PRELIMINARY AND INCOMPLETE

DO NOT CITE

SUPERSTAR CITIES

Joseph Gyourko
The Wharton School
University of Pennsylvania
gyourko@wharton.upenn.edu

Christopher J. Mayer
Graduate School of Business
Columbia University
and NBER
cm310@columbia.edu

Todd Sinai
The Wharton School
University of Pennsylvania
and NBER
sinai@wharton.upenn.edu

July 12, 2004

Dimo Pramaratov and Dan Simundza provided outstanding research assistance. We appreciate the comments and suggestions of Dennis Yao, Joel Waldfogel, Nick Souleles, Albert Saiz, Susan Guthrie, and participants at the Wharton School Applied Economics Workshop. We also benefited greatly from conversations with other friends and colleagues in the early stages of this project. Gyourko and Sinai received financial support from the Research Scholars program of the Zell-Lurie Real Estate Center at Wharton. Mayer was funded by the Milstein Center for Real Estate at Columbia Business School.

Contact authors: Sinai – 1465 Steinberg Hall-Dietrich Hall, 3620 Locust Walk, Philadelphia, PA 19104-6302 / (215) 898-5390 / sinai@wharton.upenn.edu; Mayer – 3022 Broadway, Uris Hall #808, New York, NY 10025 / (212) 854-4221 / cm310@columbia.edu.

Between 1940 and 2000, the real average house price in the San Jose, California metropolitan area grew more than eleven-fold. In nearby San Francisco, real house prices increased by a factor of nine. In third-ranked Seattle, real house prices grew to be more than six times what they were in 1940. Outside of the top ten cities, however, real house prices at most tripled over that 60-year period.

We label the phenomenon that leads to this remarkable skewness in house price appreciation over such a long time period as one of 'superstar cities.' A superstar city has a limited supply of places in which to live, and in the face of increasing demand exhibits rising prices. More importantly, a superstar city is unique enough that there is no close substitute city, thus keeping the supply inelastic. A superstar city is not necessarily everyone's top choice of a place to live, but it is preferred by some people, and it cannot be easily replicated to satisfy their demand. The more people that prefer the city, the more expensive it will become relative to others.

We do not equate urban success with superstar status unless it is accompanied by a significant rise in prices. For example, one might wish to term Las Vegas a superstar for its phenomenal population growth—which has been near or above 50 percent per decade for four of the last five decades. However, given Las Vegas' relatively low price appreciation, anyone who would like to live there, can. Our superstar cities are by their nature exclusionary – and due to the prices they command, residents have to pay a significant financial premium to live there. Thus, our focus is on the consequences of the interaction of an urban area being in inelastic supply (for whatever reason), population growth, and the skewing of the income distribution we have seen in the post-WWII era. The consequences of these events have been selective price appreciation and spatial sorting along income lines.

We show in a simple model that the residents of an inelastically supplied city will be disproportionately made up of the right tail of the national income distribution. These households can outbid the poorer ones for the scarce good, and land prices are the mechanism by which spaces in the city are allocated. As the national population –especially the high-income population – grows, the superstar city will become more expensive and the composition of the population will change. Increased demand leads to higher house prices and eventually higher income households will displace poorer ones as the bidding for the limited places to live escalates. We show that this process can occur without high-income households having different preferences than low-income households.

This pattern is exactly the one observed in the data. Using metropolitan area-level census data from 1960 to 2000, we document how income distributions evolve in superstar cities relative to high-demand, but elastically supplied cites. While every high-demand city gains high-income households as the national population grows, only superstar cities experience an increase in the share of their population that is high-income. Elastically-supplied cities experience an increase in the number of poor households over time, while the number and share of poor households in superstar cities goes down.

This view provides a different perspective on house price appreciation from the previous literature. For example, the quality-of-life literature (e.g., Rosen (1979); Roback (1980, 1982)) shows that house prices can be permanently higher in one city than another due to differences in amenities, but the gap is capped by the value of the marginal utility those differences provide. In contrast, a superstar cities theory predicts that the price differences can be ever-increasing, at least as long as there is population growth among the top portion of the national income distribution, and the city remains desirable and in limited supply.

Other theories explaining differences in urban price growth, such as those involving agglomeration, have typically emphasized productivity differences across metropolitan areas. For example, external economies in production can lead to valuable agglomeration benefits and greater growth (e.g., Glaeser, et. al. (1992); Henderson, et. al. (1995); Rosenthal and Strange (2001, 2003)). There also may be human capital externalities, so there would be productivity spillovers associated with a concentration of highly-educated people (e.g., Glaeser, et. al (1995); Glaeser and Saiz (2003); Moretti (2003, 2004a, 2004b), Rauch (2003)).[1]

Our goal is not to disprove these theories, *per se*. In fact, productivity-based theories are complementary with ours in the sense that scale economies or labor complementarities can create a lack of substitutability across cities that helps generate the inelastic supply that drives the superstar cities outcome. However, agglomeration or other productivity-based theories cannot easily explain the complete set of facts that we will present, below. For example, one prediction of superstar cities is that rich people are more likely to move into a superstar MSA because they can more easily afford the premium. Conversely, agglomeration theories imply that once someone moves to a city, they will become richer, since they will be the beneficiary of the production externalities. We test this distinction by looking a the distribution of movers in and out of superstar MSAs, and find that high-income families are more likely to move in.

A second distinction between superstars and production agglomeration is that the "fractal" version of superstar cities also should hold. That is, the same superstar pattern should hold at the suburban level within a MSA, as it does across MSAs, as long as the suburbs of

[1] This is not the place for a review of the agglomeration literature and these more recent cites do not encompass the whole of that large body of research. Alfred Marshall (1890) generally is credited with bringing the issue to the general attention of economists in his famous 19th century textbook. Chinitz (1961) and Jacobs (1969) wrote influential pieces on the topic in the modern era. More recent research is indebted to Romer's (1986) and Lucas' (1988) work on increasing returns, which suggested that cities would provide an excellent laboratory for testing such models. The most recent volume (#4) of the *Handbook of Regional and Urban Economics* contains some excellent reviews of the recent literature. The interested reader should see Duranton and Puga (2004) on the theoretical foundations of agglomeration economies and Rosenthal and Strange (2004) on recent empirical contributions.

MSAs cannot easily expand their boundaries and those suburbs have some attributes that cannot easily be replicated. As long as workers can commute from one part of the MSA to another, agglomeration benefits should help all workers equally within an MSA, even if production externalities occur at a very tight level geographically as reported in Rosenthal and Strange (2003). We examine evidence from places within MSAs in the 1970-2000 US Censuses. Our evidence shows that superstar places exist as well; that is, as the number of wealthy workers in an MSA grows, high price-growth places capture an increasing percentage of those high-income workers.

Another possibility that we will only address indirectly is increasing returns to density or scale associated with consumption, leading to greater demand to live in a vibrant area with people of similar tastes (e.g., Glaeser, Kolko and Saiz (2001), Waldfogel (2003)) or that high-income households have different preferences than low-income households. Our model is based on supply restrictions combined with preferences by some households to live in a superstar city. To the extent that high income households like the superstar city more than other households, possibly because the superstar city already has a large number of high income households (and corresponding amenities such as high-end restaurants or art galleries), the results in our paper are amplified.

Another way of framing the difference between superstar cities and these agglomeration-based theories is that superstar cities explains the evolution of house prices without appealing to anything but population growth (especially among higher income households). For consumption or production-based agglomeration theories to explain the time path of house prices, amenities must be increasing rapidly over time or significant productivity growth occurring that happens to coincide geographically with inelastically supplied cities.

The suggestion that returns to being in a city might not be due to the creation of economic value is potentially important itself. For economists, the perception of cities is quite different if one believes that the growth in incomes and house prices in cities primarily is because agglomeration facilitates spillovers, or instead because some cities are scarce factors with high house prices effectuating sorting. For the lay person, the superstar city logic, while intuitive, leads to conclusions that do not always sit well. Most people think they should be able to live where they want and bemoan the rising cost of housing. But the superstar cities model argues that living in certain places such as the Bay Area is a luxury, not a right.[2] Our model suggests that we will see increasing segregation by income across urban areas, not just within them, with the top places possibly becoming more like expensive resort communities such as Aspen, Colorado. If so, this would constitute a significant change in how we live and interact with one another on a daily basis.

In the next section of this paper, we document house price growth in U.S. metropolitan areas to show that a few cities have very high and persistent long-run real rates of growth. Section 2 presents a model of superstar cities that makes predictions about the evolution of the income distribution within and across cities. In section 3, we turn to the data, first categorizing cities as superstars or not, then testing whether the empirical income patterns match those predicted by the data. Section 4 briefly concludes.

[2] One of the co-authors was having lunch at a conference where the person sitting next to him was complaining that his daughter could not afford to live in the Boston area, where she had grown up. To which the co-author responded, "You wouldn't be angry that your daughter couldn't afford to buy a Mercedes. Yet you're upset that she can't afford to live in Boston? There's no difference." One hears similar complaints about the cost of living in highly-valued suburbs within many MSAs.

1. Patterns of long-run house price growth

There are two outstanding stylized facts about house price growth in the postwar period. First, real appreciation has been highly skewed, with just a handful of cities experiencing outsized average house price growth. Second, the real increase in house prices for the top set of cities has been remarkably high. It is noteworthy that these patterns are not driven by recent rises in house prices. Those cities that have experienced relatively high price growth tend to have done so consistently over many decades.

