where F(w) first order stochastically dominates G(w). First

order stochastic dominance of F over G captures the fact that F has a thicker right

tail than G, i.e. $\int_a^\infty f(w_i)dw_i \geq \int_a^\infty g(w_i)dw_i \;\; \forall w_i$ and $\int_a^\infty f(w_i)dw_i > \int_a^\infty g(w_i)dw_i$

for some w_i. Let $E_F[w^B]$ denote the expected wage of individuals in B under the

wage distribution F, and $E_G[w^B]$ the expected wage of individuals in B under the

wage distribution G. We now show that $E_F[w^B \mid c_i] \geq E_G[w^B \mid c_i] \forall c_i \in [0,1]$, which

implies $E_F[w^B] \geq E_G[w^B]$. As in Proposition 3, we define a binary variable D_i

such that $D_i = 1$ if individual *i* lives in B and $D_i = 0$ if individual *i* lives in A.

Taking the conditional expectation of D under the wage distributions F and G

yields $E_F[D \mid c_i] = 1 - F(\underline{w}(c_i))$ and $E_G[D \mid c_i] = 1 - G(\underline{w}(c_i))$. By first order

stochastic dominance of F over G, $E_F[D \mid c_i] > E_G[D \mid c_i]$. Moreover, by the

intermediate value theorem $\exists \hat{w}(c_i) \in [0, \underline{w}(c_i)]$ such that

$F(\underline{w}(c_i)) = G(\hat{w}(c_i)) \leq G(\underline{w}(c_i))$. Using this result in concert with the definition of

the conditional wage functions given in equation (3) we complete the proof as

follows:

$$E_F[w^B] - E_G[w^B] = \int_0^1 \left[\int_{\underline{w}(c_i)}^\infty \frac{w_i f(w_i)}{1-F(\underline{w}(c_i))} dw_i \right] h(c_i)dc_i - \int_0^1 \left[\int_{\underline{w}(c_i)}^\infty \frac{w_i g(w_i)}{1-G(\underline{w}(c_i))} dw_i \right] h(c_i)dc_i$$

$$= \int_0^1 \left[\int_{\underline{w}(c_i)}^\infty \frac{w_i f(w_i)}{1-G(\hat{w}(c_i))} dw_i \right] h(c_i)dc_i - \int_0^1 \left[\int_{\underline{w}(c_i)}^\infty \frac{w_i g(w_i)}{1-G(\underline{w}(c_i))} dw_i \right] h(c_i)dc_i$$

$$\geq \int_0^1 \left[\int_{\underline{w}(c_i)}^\infty \frac{w_i f(w_i)}{1-G(\hat{w}(c_i))} dw_i \right] h(c_i)dc_i - \int_0^1 \left[\int_{\underline{w}(c_i)}^\infty \frac{w_i g(w_i)}{1-G(\hat{w}(c_i))} dw_i \right] h(c_i)dc_i$$

$$\geq \int_0^1 \left[\int_{\underline{w}}^\infty \frac{w_i f(w_i) - w_i g(w_i)}{[1-G(\hat{w}(c_i))]} dw_i \right] h(c_i)dc_i$$

$$\geq \int_0^1 \left[\frac{1}{[1-G(\hat{w}(c_i))]} \int_{\underline{w}}^\infty (w_i f(w_i) - w_i g(w_i)) dw_i \right] h(c_i)dc_i$$

$$> 0.$$

Since $E_F[w^B] - E_G[w^B] > 0$, it follows directly that $E_F[w^B] \geq E_G[w^B]$. We now show that r_F, the equilibrium rent under the wage distribution F(w) is greater than r_G, the equilibrium rent under the wage distribution G(w). Using the result in equation (12), the equilibrium rents r_F and r_G are given by:

(15)
$$K(r_F) = N\left(1 - \int_0^1 F(\underline{w})h(c_i)dc_i\right)$$

(16)
$$K(r_G) = N\left(1 - \int_0^1 G(\underline{w})h(c_i)dc_i\right)$$

Applying the first order stochastic dominance condition to equation (16) we obtain:

(17)
$$K(r_F) = N\left(1 - \int_0^1 F(\underline{w})h(c_i)dc_i\right)$$
$$> N\left(1 - \int_0^1 G(\underline{w})h(c_i)dc_i\right)$$
$$= K(r_G) > K(r_G).$$

Since $K(r)$ is an increasing function of r, $K(r_F) > K(r_G) \Leftrightarrow r_F > r_G$. This completes the second part of Proposition 4.