To document this long-run growth, we turn to average house price data from the decennial censuses. We will describe these data in more detail in section 3, but for now we are using decadal data from 1940, and 1960-2000. The 1940 data is constructed differently from the rest, so below we will often focus on the 1960-2000 period, though our conclusions are similar either way. We adopt the Metropolitan Statistical Area (MSA) as our primary unit of geography, using a 1990 definition that is county-based, since we can measure house prices at the county level consistently over time.[3] We work at the metropolitan area level because that reflects the size of the labor market area in which key location decisions tend to be made.[4] Later in the paper, we provide complementary evidence on the evolution of income and price growth for places within MSAs.

The top-performing MSAs between 1960 and 2000 are all located in the San Francisco Bay area, and they dominate the competition. As shown in the left panel of Table 1 – which reports real annual house price growth for the top-ranked 15 MSAs – San Jose, San Francisco, and Oakland had average annual real house price growth rates of 3.96, 3.84, and 3.16 percent,

[3] In some areas, we use Primary Metropolitan Statistical Areas (PMSAs), which are components of Consolidated Metropolitan Statistical Areas (CMSAs). For ease of language, in the text we refer to the PMSAs as MSAs.
[4] MSAs are also where Americans live. According to Rosenthal and Strange (2004), 75 percent of Americans live in urban places constituting 2 percent of the land in the lower 48 states.

respectively.[5] The next closest metropolitan areas, Seattle and San Diego, grew their real house prices at annual rates of 2.78 and 2.71 percent, respectively. These annual differences compound to very large differences over time. For example, San Jose's 3.96 percent annual house price appreciation implies a 372 percent increase in real dollars over the 40 year period. In the fifth-ranked city of San Diego, real house prices rose 'just' by 191 percent. The remaining areas on the top-15 list, including Los Angeles, Boston, and the New York area[6], all experienced real average house price growth rates in excess of 2 percent, but the skewness in the overall distribution flattens out quickly. Still, these areas performed quite well, as house prices across all metropolitan areas in the country grew on average at an annual real rate of 1.77 percent, or 102 percent growth over the last 40 years.

A similar pattern can be seen in the right or second panel of Table 1, which reports average annual real house price growth rate over the 1940-2000 period. While the sample of metropolitan areas is smaller (at 129 areas) because the 1940 price data are obtained from a different source, the very top metropolitan areas are virtually the same, although there are some smaller, primarily southern areas that clearly experienced rapid house price growth between 1940 and 1960.[7] We suspect that the quality of their housing stocks increased dramatically right after the Second World War.[8] The more important point of this part of Table 1 is that the

---

[5] The MSAs are ranked by their real annual house price growth rate. For this table, we restrict our attention to the 129 identifiable areas with at least 170,000 residents in 1940.

[6] The New York PMSA, which includes New York City, is excluded from this list because the Census did not report report owner-occupied house price information for it in 1960. See footnote 15 for more on that. Its neighboring areas such as Trenton and Nassau-Suffolk did experience relatively high house price growth, as we expected would be the case.

[7] Mean house prices for 1940 come from the Integrated Public Use Micro Samples (IPUMS) files maintained at the University of Minnesota. The 1940 census file only identifies 129 metropolitan areas, and certain places such as Oakland, cannot be disaggregated from the San Francisco area in that year. See the discussion in the data section for those details.

[8] That suspicion cannot be confirmed by constructing a constant quality price series for that time period. The first Census of Housing was conducted in 1940, but all information on house traits was lost. The IPUMs only contains price data for that year. The IPUMS reports no house price data for 1950.

extremely high price real price appreciation in the Bay Area, Seattle, and San Diego is not restricted to the last one or two decades associated with the high tech boom.[9]

These data also imply increasingly large gaps across MSAs in real average house prices, which are graphed over time for five places in Figure 1. In Austin, which was the tenth-ranked MSA from 1960 to 2000, average real house prices (in 2000 dollars) rose from nearly $40,000 in 1940 to $164,000 in 2000. In contrast, San Francisco started at $60,162 in 1940 but by 2000 the average house price was almost $550,000. In relative terms, the mean house in San Francisco cost 1.5 times that in Austin in 1940, but 3.4 times Austin in 2000. The Boston and New York MSAs are in between these two. Naturally, the differences are even larger when one considers metropolitan areas outside the top 15. In 1940, Cincinnati's house prices were above San Francisco's, at nearly $65,000. But by 2000, they had risen just 136 percent, to $145,000.

House price growth, too, is highly skewed. Figure 2 plots the distribution of average real annual house price growth (computed over the 1940-2000 period) for those MSAs with population of at least 170,000 in 1940. While there is some heterogeneity across MSAs in their appreciation rates, the slope of the distribution is pretty linear until the two percent annual price growth threshold, at which point the differences in average growth rates start to accelerate.

The list of MSAs that exhibit above-average growth is surprisingly persistent over the very long term. Of course, over short horizons, MSAs experience significant price swings. In fact, the correlation in house price growth rates across contiguous decades is negative, at -0.40. But over 30-year periods, from 1940-1970 and 1970-2000, average annual percentage house price growth demonstrates very high and positive correlation, at +0.40. The root of this latter result can be seen in Table 2, which reports the transition matrix for MSAs ranked by their

[9] Of course, the Internet boom in the Bay area was really a 1990s phenomenon, yet house price growth in the 1980s exceeded the growth rate of the 1990s.

average real house price growth rates computed over the two 30-year periods of 1940-1970 and 1970-2000. Most high appreciation areas do not move very far in their relative price growth ranking. For example, of the 32 MSAs in the top quartile of annual house price growth between 1940 and 1970, half were still in the top quartile and nearly two-thirds remained ranked in the top half between 1970-2000. Outside of the top growth rate areas, there is more movement across the distribution.

2. A superstar city model that is consistent with this pattern of house price growth

To match the stylized pattern of house price growth, we need a theory that can rationalize both the fact that some MSAs have very high rates of house price appreciation, and that some people are willing to pay an ever-increasing premium to live in those places. The theory should also reflect that these high-growth places are consistently at the top of the rankings. We would also like to develop the theory without using agglomeration or other production or consumption externalities, although such factors would likely amplify the important predictions of our model.

In this section, we show that if some MSAs have an inelastic supply of places in which to live (i.e., they are capacity constrained either due to geography or political regulation), then they can exhibit very rapid house price growth rates and a widening gap between their house prices and those of other cities.[10] These MSAs need not be a universally preferable place to live. All that is required is that some people prefer to live in these areas. Once the MSA becomes full,

[10] Other work has investigated the relationship between the inelasticity of supply and house prices, but have not tied it to the long-run evolution of house prices. Nor have they examined the underlying process that causes some cities to have extraordinarily rapid price growth. Glaeser, Gyourko, and Saks (2004) find that house prices in certain markets exceed construction costs and argue that it is due to regulation that restricts building. Saks (2003) documents that MSAs with lower estimated supply elasticities experienced bigger increases in house values over the 1980-2000 period.. Mayer and Somerville (2000) find that supply restrictions, such as the time it takes to obtain a permit, leads to a lower price elasticity of supply. Looking at declining rather than growing cities, Glaeser and Gyourko (2004) note the durability of housing means that cities suffering negative demand shocks will not quickly shrink, so that the bulk of the impact gets reflected in falling prices, not in fewer people or housing units.

rich people crowd out the poor by bidding up the cost of land, which is the price of admission. With population growth or a widening of the income distribution, the number of rich people who prefer a given location goes up in absolute terms, but if the MSA is inelastically supplied, it cannot increase its absolute capacity. Over time, richer people crowd out the less rich and rents rise as the composition of the MSA changes and the average wage of the existing residents goes up.

Of course, some people prefer to live in MSAs that happen to have lots of room to grow. These elastically supplied locations cannot support any price growth because as demand rises, they create as many new spaces as they need. There is never an issue of excluding people in these metropolitan areas. Even as the population of the country grows, anyone who prefers to live in one of these MSAs can do so. Thus, they experience growth in population, rather than in house prices. If a family wants to live in one of these areas, it does not have to sacrifice anything (relatively speaking) to do so. However, the family must give up non-housing consumption to be able to live in one of the inelastically-supplied areas.

It is important to note that for a MSA to be inelastically supplied there must not be a close substitute on the dimensions that the families who prefer to live there care about.[11] Otherwise, the difference in prices would be capped by the difference in utility between the first- and second-best options. Another way to think of this is that for a MSA to be inelastically supplied, it is not sufficient merely for the existing MSA to be full of families and unable to expand. Rather, no other existing MSA can be similar enough that families would be indifferent

[11] For whatever reason, this appears to be an accurate description of some cities. Davis and Weinstein (2001, 2004) find that cities tend to recover their relative population and basic industrial mix even after they were obliterated by a nuclear bomb. This suggests that there may be something fundamentally attractive about these cities. Glaeser, Kolko, and Saiz (2001) argue that the fact that cities with good weather now have faster growth is a sign that consumption benefits matter. That kind of amenity, as well as aesthetics, may not be easily replicated.

between the two, and the MSA must not be replicable. If one of those conditions did not hold, the MSA effectively would be in elastic supply.

This view of the evolution of cities is not one where economic value needs to be created to generate higher land rents. Rather, it is a pure sorting outcome, where inelastically supplied cities are scarce factors and the rich who prefer that location can outbid poor people with an equal taste for that MSA for the limited number of spaces. If value were created, too, it could be capitalized into the land rents as well. For example, the initial sorting by income could generate positive externalities, which raises prices even more via a feedback effect. If rich people prefer to live with other rich, either because they like to exclude the poor for fiscal or peer effect reasons (e.g., Mills and Oates (1975); Bayer, et. al. (2004); Rothstein (2003)) or because a greater density of people with similar preferences improves the number and quality of consumption opportunities that match their needs (e.g., Glaeser, Kolko and Saiz (2001); Waldfogel (2003)), rich people may pay a premium to live in the area. An uneven distribution of preferences for a city could even accentuate the pattern further--if rich people preferred inelastically supplied MSAs, for example. However, these additional factors enhance the basic effects, and are not necessary to generate them.