Appendix Table A1: MSA Summary Statistics

Variable	Mean	Standard deviation
MSA time-invariant characteristics (N=279):		
Average Annual Real House Price Growth, 1950-2000	1.57	0.56
Average Annual Housing Unit Growth, 1950-2000	2.10	0.98
Average Annual Real Income Growth, 1950-2000	1.82	0.35
Ever a "superstar"	0.075	0.264
Ever "low demand"	0.738	0.440
MSA time-varying characteristics (N=1,116):		
Average 20-year Real House Price Growth	1.50	1.04
Average 20-year Housing Unit Growth	2.10	1.20
Average 20-year house price growth + housing unit growth	3.60	1.86
Average ratio of 20-year price growth to 20-year unit growth	0.869	1.148
Real house value	111,329	54,889
Average price/average annual rent	17.00	3.99
Year	# "superstars"	
1970	0	
1980	2	
1990	21	
2000	20	
Income Distribution	Mean	Standard deviation
Share of an MSA's population that is "rich"	0.033	0.021
Share "middle-rich"	0.035	0.024
Share "middle"	0.129	0.043
Share "middle-poor"	0.400	0.050
Share "poor"	0.402	0.095
National number "rich"		
1970	1,571,136	
1980	1,312,103	
1990	2,611,178	
2000	4,098,324	

Appendix Table A2: Place Summary Statistics

Variable	Mean	Standard deviation
Place time-invariant characteristics (N=3,788):		
Avg. Real House Price Growth (1970-2000)	0.015	0.011
Avg. Housing Unit Growth (1970-2000)	0.017	0.019
Avg. Real Income Growth (1970-2000)	0.007	0.007
Ever a "superstar"	0.220	0.414
Ever "low demand"	0.618	0.486
Place time-varying characteristics: (1990-2000; N=7,576))		
Average 20-year Real House Price Growth	0.015	0.017
Average 20-year Housing Unit Growth	0.016	0.021
Average 20-year house price growth + housing unit growth	0.031	0.028
Mean real house value	156,736	125,401
10th Percentile house value	90,757	79,123
Average price/average annual rent	17.76	7.60
Year	# "superstars"	
1990	653	
2000	580	
Income Distribution (1990-2000)	Mean	Standard deviation
Share of an place's population that is "rich"	0.061	0.035
Share "middle-rich"	0.069	0.030
Share "middle"	0.174	0.041
Share "middle-poor"	0.372	0.039
Share "poor"	0.323	0.085
MSA number "rich"		
1990	26,789	36,031
2000	39,582	49,513

Appendix Table B: Superstar MSAs by year

City name	1970	1980	1990	2000
Albany			X	X
Allentown			X	
Atlantic City			X	
Baltimore			X	
Bellingham	X			
Bergen-Passaic			X	X
Boston			X	X
Bremerton	X			
Detroit				X
Dutchess County			X	X
Enid		X		
Glens Falls			X	
Hartford			X	
Jersey City			X	X
Lewiston			X	
Los Angeles		X	X	
Louisville				X
Middlesex-Somerset-Hunterdon			X	
Nassau-Suffolk County			X	X
New Haven			X	X
New London			X	X
Newark			X	X
Oakland			X	X
Orange County			X	
Philadelphia			X	X
Pine Bluff	X			
Pittsfield			X	X
Portland			X	
Providence			X	X
Reading			X	
Salinas			X	X
San Francisco		X	X	X
San Jose			X	X
Santa Barbara-Santa Maria			X	X
Santa Cruz			X	X
Springfield, MA			X	X
Trenton			X	X
Ventura			X	

Notes: 241 MSAs that are never superstars are excluded from the table. Rows shaded in grey correspond to MSAs that achieve superstar status in two or more decades. This subset of MSAs are defined as superstars in our regression analysis. The empirical results are robust to defining all MSAs in this table as superstars. Expanding the definition yields slightly lower magnitudes of the estimated coefficients and slightly larger standard errors, but the results remain economically and statistically significant.

REFERENCES

Baum-Snow, Nathaniel and Ronni Pavan. 2011. "Inequality and City Size,"
        Review of Economics and Statistics, forthcoming.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny. 2010 "Spatial Sorting:
        Why New York, Los Angeles and Detroit Attract the Greatest Minds as
        Well as the Unskilled." Working paper, September.