This model of city house price growth generates some empirically testable predictions. For one, rich people should disproportionately sort into the inelastically supplied cities. Since their marginal value of consumption is lower, they are willing to pay more to live in their preferred city than someone with less money but equal desire for the location. As the right tail of the income distribution gets thicker, the superstar (high-demand, inelastically supplied) cities should gain a disproportionate share of the rich, and their house prices should rise. Since those cities cannot have much population growth, the new rich should crowd out the existing poor. In the elastically-supplied cities, as the national population grows, we should observe a growing number of rich and poor, since both rich and poor people who prefer to live there can do so. But these cities should become relatively poorer, as the poor who are unwilling to pay the entrance fee for the superstar cities must also live there.

We now formalize this model.

Labor Market:

A continuum of worker types is presumed, ranging from very low skilled to extremely high skilled. The higher the skill set, the higher the worker's wage. We assume that wages are given exogenously and are independent of a worker's location. Define n(w) as the distribution of workers who earn wage w and N as the total number of workers in the economy. Thus $\int_{w=0}^{\infty} n(w)dw = N$. A hypothetical wage distribution that will be used in the simulation analysis reported below is shown in Figure 3.

Urban Setting:

We assume the there are two cities: a red city (R) and a green city (G). The red city has stopped any future development and limits supply to K housing units. The green city has a perfectly elastic supply of housing units.

Preferences:

All workers draw a preference representing the match quality a worker receives from living in the red versus green city. The preference (c_i) is uniformly distributed between 0 and 1 (i.e., $c_i \sim u[0,1]$), where a preference of 0 means that worker i strongly prefers to live in the red city, whereas a preference of 1 means that worker i prefers the green city. Note that the preference to live in R versus G is independent of income, so all types of workers are equally likely to prefer each city.

Goods and Prices:

Workers choose three items in their consumption bundle: a variable quantity of housing (H), a city to live in (R or G), and a quantity of the composite good (A—all other goods). The price of A is normalized to 1. Housing is produced in a competitive market at a per-period rental cost of p_h per unit of H. Workers choose to live in city R or G based on the relative price of R or G. By assumption, the rental cost of land in G is 0 (land is perfectly elastically supplied in the green city). If demand is high enough, workers will pay rent $r>0$ to live in city R, with r determined based on bidding as described below. For the moment, r is defined as a per-period rent to use land in city R.

Utility/Value Function:

We assume that workers have Cobb-Douglas utility functions with parameter h to measure the relative preference for housing and all other goods. In this model, housing is a proxy for nontraded goods, while A is a proxy for traded goods. We recognize that this distinction is difficult to make in practice. For example, lumber, appliances, and cars are traded goods, but the land inputs in those goods are nontraded. Similarly, a housing structure is a mixture of traded and nontraded goods.[12] In addition, by pre-multiplying by c (or 1-c), we assume that households will get higher utility from driving a car or drinking a latte when they are in their preferred city. Thus,

[12] However, this model makes the unambiguous prediction that households in the more expensive city will spend a higher proportion of their income on housing when we measure housing to include structure and land.

$$U = \begin{cases} U^R \equiv (1\text{-c})(H^R)^h\,(A^R)^{(1\text{-h})} & \text{if worker resides in city R} \\ U^G \equiv c(H^G)^h\,(A^G)^{(1\text{-h})} & \text{if worker resides in city G.} \end{cases}$$

Solution:

Demand by workers:

A worker with wage w and preference parameter c will be indifferent between living in R versus G if and only if:

(0.1) value of living in R = value of living in G

Workers are willing to pay a rent $r \leq \max\{0, U^R - U^G\}$ to live in city R. Workers with $c \geq ½$ will always live in G. We assume that $N/2>K$, so that rents in the red city will be strictly positive (i.e., $r>0$).

With Cobb-Douglas preferences, consumers spend the fraction h of their income after rent on housing and (1-h) on all other goods (A). In city R, a worker's remaining income after rent is (w-r), whereas in city G workers can spend their entire wage on housing and the composite good. At current prices, workers' optimal consumption in each city and for each good is:

	City G	City R
Housing	Hw/p_h	$h(w-r)/p_h$
Composite good	(1-h)(w)	(1-h)(w-r)

Equilibrium:

Given the demand computed above, we derive rent r^* and corresponding cutoff value $c^*(w)$ such that workers at a given wage w and $c < c^*$ will choose to live in R and pay rent r^*. In equilibrium, rent will adjust so that the number of workers desiring to live in R is less than or equal to the number of lots in R as in:

(0.2) $$\int_{w=0}^{\infty} c^*(w) * n(w) dw \leq K.$$

Define $V^*(i)$ as optimal consumption of A and H, given income i. We also know that $U^R = (1-c)V(w-r)$ and $U^G = cV(w)$. We can solve for $c^*(r)$ in two parts. When $w \leq r$, then $c^*(w) = 0$; that is, workers cannot afford to live in R unless their wage is greater than the rent required to live in R. When $w > r$, we derive $c^*(w)$ as the solution to

(0.3) $$U^R - U^G = 0.$$

Substituting for *U* with *V*( ) when $w > r$, we obtain:

(0.4) $$(1-c^*)V(w-r) - c^*V(w) = 0,$$

or

(0.5) $$c^*(w) = \frac{V(w-r)}{V(w-r)+V(w)}.$$

Result 1: The average wage/skill level will be higher in the red city, even without explicit preferences that benefit R or any productivity differences across the cities.

Explanation for Result 1:

With strictly positive rents, some low wage workers will be unable to live in R because $r \geq w$. In fact, living in R can be considered a scarce luxury good and thus is more likely to be consumed by wealthy workers. In addition, notice that there will always be some high wage workers who live in the green city because they are indifferent or prefer that location (i.e., when their $c \geq ½$). Following equation (0.5), with finite wages ($w < \infty$) and positive rents, the cutoff value $c^*(r) < ½$. That is, even for very high wage workers, a modest rent is sufficient to cause a small number of workers who slightly prefer R to live in G.

Simulation Results:

To provide a sense of how this model works, we conduct a simple simulation exercise for an economy with 100,000 workers, K=30,000 for the number of spaces in city R, a truncated normal distribution for wages, a mean wage of $40,000 with a standard deviation also of $40,000, w_L= 0 (the lowest wage), and w_U= $150,000 (the top wage), solved using discrete $1,000 increments, P_H=1, and h=0.4. The average wage in the economy is actually $51,246 given that the lower truncation wage of $0 removes more of the low wage workers relative to the upper truncation wage of $150,000. The optimal rent in this example, following equation (0.5), is a relatively high $19,257, or almost 40 percent of the average wage. The lowest wage workers will find city R unaffordable, and even for workers earning the average wage, relatively few will choose to live in R. Thus the distribution of incomes in R will be well above the average for the economy as a whole. The mean income in R is about $66,000 versus $45,000 in G.

Figure 3 plots the assumed distribution of wages, while Figures 4 and 5 demonstrate the equilibrium for this model. Figure 4 plots the cutoff values for c*(w). Naturally, there is a discrete jump at the optimal rent level of $19,257. As suggested above, the cutoff value increases rapidly before flattening out, as a modest preference for the red city tends not to counterbalance the high rent required to live there. Figure 5 graphs the wage distributions in the red and green cities, with the number of residents in the green city being reflected by the difference between the two solid lines. Nobody earning below $19,257 lives in the red city, but a disproportionate share of higher income workers live in the red city.

Extension #1: Growth

We now extend the model to two periods.

Growth in Real Wages:

First, consider the possibility of growth in real wages of g percent. If real wages grow for all workers at the same rate, the results are trivial. Following equation (0.5), the rent to live in R versus G will increase by exactly one plus the growth rate, (1+g). Everything else will be symmetric.

Growth in Population:

Another possibility to consider is that population grows at rate g, but that the distribution of income remains the same. Clearly, c^* must fall, but the equilibrium in the new model does not just involve a proportional g% decrease in c^*. To see this more clearly, consider equation (0.5) again. In order for c^* to fall, rent r must rise. Thus, some households who previously had a strong preference to live in R will no longer have the income necessary to cover the rent r. With growth in population, households with the lowest incomes are disproportionately priced out of city R.

Figure 6 depicts the new 'national' wage distribution assuming that population doubles from 100,000 to 200,000. The equilibrium cutoff values for $c^*(w)$ shift to the right and flatten, as shown in Figure 7. Figure 8 illustrates how the wage distribution in the Red city shifts to the right. The minimum wage to live in Red rises from less than \$20,000 to almost \$45,000. The equilibrium rent increases by more than one hundred percent from \$19,257 to \$41,263.[13] All former residents with wages between the previous minimum and the new one move from Red to

[13] The specific percentage change in rent, however, is a dependent on the particular values used for this simulation. Eventually, as population grows large enough in this simulation, the cutoff rent must grow at a slower rate than the growth in population because the upper support of the truncated normal distribution does not change with population growth.

Green, and a large fraction of those with incomes between $45,000 and $60,000 do as well. Across the rest of the distribution, the number of people in Red rises, but disproportionately so at the high end of the income distribution. Thus, population growth leads to a more skewed distribution of wages for workers living in R relative to G, even without a change in the overall distribution of incomes.