Epple, Dennis and Glenn J. Platt. 1998. "Equilibrium and Local Redistribution in
        an Urban Economy when Households Differ in both Preferences and
        Incomes," *Journal of Urban Economics*, 43(1): 23-51.

Epple, Dennis. and Holger Sieg. 1999." Estimating Equilibrium Models of Local
        Jurisdictions. " *Journal of Political Economy*, 107(4): 645-681.

Glaeser, Edward and Joseph Gyourko. 2005. "Urban Decline and Durable
        Housing," *Journal of Political Economy,* 113(2): 345-375.

Glaeser, Edward and Kristina Tobio. 2008. "The Rise of the Sunbelt," *Southern
        Economic Journal*, 74(3): 609-643.

Gyourko, Joseph and Albert Saiz. 2006. "Construction Costs and the Supply of
        Housing Structure," *Journal of Regional Science*, 46: 661-680.

Gyourko, Joseph, Albert Saiz, and Anita Summers. 2008. "A New Measure of the
        Local Regulatory Environment for Housing Markets," *Urban Studies*,
        45(3): 693-729.

Lee, Sanghoon. 2010. "Ability Sorting and Consumer City," *Journal of Urban
        Economics*, 68(1): 20-33.

Meese, Richard and Nancy Wallace. 1994. "Testing the Present Value Relation
        for Housing Prices: Should I Leave My House in San Francisco?" *Journal
        of Urban Economics*, 35(3): 245-266.

Moretti, Enrico. 2011. "Real Wage Inequality," NBER Working Paper No. 14370.
        Mimeo, UC Berkeley.

Ortalo-Magné, Francois, and Andrea Prat. 2011. "On the Political Economy of Urban Growth: Homeownership versus Affordability," CEPR Discussion Papers 8243, University of Wisconsin working paper.

Ortalo-Magné, Francois and Sven Rady. 2008. "Heterogeneity within Communities: A Stochastic Model with Tenure Choice," *Journal of Urban Economics*, 64(1): 1-17.

Paciorek, Andrew. 2011. "Supply Constraints and Housing Market Dynamics," The Wharton School working paper.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90(6): 1257-78.

Rosen, Sherwin. 1981. "The Economics of Superstars," *The American Economic Review*, 71(5): 845-858.

Rosen, Sherwin. 1979. "Wage-Based Indexes of Urban Quality of Life". In *Current Issues in Urban Economics*, edited by Peter Mieszkowski and Mahlon Straszheim. Baltimore: Johns Hopkins Univerity Press, 1979.

Rosenthal, Stuart and Willam Strange. 2003. "Geography, Industrial Organization, and Agglomeration," *Review of Economics and Statistics*, 85(2): 377-393.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125(3): 1253-96.

Sinai, Todd and Nicholas Souleles. 2005. "Owner Occupied Housing as a Hedge Against Rent Risk," *Quarterly Journal of Economics*, 120(2): 763-89.

Van Nieuwerburgh, Stijn and Pierre-Olivier Weill. 2010. "Why Has House Price Dispersion Gone Up?," *Review of Economic Studies*, 77(4): 1567-606.

Waldfogel, Joel. 2003. "Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated Product Markets," *RAND Journal of Economics*, 34(3): 557-68.

Table 1: Real annualized house price growth, 1950-2000,

Top and Bottom 10 MSAs with 1950 population>500,000

Top 10 MSAs by Price Growth Annualized growth rate, 1950-2000		Bottom 10 MSAs by Price Growth Annualized Growth Rate, 1950-2000	
San Francisco	3.53	San Antonio	1.13
Oakland	2.82	Milwaukee	1.06
Seattle	2.74	Pittsburgh	1.02
San Diego	2.61	Dayton	0.99
Los Angeles	2.46	Albany (NY)	0.97
Portland (OR)	2.36	Cleveland	0.91
Boston	2.30	Rochester (NY)	0.89
Bergen-Passaic (NJ)	2.19	Youngstown-Warren	0.81
Charlotte	2.18	Syracuse	0.67
New Haven	2.12	Buffalo	0.54