Fatter Tails in the Income Distribution:

Another interesting possibility is that the variance in the income distribution rises relative to the mean, leading to fatter tails of relatively high or low wage households. Clearly, the only change that matters for the distribution of wages in R versus G is in the right tail of the distribution; more specifically, an increase in the number of workers with wages above r. To get at this issue, we solve for a solution to our simulation by doubling the standard deviation of wages from $40,000 to $80,000, as shown in Figure 9. The new equilibrium rent is just below $24,000, for an increase of 25 percent. The intuition for this result is provided by Figure 10 which plots the change in c^*. Using the c^* levels from the base model, total demand for R would grow from 30,000 to 33,000 workers. Thus, only a moderate increase in the rent level is required to equilibrate supply and demand for R. Even with a relatively modest change in c^*, the Red city has a greater density of high income families, as shown in Figure 11. Both cities have an increase in the number of high-wage workers (also see Figure 12). However, with more high-wage workers in the economy, the R city attracts a slightly *higher* percentage of them relative to the previous equilibrium.

Increase in the right support of the distribution of wages:

As is probably clear by now, an increase in the right support for the distribution of wages has very little impact on the overall equilibrium. Such an increase slightly raises the number of workers with wages above r, leading to a very small increase in the equilibrium rent.

An Alternative Model: Tournaments and Correlated Preferences

Some may argue that the correct model is not one in which preferences for R versus G are randomly drawn, but instead that most (or all) workers have a preference to live in R versus G. According to this perspective, what differentiates San Francisco from Las Vegas is not the supply constraints in San Francisco *per se*, but that workers systematically prefer to live in San Francisco. Thus, high-wage workers end-up in San Francisco because they are the only workers that can afford to live there.

We consider this case by examining the implications of this model relative to our base case. If all workers prefer R to G, then only the wealthiest workers will be able to afford R. No poor workers will live in R. Instead of the observation that a cutoff preference is increasing with the wage, in the extreme version of this alternative model, there is a marginal wage earner. Anyone earning less than the marginal wage will live in G and anyone earning more than the marginal wage lives in R. This leads to pure sorting by incomes, as depicted in Figures 13, 14 , and 15. As we will see later in the paper, a pure sorting model does not fit the data very well. Of course, intermediate cases are possible, so that all workers have a moderate preference for R versus G, but that their relative strength of preference differs across individuals and is uncorrelated with income.

3. More evidence on Superstar Cities

One of the important implications of the superstar cities model, beyond the stylized facts reported above, is that high-demand, inelastically-supplied cities should draw a relatively large proportion of their populations from the high end of the national income distribution. Then, as the population grows, or as the right tail of the national income distribution gets thicker, superstar cities should experience an increase in their average incomes and house prices. The composition of households in a superstar city should shift to the right in the income distribution, with new rich crowding out the poor. In addition, a disproportionate share of the growth in rich households should go to superstar cities.

In examining these predictions, we try to distinguish between superstar cities and agglomeration-based theories. Given that any production-based externality is tied to a specific location, an agglomeration model predicts that in-migrants to San Francisco will become more productive once there. In contrast, superstar cities predicts that more productive workers will move to San Francisco, since higher-wage workers are more likely to go to scarce locations.[14] It is always possible that higher-wage workers are also the ones who benefit from the production externalities, so this test is suggestive rather than definitive.

3.1 Data description

Virtually all of the data used in our analysis comes directly or indirectly from the decennial censuses of the United States. Observations on the metropolitan U.S. were collected for population, the number of housing units, educational achievement, as well as the distributions of house values and family incomes (described more fully below), and geographical identifiers. We obtain this data from the *Census of Population and Housing* from 1960-2000, which is based on 100 percent population counts. More specifically, the variables for 1970-2000 were taken

[14] Consumer cities theories also predict that high-productivity workers will move to more attractive places.

from CD-ROMs prepared by Geolytics, Inc. Information for 1960 was hand collected from hard copy volumes of the 1960 *Census of Population and Housing*.

Information for 1940 is more difficult to obtain. We collected population and housing unit data based on 100 percent counts, but housing values for that year are averages from the 1940 sample provided by the Integrated Public Use Micro Samples (IPUMs) housed at the University of Minnesota. We do not yet use any family income data for 1940, but some is available for that year and will be examined in future drafts.

Much of our analysis is done at the metropolitan area level—that is, at the level of the labor market. The 1990 county-based definitions of metropolitan areas provided by the Office of Management and Budget are used, with county-level data aggregated to the metropolitan statistical area (MSA) or primary metropolitan statistical area (PMSA) level in the case of consolidated metropolitan areas (CMSAs). For the five censuses from 1960-2000, we have data for all 845 metropolitan counties in the lower 48 states, which then are aggregated into 316 MSAs or PMSAs.[15] For 1940, our sample is restricted to 129 metropolitan areas listed in the 1940 IPUMs.[16]

We also perform some household- and person-level analysis. Whenever micro data are used, the IPUMs is the source. These data provide a set of matched variable definitions across decades. All micro-level data on migration to and from MSAs is also from the IPUMs.

House values and family income play an especially prominent role in our analysis and we use information on their distributions from different census years. However, top-coding of the

[15] The New York PMSA is missing crucial house price data for 1960, and, therefore, is excluded from some of the empirical analysis reported below. Because of the preponderance of cooperative units in Manhattan for that year, the census did not report any value data because it did not think it could accurately assess owner-occupied value for that ownership structure.

[16] The IPUMs is the only source for house values for that year, hence the restriction. It also should be noted that we cannot be sure that the IPUMs metro definitions for 1940 precisely match those created for the other years. Hence, there is likely to be some measurement error when using those data in conjunction with information from later censuses.

data presents potential problems in comparing Census data across years. In order to represent the distributions in a parsimonious manner, five categories (or bins) are created in each year based on the topcoded value of the house value or family income variable in a given year. For house values, we create separate standardized bin distributions based on the 1960 and 1980 distributional information provided by the census. For family income, we work with the 1960 and 1970 distributional data in the census. As will become more apparent in the empirical presentation below, working with these particular breakdowns of the data allows us to see interesting variation in income and value distributions across metropolitan areas and over time.

We define the boundaries of the five bins based on the real ($2000) topcode in one decade, with the lowest bin in any year containing real values between zero and 25 percent that topcode figure. The next bin covers 25-50 percent, then 50-75 percent, 75-100 percent, and 100 percent or greater.[17] We then work with the actual house value or family income bins reported in each census year and convert their cutoff points into constant 2000 dollars ($2000).[18] Assuming a uniform distribution within the census bins, our five standardized bins then are populated using the weighted average of the actual bins.

3.2 How do we determine if a city is a superstar?

In order to make comparisons between MSAs, we need a way to identify which ones are "superstars." Section 2 provides some guidance. Superstar cities, in the face of rising demand, experience high house price growth and low growth in the number of housing units because they

---

[17] Values can be greater than the topcode if the topcode in a given decade is higher than the topcode we applied. This often is the case since we choose from the lowest topcodes in order to make the data more comparable.

[18] Ultimately, this procedure results in the following breakdown of bins, with all values in $2000. Family Income based on the 1960 Top Code: <$14,553, $14,553-$29,107, $29,108-$43,661, $43,662-58,215, >58,215; Family Income Based on the 1980 Top Code: <$39,179, $39,179-$78,358, $78,359-$117,537, $117-538-$156,716, >$156,716; House Value Based on the 1960 Top Code: <$36,383, $36,383-$72,767, $72,768-$109,152, $109,153-$145,536, >$145,536; House Value Based on the 1970 Top Code: <$55,434, $55,434-$110,868, $110,869-$166,302, $166,303-$221,736, >$221,736.

are inelastically supplied. Other MSAs with increases in demand may experience growth in population or the number of houses but not in price. While those are vibrant, growing areas, they do not meet our definition for being a superstar because there is no material entry fee required to reside there. Of course, MSAs with little upswing in demand have neither housing unit nor house price growth.

Since we do not directly observe either the demand for cities or their elasticities of supply, determining which are the typically high-demand, low-elasticity MSAs is not immediately obvious. However, if one plots MSAs along both house price growth and housing unit growth dimensions, high-demand MSAs (both elastictically and inelastically supplied) should form a frontier in the top right quadrant. Figure 16 graphs this relationship. For each MSA in our sample, we compute the average annual real house price growth between 1960 and 2000, and the comparable growth rate for housing units. Within the (arbitrarily drawn) frontier between the two green lines, MSAs have relatively higher demand. Some, like San Francisco and Santa Cruz, have very high house price growth but low unit growth. Some, like Las Vegas and Fort Meyers, experience large percentage growth in units, but not in prices. Places like Phoenix are in-between. MSAs closer to the origin presumably have low demand, since (relatively speaking) they are gaining neither house value nor population.

In this graph, San Francisco or San Jose would be superstar cities, as they have high long-run price growth. Las Vegas or even Phoenix would not because, even though they demonstrate high demand through unit growth, they have elastic supply and thus much lower price growth.

In the empirical work, below, we will identify superstar cities as those with higher long-run house price growth, measured over the 1960-2000 period.[19] High demand cities that are not superstars are presumed to have higher housing unit growth between 1960 and 2000. In the regression analysis we experiment with two functional forms of the long-run growth rates, letting them enter linearly or using an indicator variable for the MSA being in the top quartile of growth.

3.3 Evidence from income distributions

One of the predictions of the model in Section 2 is that the across- and within-MSA income distributions should vary with the degree to which a city is a superstar. A superstar city's population should be skewed to the right of the US income distribution, and, if the right tail of the national income distribution gets thicker, the right-ward shift for superstar cities should be more pronounced. Analogously, as the number of the nation's rich increases, the superstars should also get a disproportionate *number* of the new rich.

The right tail of the national income distribution has indeed been getting thicker over time, which could explain the growth in house prices in inelastically supplied MSAs according to our model. Figure 17 uses data from Saez (2004) to graph the share of U.S. income by population percentile over time. The tax return data Saez (2004) uses provides a very good picture of changes at the high end of the income distribution. The share of income held by the very top fractiles of the U.S. population – the top one-hundredth or 0.01 percentile, the 0.1 to 0.01 percentile, and the 0.5 to 0.1 percentile – all increased dramatically over the last 40 years. The income share of the top 1% grew from under 10% in 1960 to almost 17% in 2000. Even the

[19] We use the 1960-2000 period so as to include as many MSAs as possible and to mitigate the issue of unobserved increases in house quality in southern cities. However, we obtain similar results when we use growth rates computed over the 1940-2000 period.

share of income held by the first to tenth percentiles of the population went up, from about 23 percent in 1960 to about 27 percent in 2000.

When we try to replicate the Saez (2004) results using our Census data, we see a very similar pattern, although with less detail on the top end due to the topcoding of incomes. The left panel of Figure 18 shows the distribution of incomes obtained by adding up all households in metropolitan areas. The top chart uses bins defined relative to the 1960 topcode amount of $58,215 in 2000 dollars. Since 1960 had the lowest topcode of our entire sample period, this definition has a lot of granularity at the bottom end of the income distribution and every year is defined. The bottom chart uses the 1980 topcode. In that chart, the 1960 data populate only the bottom two bins, since it is topcoded below $78,358. But the lowest bin is still measured properly. Whichever chart one examines, it is clear that the nation's income distribution has been shifting to the right in real dollars.

That shift in the income distribution has occurred because the right tail of the income distribution has grown, not simply because the income distribution has experienced a mean-shift. The right panel of Figure 18 displays the evolution of the number of people across all MSAs in each of the income bins. In the top chart, the population in each income category is fairly constant over the 1960-2000 period, except for those making in excess of $58,215 in 2000 dollars (i.e., those in the top bin). The number in that category grew tremendously, more than tripling from about 7 million to over 25 million households. In the bottom chart, we can see that most of that growth occurred for people making above $78,358.

When we compare these national changes in the income distribution with those for San Francisco (our canonical superstar city and representative of the red city in the model) in Figure 19 and Las Vegas (a high-demand, elastically supplied MSA that is like the green city) in Figure

20, we see that the predicted relationships hold for this pair of MSAs. The first thing to notice is that the cross-sectional prediction that a superstar city should draw more from the right portion of the income distribution is borne out in the data. In any decade, San Francisco is relatively more rich and relatively less poor.

What is more interesting is what happens over time as the national income distribution shifts right. The share rich in San Francisco (left panel of Figure 19) increases and the share poor declines, both much more than for the nation as a whole. In fact, only the richest groups – those with incomes of $78,358 and above – increased their share of the population in the San Francisco MSA. In contrast, the left panel of Figure 20 shows that Las Vegas had little change in its income distribution over the 1960-2000 time period.[20] While it became slightly more rich, it did so much less than the MSA average.

The right panels of Figures 19 and 20 show why. San Francisco, though clearly in demand, only slightly expanded its population between 1960 and 2000. Relatively rich households entered the MSA and displaced lower income ones.[21] In contrast, Las Vegas experienced explosive population growth. It grew from fewer than 50,000 households in 1960 to the size of San Francisco by 2000. But it added households across the income distribution, rich and poor, proportionally to the distribution in 1940. Relative to the national income distribution, which shifted right, the growth in Las Vegas has been skewed towards the relatively poor.

3.3.1 Within-MSA income distribution changes

[20] Recall that the top three income categories are encompassed in the second bin for 1960 in the bottom panel.
[21] We will reject the possibility below that this is solely because incomes in San Francisco grew a lot across the entire distribution.

While the stark differences between San Francisco and Las Vegas are informative, they are potentially extreme examples. To see if this pattern holds across all MSAs, we being by estimating the following regression for MSA *i* in year *t*:

$$\frac{\#\text{ of Rich}_{\text{it}}}{\#\text{ of Households}_{\text{it}}} = \beta(\text{Price Growth}_{\text{i}}) + \gamma(\text{Unit Growth}_{\text{i}}) + \delta_{\text{t}} + \varepsilon_{it}$$

This equation relates the share of an MSA's households that are rich to the degree it is a superstar, as measured by its average house price growth over the past 40 years, and controlling for whether it is a high-growth MSA. We measure the share rich as the proportion of households in the top income bin, using the 1980 topcode-based definition. We also repeat the estimation using the share poor, defined as the households in the bottom bin divided by total households in the MSA.[22]

We find that the income composition of MSAs as a whole matches the model quite well. In the first column of the top panel of Table 4, MSAs with higher long run average price growth have a greater share of their households that are rich. A one standard deviation increase in the average annual house price growth rate of 0.66 is associated with a 0.73 percentage point increase in the share of the MSA that is rich. Compared to the average MSA share rich of 2.7 percent (see Table 3), this is a 27 percent increase. There is no statistically significant relationship between the 40-year average growth in housing units and the average share rich over the period. We obtain similar results when using indicator variables for the top quartiles of price and income growth instead of letting these variables enter linearly. MSAs in the top quartile of average price growth have a 1.6 percentage point greater share of their households that are rich.

Conversely, superstar cities also have a much smaller proportion of the poor. In the first column of the bottom panel, the estimated relationship between price growth and the poor share

[22] Summary statistics for a host of variables used throughout our analysis are reported in Table 3.

of the MSAs households is -4.84 (with a standard error of 0.40), corresponding to about a 3.2 percentage point decline in the share poor when the price growth rate is one standard deviation higher. MSAs with higher growth in the number of units do get a relatively larger share of the poor. A one standard deviation increase in the 40-year average growth rate of housing units (1.14) is associated with about a 2 point increase in the share poor. Using a dummy variable for being in the top quartile of price growth, superstar cities have about a 7 point decline in the share poor while cities in the top quartile of unit growth are associated with a 4 point rise in the share poor.

Over time, when the national population was growing, especially at the high-income end of the distribution, these relationships became more pronounced, with superstar cities growing more rich and less poor. In the rich share of the MSA regressions reported in the top panel, the coefficient on long run price growth was 0.51 (0.09) in 1970. By 1990, it had risen to 1.36 (0.13), and then to 2.27 (0.18) in 2000. Similarly, MSAs with higher average price growth had an ever-smaller share poor between 1970 and 2000, with the estimated coefficient dropping from -3.92 (0.93) to -6.07 (0.69) in 2000.

3.3.2. The relationship between MSA income composition and the national distribution

The model in section 2 makes predictions not only about how the income distributions of superstar and non-superstar MSAs should appear, but also how they should evolve in response to changes in the national income distribution. In this subsection, we examine whether superstar cities get more than their proportionate share of any increase in the nation's rich, as would be predicted by the theory.

We estimate the following regression to see if, when the right tail of the national income distribution gets thicker, the population make-up of superstar cities becomes relatively 'more rich':

$$\frac{\Delta \text{Rich}_{\text{it}}}{\#\text{ of Families}_{\text{it}}} = \begin{matrix} \alpha + \beta\left(\dfrac{\Delta \text{Rich}_{\text{all MSAs,t}}}{\#\text{ of Families}_{\text{all MSAs,t}}} \times \text{Price growth}_{\text{it}}\right) \\ +\gamma\left(\dfrac{\Delta \text{Rich}_{\text{all MSAs,t}}}{\#\text{ of Families}_{\text{all MSAs,t}}} \times \text{Unit growth}_{\text{it}}\right) + \delta_i + \mu_t + \varepsilon_{it} \end{matrix}$$

The left-hand-side variable is the change in the number of rich, normalized by the beginning-of-period number of families to account for differences in scale among MSAs. The regressors are the change in the number rich across all metropolitan areas divided by the total number of families in those areas, interacted with both the price growth variable and the housing unit growth variable used above. The equation is estimated on the pooled 1970 to 2000 data and the specification includes year and MSA dummies. The metro dummies pick-up any direct effects of price growth and unit growth on the proportional change in the number of rich, as these variables are constant over time within a MSA.

As predicted, superstar cities become disproportionately richer, this time in terms of flows rather than just at a point in time. In the top-left panel of Table 5, when the number of rich in all the MSAs in total goes up, superstar MSAs obtain an above-average share of those rich. This effect is increasing in the degree of long-run price appreciation. The relationship between high housing unit growth and the change in the proportion of rich is much smaller and opposite in sign. Thus, fast growing places in terms of housing units gain a smaller-than-average share of the national increase in rich, controlling for house price appreciation.

This pattern is not sensitive to our measure of the change in the MSA's income distribution relative to that for the nation. For example, column 2 reports a similar positive relationship when we relate the change in the share of the MSA's population that is rich to the change in the share of the all-metro's population that is rich. Finally, column 3 shows that superstar cities even get a disproportionate share of the increase in the *number* of rich in the all-MSA group. This result is striking because superstar cities, by definition, have very inelastic supply. Column (3) shows that the composition effect of rich displacing poor is large enough to overcome the fact that superstar cities have an ever-shrinking proportion of overall housing units.