Population-weighted average of the 48 MSAs in this sample: 1.71

Table 2: MSA and Place-level differences in superstar MSAs

Left-hand-side variable:	Log House Value	Log 10th Percentile House Value	Log mean income	Share of families in the ____ category				
				Rich	Middle-rich	Middle	Middle-poor	Poor
(A) Pooled MSA cross-sections, Year fixed effects:								
$Superstar_i$ [Relative to the mean of the LHS variable]	0.6053 (0.0729)	0.7844 (0.0855)	0.2360 (0.0308)	0.0339 (0.0058) [1.017]	0.0284 (0.0039) [0.804]	0.0524 (0.0066) [0.405]	-0.0094 (0.0061) [-0.023]	-0.1053 (0.0137) [-0.262]
Adj. R^2	0.4162	0.3149	0.4605	0.4151	0.5952	0.3584	0.1177	0.1969
Mean of LHS:	11.54	10.64	10.84	0.033	0.035	0.129	0.400	0.402
(B) Pooled Place cross-sections, MSA × year fixed effects:								
$Superstar_k$ [Relative to the mean of the LHS variable]	0.3668 (0.0225)	0.2723 (0.0163)	0.2448 (0.0118)	0.0759 (0.0032) [1.231]	0.0167 (0.0015) [0.257]	-0.0006 (0.0022) [-0.004]	-0.0427 (0.0028) [-0.114]	-0.0493 (0.0047) [-0.148]
Adj. R^2	0.4766	0.6804	0.3795	0.2535	0.4258	0.4040	0.2779	0.4381
Mean of LHS:	11.72	11.16	11.01	0.062	0.065	0.166	0.375	0.332
(C) MSA panel, MSA and year fixed effects:								
$Superstar_{it}$ [Relative to the mean of the LHS variable]	0.4427 (0.0304)	0.2744 (0.0511)	0.1224 (0.0127)	0.0325 (0.0021) [0.977]	0.0316 (0.0020) [0.895]	0.0082 (0.0032) [0.063]	-0.0543 (0.0062) [-0.136]	-0.0180 (0.0081) [-0.045]
Adj. R^2	0.8679	0.8311	0.9121	0.8203	0.8683	0.8948	0.7046	0.8584

Notes: In Panels (A) and (C), the number of observations is 1,116, for four decades (1970-2000) and 279 MSAs. In Panel (B), the sample period is 1990-2000 and covers 3788 Census Places over two decades (N=7,576). Standard errors are in parentheses. Standard errors are clustered by MSA in panel (A). The specification in panels (A) and (B) is $Y_{it} = \beta_1 Superstar_i + \delta_t + \varepsilon_{it}$, where the superstar variable is not time-varying and is defined at the MSA level in panel (A) and the Place level in panel (B). In panel (C), the specification is $Y_{it} = \beta_1 Superstar_{it} + \delta_t + \gamma_i + \varepsilon_{it}$, where the superstar variable is defined at the MSA level and varies over time. The means of the left-hand-side variables in panel (C) are the same as in panel (A).

Table 3: How Changes in the Aggregate Number of Rich Households Differentially Affect Superstar MSAs and Places

Left-hand-side variable:	Log House Value	Log 10th Percentile House Value	Log mean income	Share of families in the ___ category				
				Rich	Middle-rich	Middle	Middle-poor	Poor
(A) MSA panel, MSA and year fixed effects:								
$Superstar_i \times \log(\# Rich_t)$ [Relative to the mean of the LHS variable]	0.3943 (0.0356)	0.1992 (0.0578)	0.1292 (0.0143)	0.0407 (0.0022) [1.222]	0.0310 (0.0023) [0.877]	-0.0003 (0.0036) [-0.002]	-0.0624 (0.0069) [-0.156]	-0.0090 (0.0091) [-0.022]
Adj. R^2	0.8555	0.8277	0.9110	0.8336	0.8591	0.8940	0.7060	0.8577
Mean of LHS:	11.54	10.64	10.84	0.033	0.035	0.129	0.400	0.402
(B) Pooled Place cross-sections, MSA × year and Place fixed effects:								
$Superstar_i \times \log(\# Rich_{kt})$ [Relative to the mean of the LHS variable]	0.1565 (0.0126)	0.0972 (0.0226)	0.0857 (0.0114)	0.0292 (0.0031) [0.473]	-0.0009 (0.0023) [-0.014]	-0.0116 (0.0040) [-0.070]	0.0016 (0.0056) [0.004]	-0.0182 (0.0056) [-0.055]
Adj. R^2	0.8123	0.6439	0.5804	0.4447	0.3689	0.3644	0.3412	0.4056
Mean of LHS:	11.72	11.16	11.01	0.062	0.065	0.166	0.375	0.332

Notes: In Panel (A), the sample period is 1970-2000 and covers 279 MSAs over four decades (N=1116). In Panel (B), the sample period is 1990-2000 and covers 3788 Census Places over two decades (N=7,576). Standard errors are in parentheses. In Panel (A), the specification is $Y_{it} = \beta_1 Superstar_i \times \ln(\# Rich_t) + \gamma_i + \delta_t + \varepsilon_{it}$, where an indicator variable for an MSA ever being a "superstar" during the entire 1970-2000 period is interacted with the log national number of families in the "rich" category. In Panel (B), the specification is $Y_{it} = \beta_1 Superstar_k \times \ln(\# Rich_t) + \gamma_{it} + \delta_k + \varepsilon_{ikt}$, where an indicator variable for a Census Place ever being a "superstar" during the 1990-2000 period is interacted with the MSA's log number of families in the "rich" category. In both panels, the uninteracted variables are subsumed by the fixed effects.

Table 4: Price-to-Rent Ratio Results

Table #:	2A	2B	2C	3A	3B
Variation	Pooled Cross Section	Pooled Cross Section	Within-MSA changes	National # Rich	MSA # Rich
Geography	MSA	Place	MSA	MSA	Place
Superstar_i	0.3145 (0.0437)	0.2661 (0.0196)			
Superstar_{it}			0.2717 (0.0216)		
$\text{Superstar}_k \times \log(\text{\# Rich}_{it})$				0.2222 (0.0253)	0.1229 (0.0143)
Fixed effects:	Year	MSA × year	MSA, year	MSA, year	MSA × year, place
Adj. R^2	0.4030	0.2867	0.7936	0.7754	0.6639
N	1,116	7,576	1,116	1,116	7,576

Notes: The left-hand-side variable is the log average price/rent ratio.

Table 5: Robustness to alternative definitions of Superstar

Construction of Superstar$_i$	Log House Value	Log 10th Percentile House Value	Log mean income	Share of MSA families in the ____ category				
				Rich	Middle-rich	Middle	Middle-poor	Poor
(A) Right-hand-side variable: Superstar$_i$ (corresponds to Panel (A) of Table 2)								
Saiz supply elasticity + January temp	0.5101 (0.0965)	0.6349 (0.1264)	0.1023 (0.0381)	0.0174 (0.0054) [0.524]	0.0134 (0.0045) [0.380]	0.0198 (0.0092) [0.153]	-0.0182 (0.0049) [-0.045]	-0.0325 (0.0187) [-0.081]
Saiz supply elasticity + price growth 1950-70	0.5690 (0.0886)	0.7539 (0.1091)	0.1196 (0.0373)	0.0185 (0.0055) [0.556]	0.0145 (0.0045) [0.411]	0.0230 (0.0092) [0.178]	-0.0097 (0.0069) [-0.024]	-0.0464 (0.0186) [-0.115]
Fixed effects:	Year	Year	Year	Year	Year	Year	Year	Year
(B) Right-hand-side variable: Superstar$_i$ × log(# Rich$_{kt}$) (corresponds to Panel (A) of Table 3)								
Saiz supply elasticity + January temp	0.2685 (0.0361)	0.1093 (0.0568)	0.0836 (0.0144)	0.0239 (0.0024) [0.717]	0.0144 (0.0024) [0.407]	-0.0043 (0.0035) [-0.033]	-0.0168 (0.0070) [-0.042]	-0.0171 (0.0089) [-0.043]
Saiz supply elasticity + price growth 1950-70	0.3026 (0.0367)	0.1217 (0.0580)	0.0938 (0.0147)	0.0253 (0.0025) [0.760]	0.0159 (0.0025) [0.451]	-0.0018 (0.0036) [-0.014]	-0.0156 (0.0072) [-0.039]	-0.0239 (0.0091) [-0.059]
Fixed effects:	MSA, year	MSA, year	MSA, year	MSA, year	MSA, year	MSA, year	MSA, year	MSA, year
Mean of LHS:	11.54	10.64	10.84	0.033	0.035	0.129	0.400	0.402

Notes: Number of observations is 1,116, for four decades (1970-2000) and 279 MSAs. Standard errors are in parentheses. The specification in the first panel is $Y_{it} = \beta_1 Superstar_i + \delta_t + \varepsilon_{it}$, where the superstar variable is defined at the MSA level and is not time-varying. The specification in the second panel is $Y_{it} = \beta_1 Superstar_i \times \ln(\# Rich_t) + \gamma_i + \delta_t + \varepsilon_{it}$, where an indicator variable for an MSA ever being a "superstar" during the entire 1970-2000 period is interacted with the log national number of families in the "rich" category. The uninteracted variables are subsumed by the fixed effects. Marginal effects relative to the mean of the LHS variable are in square brackets.