3.3.3. Are superstar MSAs filling up over time?

The results so far suggest that the estimated effects of being a superstar city have been increasing over time, which we have attributed to an upward trend in the number and national share of the rich. However, it is also possible that as the national population grows, the superstar cities are running out of space (for whatever reason), and the inelasticity of supply is becoming more binding.

To see if the data is consistent with this hypothesis, we examine whether there is a nonlinear relationship between price and unit growth. We regress the decadal growth in housing units on the decadal growth in house prices, by MSA, allowing a different slope and intercept for those MSAs in the top quartile, ranked by house price growth. The regression estimated is:

$$\%\Delta(\#\text{ Units}) = \alpha + \beta(\%\Delta P_{it}) + \gamma(\text{Top quartile }\%\Delta P_{it}) + \theta(\%\Delta P_{it} \times \text{Top quartile }\%\Delta P_{it}) + \varepsilon_{it}$$

In essence, we identify the long run superstar cities and see if their growth in housing units declines over time. With elastic supply, high housing demand should lead to new housing units. With inelastic supply, it should generate higher prices. If an MSA becomes more inelastic, the positive relationship between prices and units should become negative or go away.

In table 6, we see that the positive relationship between price growth and unit growth has disappeared for the top quartile of the price growth distribution between 1970 and 2000. For the bottom 75 percent of the price growth distribution, the relationship between average price growth and unit growth is positive and, with the exception of 1980, flat over the decades. However, the MSAs in the top price growth quartile start in 1970 with a slightly less positive correlation than for the lower 75 percent (11.12-3.12 = 8.0). By 1980, the top quartile has, if anything, a slightly negative correlation between price growth and housing unit growth (17.18-18.14=-0.96). The bottom 75 percent has a very large positive relationship. That negative correlation for the top quartile increases over time, suggesting that the top MSAs have effectively run out of space (either for fundamental geographic reasons and/or by political design via regulation). This fact likely contributes to the temporal evolution of house prices in superstar cities.

3.4 Mobility and superstar cities

While we have documented that superstar cities seem to be a magnet for the rich, the pattern of house price growth and income sorting we have observed could conceivably be due to agglomeration benefits in certain MSAs raising the productivity of workers there. The resulting higher wages would then be capitalized into land prices in inelastically supplied places. While this alternative story strikes us as less likely than the superstars hypothesis – for agglomeration to be the sole explanation, a handful of inelastically supplied MSAs would have to enjoy production externalities that were increasing in the same MSAs over time, growing at very high rates, and accruing mainly to the high-income end of the spectrum – we turn to the data to try to sort out the two options.

One implication of the superstar cities theory is that highly skilled workers locate in superstar cities. The workers had high-income before they went to the MSA, so to the extent that an MSA appears productive, it is because skilled people migrated there. Conversely, production agglomerations imply that any given worker will be more productive, and thus earn a higher wage, when she works in an agglomerated MSA. With agglomeration, either through firm or human capital spillovers, a place makes people more productive.

Ideally, we would test this distinction by comparing workers' wages before and after exogenously moving to a superstar city. If the superstar city had a positive influence on wages over and above what the worker earned prior to moving, that would provide evidence in support of agglomeration benefits. But if the type of worker who moved to the city was already high-income so that its arrival changed the income distribution of the destination city, the pattern would be more consistent with superstar cities.

Since we cannot observe preexisting wages (or exogenous moves), we compare the distributions of in-migrants and out-migrants in superstar cities vs. non-superstars. To do so, we use the individual-level IPUMS data, which reports the MSA of residence five years prior. If the resident lived in a different MSA than the current one five years ago, we label him an in-migrant to the current MSA. We also say he is an out-migrant from the MSA he lived in five years ago. Matching this information with incomes reported in the IPUMS, we estimate the distribution of in-migrants and out-migrants across the five income bins. By comparing these distributions in superstar vs. non-superstar cities, we are in effect asking whether the share of in-migrants who are currently rich is higher in superstar cities.[23]

[23] That is, we look at the percent of in-migrants who are rich. We do not consider the percent of rich who are in-migrants to be very informative since elastically supplied locations have a high percentage of in-migrants – that is how they grow – and of course a large absolute number of rich.

The model in section 2 would predict that we should see relatively more rich households moving into superstar cities. The top panel of Table 7 shows exactly that result. Each column corresponds to the regression of the share of the MSA's movers (in-migrants or out-migrants, depending on the row) in that income bin on a measure of price growth and a measure of unit growth. Bin 1 is the poorest and bin 5 is the richest.[24] MSAs with greater house price growth have fewer low-income in-migrants and more middle-to-upper income ones. In fact, a one standard deviation increase in the annual rate of house price growth leads to a 2.5 percentage point (-0.038 * 0.664) decline in the share of movers-in who are poor. Conversely, high unit growth MSAs have a larger proportion of their movers in the poor category, as we would expect from the model, and fewer of their movers from the richer bins.

This pattern holds when we compare MSAs in the top quartile of house price or unit growth to those in the bottom 75 percent (see the third panel of Table 7). Being in the top quartile of house price growth leads a MSA to have 6.3 percentage points fewer poor in-migrants, on a base of 34.8. The elasticity is steadily decreasing across the income range, until in the rich category, being in the top quartile leads to a 1.6 percentage point increase in the share of movers that are rich. On a base of 3.2 percent, that corresponds to a 50 percent rise.[25] MSAs in the top quartile of unit growth have steadily declining shares of the movers as the income category goes up, compared to the remaining 75 percent of the MSAs.

We do not see such a compelling pattern for out-migrants (panels 2 and 4). There is no discernable effect of being a high house price appreciation MSA on outmigration. MSAs with

[24] We restrict the sample to 1980 and 1990 because the 'MSA-five-years-ago' variable did not exist prior to 1980 – except for some MSAs in 1940 – and in 2000 the number of named MSAs was drastically reduced.

[25] While the 0.037 estimated coefficient in the middle income category looks large statistically, it represents only a 26 percent increase (i.e., 0.037/0.148)

large growth in units do see an increase in their share poor who outmigrate, but there are no other statisticially significant relationships.

To rule out possible unobserved MSA-level heterogeneity that is masked by running separate regressions for rich and poor households, we also estimate whether the superstar effect on the share of movers in the income bin differs for the rich and poor categories, controlling for unobserved MSA-level heterogeneity with MSA effects. Again, superstar cities have a larger share of rich in-migrants and high unit-growth MSAs have a lower share, as can be seen in Table 8. Out-migrants are not affected by superstar city status but are less likely to be rich as the growth in the number of units rises.

3.5 "Superstar Places"

While the evidence thus far is consistent with the persistent differences in house price growth rates across metropolitan areas being due to the phenomenon we characterize as superstar cities, skeptics still might present an alternative explanation by which agglomeration still fits the facts. For example, the growth in rich residents in superstar cities and the in-mobility of rich residents both might be due to high human capital workers who are attracted to high productivity cities, where their productivity grows further. We address this issue further by examining additional implications of our superstar cities model for the behavior of prices and the distribution of incomes in suburbs within a MSA—a "fractal" version of superstar cities.

We replicate some of the major findings from the previous section on the distribution of the rich (and changes in the distribution of the rich) to suburbs or places within a MSA. Examining changes in house prices and income distributions in suburbs has some clear advantages (and some disadvantages) relative to looking at MSAs within the nation. The most important advantage is that the changes in the local concentration of the rich would take place within a common labor market, so agglomeration or other production externalities cannot easily explain these findings.

Another appreciable advantage is that MSAs exhibit much more variation in changes in the overall income distribution. For example, superstar cities exhibit dramatic growth in the percentage of rich residents while non-superstar cities have much slower than average increases in the proportion rich. Some MSAs even have decadal declines in the proportion of rich residents.

However, within-MSA changes have some disadvantages. For example, the results are likely to be much noisier. Using a large number of relatively small places exacerbates sampling errors and small numbers problems. We limit our sample to places with at least 500 families as a compromise. While the basic results reported here remain even if we include all places, they get appreciably stronger (i.e., with even more reasonable coefficients and smaller standard errors) if we restrict the sample further to places with at least 3000 families.

A second complication has to do with the types of suburbs that exhibit growth in housing units and the process of filtering. When demand for housing units increases, builders tend to construct above-average quality properties. Thus, suburbs with new construction may have a relatively well-to-do population, displacing the poor to other suburbs and central city locations with lower-quality housing. These effects are outside the superstar cities framework and potentially bias the within-MSA or suburb regressions against finding results that are consistent with our model. While the superstar city model fits the data for the proportion of wealthy residents, it performs less well in explaining the distribution of poor residents within a MSA.

3.5.1 Data and place definitions

Ideally, we would like to measure patterns of development and income distribution at the local jurisdiction level. Differences across states in legal structure and urban form make this a difficult problem. States differ in the extent to which local jurisdictions control new construction and even whether jurisdictions can expand their boundaries over time. In the interest of comparability, we examine census designated places—places for short—which are supposed to be independent political jurisdictions.[26] We always include MSA or place dummy variables in an attempt to control for the many unobserved differences across MSAs and places.

The data are derived from a package of CD-ROMs compiled by GeoLytics from long-form data from the 1970, 1980, 1990, and 2000 decennial *Censuses of Housing and Population*. For the census year 1970, the CD-ROMs only include 6,963 (out of 20,768) places.[27] However, these 6,963 places account for more than 95 percent of the U.S. population. One can assess how representative this sample is by considering that in 2000 there are 161 million people living in those places, 206 million people in all the places, and 281 million people in the entire United States. We further limit the sample to places within a MSA. As a robustness check, we have performed the regressions reported below with a balanced sample that consists only of places that are available in all four Census years. Not only do our conclusions remain unchanged, but the results become stronger with larger coefficients and lower standard errors.

3.5.2 Findings

[26] Future research may further restrain the sample, removing MSAs where places serve as a poor proxy for jurisdictions that can constrain supply or where places can expand their boundaries or merge over time.
[27] Conversations with the Census Bureau suggest that the micro data on the remaining places has been lost or is not readily available.

We begin by replicating the basic results for superstar cities. We define a superstar place as a high price growth, low unit growth jurisdiction. All variables are defined based on the local analog to the national regressions. Percentage price growth is based on the annualized change in the average valuation of a housing unit between decades. Unit growth is defined as the annualized change in the number of housing units in a place. The upper quartile dummies for percentage growth in valuation and housing units are based on all places within a MSA with at least 500 families. Finally, we define the income bins based on the 1980 income topcodes, as described above for the MSA data.

Despite potential difficulties, the results in Table 9 suggest that the fractal version of superstar cities fits the data quite well. High price growth (superstar) places are correlated with a much higher proportion of rich residents, while high unit growth places have a much lower proportion of rich residents. Our base regression (column 1) is a pooled regression over all four Census years and includes a full set of MSA and year dummies.[28] The coefficient estimate suggests that a one standard deviation increase in the annual average growth rate in valuation (1.56 percent) is correlated with a relatively large 2.2 percent increase in the proportion of rich residents, compared to a mean proportion rich of 4.4 percent. By comparison, column 2 shows that these findings become larger if the sample is limited to a balanced panel of places that exist in all four Census years. Of additional interest, other than in the relatively small 1970 sample, the coefficient estimates rise between 1980 and 2000, suggesting a greater concentration of the rich in superstar places over the past two decades. This is a similar, but less restrained pattern seen in the national data (Table 4). The lower panel documents quite similar findings when we use dummies for the top quartile of price and unit growth rates instead of the continuous variables.

[28] The results are virtually unchanged if we include interactions between year and MSA dummies as well.

When we examine changes in the proportion of rich residents across decades (Table 10), our results remain consistent with superstar places attracting a disproportionate share of rich families. The coefficients on the interaction between growth in MSA rich and annualized growth in average values is always economically large, positive, and statistically significant. A one-standard deviation increase in the growth rate of housing values leads to an increase in the proportion of wealthy residents of nearly fifty percent. Less consistent with superstar cities is the finding in the top panel of column one that places with high unit growth also have an increase in the proportion of rich families. That said, the estimated effect of a one-standard deviation increase in the growth rate of units is only about 40 percent as large as a standard deviation increase in valuation growth.

The results from the specifications reported in columns 2 and 3 of Table 10 are more consistent with the findings for superstar cities. The coefficients suggest that the change in percentage rich and the absolute number of rich increase in superstar cities, but are flat or fall in elastically supplied places. The coefficients on the dummy variables in the lower panel are still consistent with superstar cities attracting a disproportionate share of the wealthy, but also show wealthy residents moving to places in the top quartile of unit growth. Given the likely biases associated with rich residents moving to high-quality new homes, these results still suggest the existence of superstar places.

Finally, Table 11 examines the possibility that superstar places “fill-up” in the same sense that superstar cities have filled up. Here the findings are decidedly mixed. While the coefficient on the interaction between top quartile of price growth and the price growth is negative and significant in 1980, it becomes small and not statistically different from zero in 1990 and 2000.

Put together, these findings are consistent with a "fractal" version of superstar cities within MSAs for places. High price growth (superstar) places are associated with a large and growing share of the wealthiest families in a MSA. High unit growth places are correlated with below average ratios of wealthy families and reductions in the number of wealthy families and a lower ratio of wealthy families, although the latter results are not as robust as the former findings. Finally, the data do not show that superstar places have completely filled-up, as superstar cities clearly have.

4. Conclusion

With an increasing national population, people have to choose to live somewhere, and most choose to live in metropolitan areas. What is remarkable is that a significant fraction are willing to pay a premium to live in certain places, and that gap is increasing over time.

The foundations of our explanation for this phenomenon are as follows: if some cities are inelastically supplied in that they have geographically or politically limited growth and are not readily replicable, and as long as some fraction of the population prefers to live there, the city can command positive rents. As the high income population grows, the mass of households willing to pay the entrance fee for the scarce city goes up, but the city does not have any additional spaces. Thus, the cutoff rent (or house price) goes up over time.

This theory is well supported by the data. We find that superstar cities, those with high long-run house price growth due to scarcity, have residents with income distributions shifted to the right of non-superstar cities. As the national high-income population grows, superstars obtain more than their *pro-rata* share of the additional rich.

We also present evidence that superstar cities must be at least a significant part of the explanation for the evolution of house prices that we observe, despite there being alternative theories of such growth. We do not try to disprove these alternatives, but show that they are insufficient to explain all the patterns in the data without superstar cities. First, in superstar cities a larger fraction of the in-migrants are rich. This is what one would expect if skilled people moved to superstar cities (they can outbid others), but would not necessarily expect to find if production externalities made people who move to the MSA more productive. In addition, controlling for MSA effects, we find a "superstar" pattern among places within MSAs. Inelastically supplied places experience a different composition of income compared to other high-demand, but high-supply elasticity places. By controlling for unobserved MSA heterogeneity, we control for agglomeration benefits.

Given the rapid rise in house prices in the Bay Area, one naturally wonders whether it can continue. Classic economics would argue that it could not, as residents would substitute to other MSAs if the price gap grew too high. Agglomeration economics would say continued growth was possible as long as real productivity growth were 4 percent per year. Superstar cities, on the other hand, points out that all that would be necessary is for the national population to keep growing and for the Bay Area to remain attractive on some dimensions that cannot easily be duplicated elsewhere. If there is continued dispersion of incomes in the U.S., the price rises can be even faster. However, it appears that a large part of the growth in incomes and prices in the Bay Area has been from a composition change: the rich arrivals have crowded out the poor. Once the area is filled with rich families, growth in house prices will be limited by the growth in income among the rich.

Selected References (Incomplete, more to be added.)

Bayer, Patrick, Fernando Ferreira and Kim Rueben (2004). "Tiebout Sorting, Social Multipliers, and the Demand for School Quality", Yale University Economics Department Working Paper.

Chinitz, Benjamin (1961). "Contrasts in Agglomeration: New York and Pittsburgh", *American Economic Review*, 51: 279-89.

Davis, Donald and David Weinstein (2001). "Bones, Bombs and Break Points: The Geography of Economic Activity", NBER Working Paper 8517.

______________________________ (2004). "A Search for Multiple Equilibria in Urban Industrial Structure", NBER Working Paper 10252.

Duranton, Gilles and Diego Puga (2004). "Microfoundations of Urban Agglomeration Economies" in J.V. Henderson and J-F. Thisse (eds.) *Handbook of Regional and Urban Economics, Volume 4*.

Glaeser, Edward and Joseph Gyourko (2004-revised), "Durable Housing and Urban Decline", NBER Working Paper 8598.

Glaeser, Edward, Joseph Gyourko, and Raven Saks (2004-revised), "Why Is Manhattan So Expensive? Regulation and the Rise in House Prices", NBER Working Paper 10124.

Glaeser, Edward, Hedi Kallal, Jose Scheinkman, and Andrei Shleifer (1992). "Growth in Cities", *Journal of Political Economy*, 100(6): 1126-1152.

Glaeser, Edward, Jed Kolko and Albert Saiz (2001). "Consumer Cities", *Journal of Economic Geography*, 1(1): 27-50.

Glaeser, Edward, Jose Scheinkman, and Andrei Shleifer (1995). "Economic Growth in a Cross-Section of Cities", *Journal of Monetary Economics*, 36: 117-143.

Glaeser, Edward and Albert Saiz (2003). "The Rise of the Skilled City", NBER Working Paper 10191.

Henderson, J. Vernon (1997). "Externalities and Industrial Development", *Journal of Urban Economics*, 42: 449-470.

Henderson, J. Vernon, A. Kuncoro, and M. Turner (1995). "Industrial Development in Cities", *Journal of Political Economy*, 106(4): 667-705.

Jacobs, Jane (1969). *The Economy of Cities*. New York: Vintage.

Kopczuk, Wojciech and Emmanuel Saez (2004), "Top Wealth Shares in the United States, 1916-2000: Evidence from Estate Tax Returns", *National Tax Journal,* LVII(2), part 2, 445-488.

Lazear, Edward and Sherwin Rosen (1981), "Rank-Order Tournaments as Optimum Labor Contracts", *Journal of Political Economy*, 89(5): 841-864.

Lucas, Robert (1988). "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22: 3-42.

Mayer, Christopher and Tsur Somerville (2000), "Land Use Regulation and New Construction", *Regional Science and Urban Economics*, 30(6), 639-62.

Mills, Edwin and Wallace Oates (1975). *Fiscal Zoning and Land Use Controls*. Lexington, MA: Heath-Lexington.

Moretti, Enrico (2004a). "Estimating the External Return to Higher Education: Evidence from Cross-Sectional and Longitudinal Data", *Journal of Econometrics*, Vol. 121.

____________ (2004b). "Human Capital Externalities in Cities" in J.V. Henderson and J-F. Thisse (eds.) *Handbook of Urban and Regional Economics*, Volume 4.

____________ (2004c). "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions", *American Economic Review*, forthcoming.