Figure 1: Density of 1950-2000 Annualized Real House Price Growth Rates
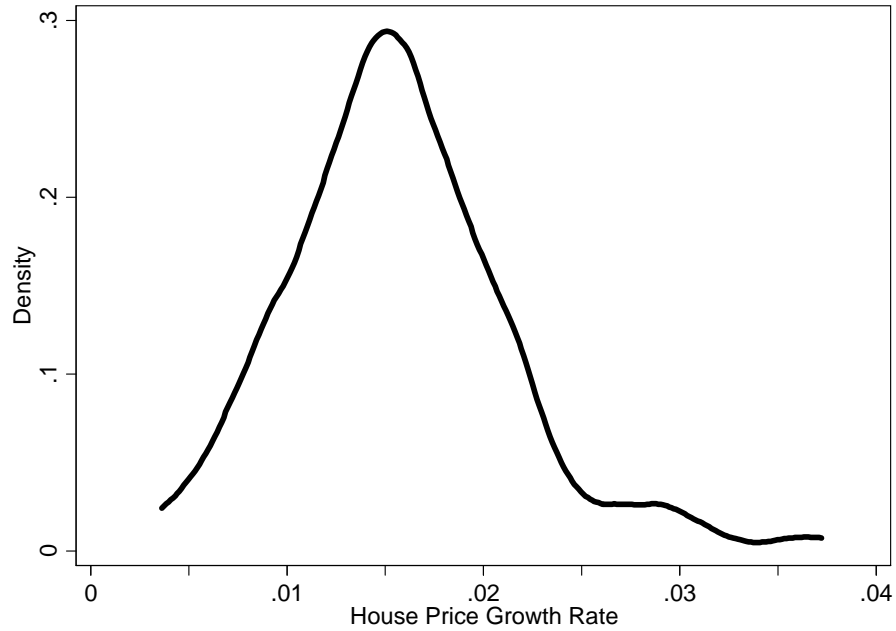Across MSAs with 1950 population > 50,000

Figure 2: Density of Mean House Values Across MSAs (1950 vs. 2000)
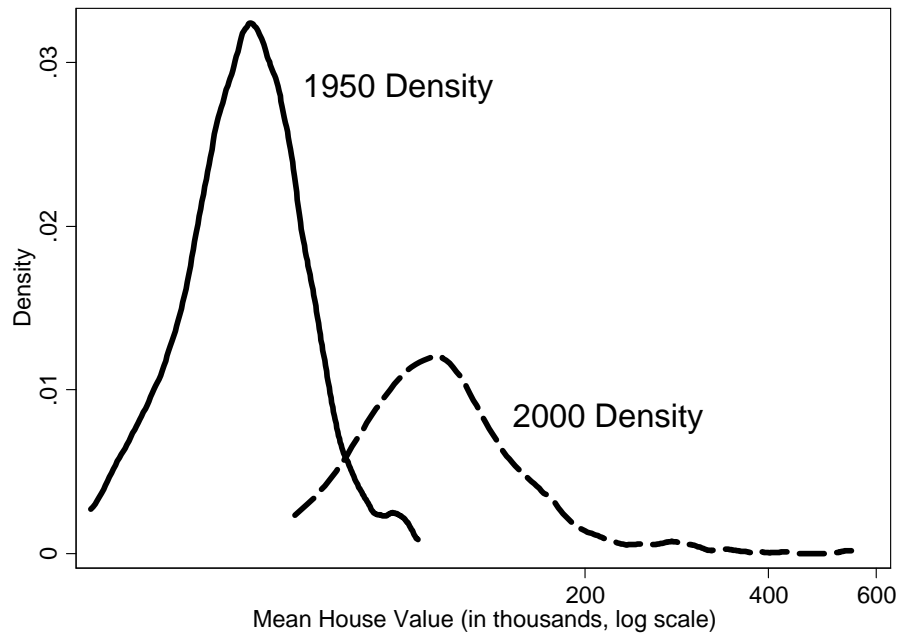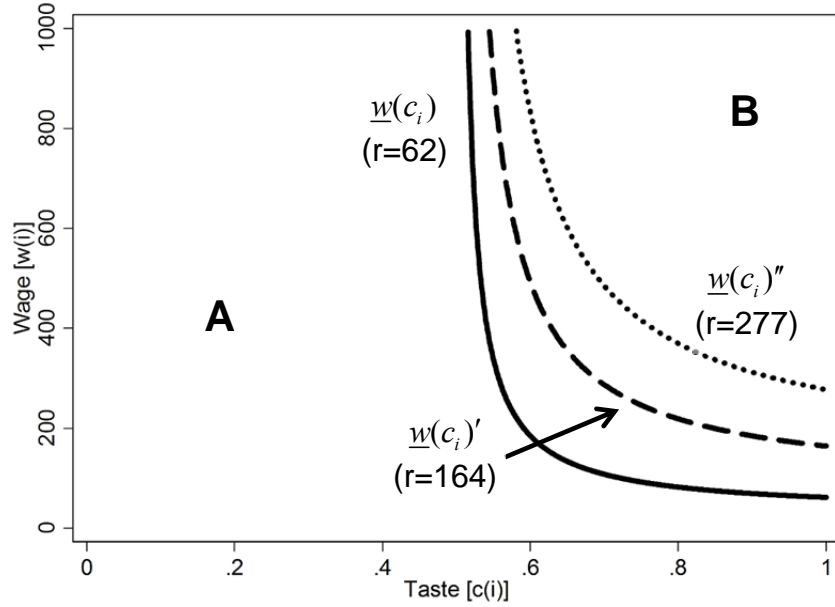
Figure 3: Sorting between locations A and B as population increases and the wage distribution shifts

Notes: $\alpha = 0; \beta = 1$. Tastes are distributed uniform $(0,1)$. $K'(r) = 0$ and capacity in B is $0.33N$. Rent is computed in equilibrium based on Appendix Equation 12. For $\underline{w}(c_i)$ and $\underline{w}(c_i)'$, wages are distributed lognormal $(0,1)$. $\underline{w}(c_i)'$ corresponds to the case when N doubles and capacity in B is unchanged. $\underline{w}(c_i)''$ corresponds to the case when N has doubled and wages are distributed lognormal $(0.7, 0.9)$.

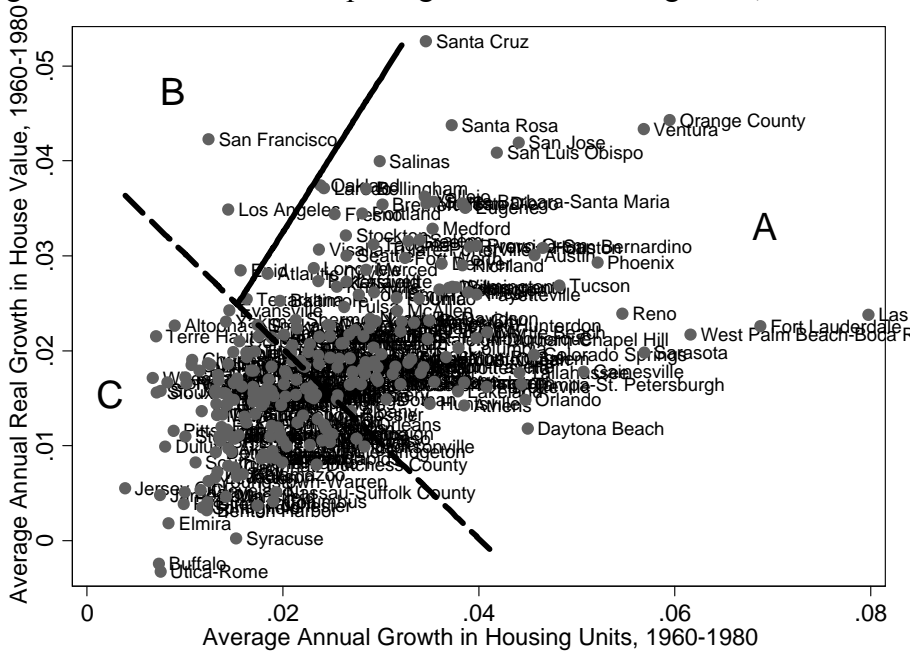Figure 4: Real annual house price growth versus unit growth, 1960-1980

Figure 5: Real annual house price growth versus unit growth, 1980-2000