Rauch, Jonathan (1993). "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities", *Journal of Urban Economics*, 34: 380-400.

Roback, Jennifer (1980). "The Value of Local Urban Amenities: Theory and Measurement", Ph.D. dissertation, University of Rochester.

_____________ (1982). "Wages, Rents, and the Quality of Life", *Journal of Political Economy*, 90(2): 1257-78.

Romer, Paul (1986). "Increasing Returns and Long-Run Growth", *Journal of Political Economy*, 94(5): 1002-37.

Rosen, Sherwin (1979). "Wage-Based Indexes of Urban Quality of Life" in P. Mieszkowski and M. Strasheim (eds.), *Current Issues in Urban Economics*. Baltimore, MD: Johns Hopkins University Press.

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

Rosenthal, Stuart and William Strange (2001). "The Determinants of Agglomeration", *Journal of Urban Economics*, 50: 191-229.

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

Rosenthal, Stuart and William Strange (2004). "Evidence on the Nature and Sources of Agglomeration Economies" in J.V. Henderson and J-F. Thisse (eds.) *Handbook of Regional and Urban Economics, Volume 4*.

Rothstein, Jesse (2003). "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition Among Jurisdictions", Princeton University Economics Department Working Paper.

Saez, Emmanuel (2004). "Reported Incomes and Marginal Tax Rates, 1960-2000: Evidence and Policy Implications", NBER Working Paper 10273.

Saks, Raven (2003). "From New York to Denver: Housing Supply Restrictions Across the United States", Economics Department Working Paper, Harvard University.

Saxenian, AnnaLee (1999). R*egional Advantage: Culture and Competition in Silicon Valley and Route 128.* Cambridge, MA: Harvard University Press.

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

Table 1: Real annual house price growth, for the top 15 MSAs, 1940 population>170,000

Rank	MSA	Annualized growth rate, 1960-2000	MSA	Annualized growth rate, 1940-2000
1	San Jose	3.96	San Jose	4.08
2	San Francisco	3.84	San Francisco	3.75
3	Oakland	3.16	Seattle	3.18
4	Seattle	2.78	San Diego	3.09
5	San Diego	2.71	Phoenix	3.01
6	Portland	2.60	Portland	2.92
7	Los Angeles	2.47	Los Angeles	2.91
8	Boston	2.47	Tacoma	2.80
9	Tacoma	2.45	Denver	2.67
10	Austin	2.40	Knoxville	2.65
11	Raleigh-Durham	2.32	Salt Lake City	2.49
12	Charlotte	2.29	Sacramento	2.48
13	Trenton	2.25	Boston	2.48
14	Denver	2.19	Atlanta	2.43
15	Nassau-Suffolk Cty	2.16	Mobile	2.40
	All metro areas	1.77	All metro areas	2.09
	All, excluding top 15	1.66	All, excluding top 15	1.87

Table 2: 30-year growth rate transition matrix

		1970-2000			
		Top Quartile	**2nd**	**3rd**	**4th**
1940-1970	**Top Quartile**	**16**	**6**	**6**	**4**
	2nd	**8**	**8**	**7**	**9**
	3rd	**4**	**7**	**7**	**14**
	4th	**4**	**11**	**12**	**6**

Notes: Uses the 129 MSAs with house price data in 1940

Table 3: Summary statistics

Variable	Mean	Standard deviation
MSA Level:		
Average Annual House Price Growth, 1960-2000	1.447	0.664
Average Annual Housing Unit Growth, 1960-2000	2.081	1.137
Average Annual Income Growth, 1960-2000	1.323	0.366
Share of an MSAs population that is "rich"	2.716	1.853
Share of an MSA's population that is "poor"	44.311	9.883
(Change in MSA's rich)/MSA's families	0.009	0.015
Change in (MSA rich/MSA families)	0.027	0.019
MSA number rich	5,888	14,734
Percent of in-migrants in each income category:		
Poorest (1)	34.8	16.5
Less poor (2)	43.9	13.0
Middle (3)	14.3	9.3
Less rich (4)	3.8	4.1
Rich (5)	3.2	5.7
Percent of out-migrants in each income category:		
Poorest (1)	30.2	13.8
Less poor (2)	44.7	13.5
Middle (3)	17.0	11.1
Less rich (4)	4.6	5.2
Rich (5)	3.5	7.5
Place Level:		
Average Annual House Price Growth, 1960-2000	1.310	1.566
Average Annual Housing Unit Growth, 1960-2000	1.896	3.096
Share of a place's population that is "rich"	0.044	0.075
(Change in place's rich)/place's families	0.019	0.057
Change in (place's rich/place's families)	0.011	0.036

Table 4: Evolution of the income distribution in superstar cities

Period	1970-2000	1960	1970	1980	1990	2000
Rich share of MSA						
Price growth	1.100 (0.065)		0.512 (0.093)	0.254 (0.055)	1.360 (0.126)	2.274 (0.176)
Unit growth	-0.036 (0.038)		-0.011 (0.054)	0.064 (0.032)	-0.079 (0.074)	-0.118 (0.103)
Adj. R^2	0.4310		0.0950	0.1052	0.2862	0.3698
Top quartile price growth	1.600 (0.096)		0.942 (0.134)	0.502 (0.081)	1.924 (0.191)	3.033 (0.274)
Top quartile unit growth	-0.124 (0.096)		-0.199 (0.134)	-0.043 (0.081)	-0.175 (0.191)	-0.057 (0.274)
Adj. R^2	0.4173		0.1320	0.1102	0.2497	0.2956
Poor share of MSA						
Price growth	-4.837 (0.399)	-3.220 (0.894)	-3.919 (0.927)	-2.662 (0.764)	-6.693 (0.754)	-6.072 (0.686)
Unit growth	1.928 (0.233)	2.643 (0.522)	3.020 (0.541)	2.164 (0.446)	1.482 (0.440)	1.047 (0.400)
Adj. R^2	0.2355	0.0788	0.0982	0.0727	0.1970	0.1980
Top quartile price growth	-6.909 (0.590)	-5.288 (1.329)	-5.878 (1.381)	-4.519 (1.126)	-8.892 (1.141)	-8.202 (1.036)
Top quartile unit growth	4.291 (0.588)	5.362 (1.329)	6.499 (1.381)	5.138 (1.126)	3.070 (1.141)	2.486 (1.036)
Adj. R^2	0.2311	0.0673	0.0845	0.0776	0.1581	0.1625

Notes: N=1,256 in the first column, 314 in subsequent ones. All regressions include year dummies. Growth rates are the annual average over the 1960-2000 period; prices are in real dollars. "Rich" corresponds to income >= 100 percent of the 1980 topcode amount. "Poor" corresponds to income < 25 percent of the 1980 topcode amount.

Table 5: The relationship between the national income distribution
and the growth in the MSA income distribution

Change in rich divided by share of families	(Change in all metro areas' rich) / families	**Change in share of families that are rich**	Change in all metro areas' (rich/families)	**Change in rich**	Change in all metro areas' rich
(Change in MSA rich) / families		Change in MSA (rich / families)		Change in MSA rich	
× average annual growth in house values	0.369 (0.028)	× average annual growth in house values	0.503 (0.022)	× average annual growth in house values	0.0040 (0.0003)
× average annual growth in housing units	-0.076 (0.016)	× average annual growth in housing units	-0.044 (0.013)	× average annual growth in housing units	-0.0006 (0.0002)
Adj. R^2	0.6484	Adj. R^2	0.8453	Adj. R^2	0.7744
× in the top quartile of growth in house values	0.480 (0.042)	× in the top quartile of growth in house values	0.626 (0.035)	× in the top quartile of growth in house values	0.0056 (0.0005)
× in the top quartile of growth in housing units	-0.090 (0.042)	× in the top quartile of growth in housing units	0.001 (0.035)	× in the top quartile of growth in housing units	-0.0004 (0.0005
Adj. R^2	0.6270	Adj. R^2	0.8212	Adj. R^2	0.7728
N	942	N	1,256	N	1,256

Notes: Standard errors in parentheses. All regressions include year and MSA dummies. The sample includes the 314 MSAs with data over the 1970 through 2000 period. Growth rates are the annual average over the 1960-2000 period; prices are in real dollars. "Rich" corresponds to income >= 100 percent of the 1980 topcode amount.

Table 6: The Relationship Between High Price Growth MSAs and the Change in the Number of Housing Units, by Decade

	1970	**1980**	**1990**	**2000**
Average House Price Growth, 1960-2000	11.12	17.18	11.73	9.37
	(4.76)	(3.77)	(2.19)	(1.51)
In Top Quartile of Average Price Growth	6.10	35.23	31.99	24.99
	(16.02)	(12.68)	(7.38)	(5.08)
Average Price Growth × In Top Quartile	-3.12	-18.14	-15.35	-13.26
	(7.91)	(6.26)	(3.64)	(2.51)
N	314	314	314	314
Adj. R^2	0.04	0.10	0.16	0.15

Notes: The left-hand-side variable is the decadal percent change in the number of housing units. Standard errors in parentheses. To be in the top quartile, average real house price growth must have exceeded 1.75 percent over the 1960-2000 period.

Table 7: The difference among places with long-run price growth in the income distribution of movers

	Income Bins				
	(1)	(2)	(3)	(4)	(5)
In-migrants:					
Average House Price Growth, 1960-2000	-0.038 (0.012)	-0.010 (0.010)	0.026 (0.007)	0.012 (0.003)	0.010 (0.004)
Average Unit Growth, 1960-2000	0.032 (0.007)	-0.014 (0.006)	-0.012 (0.004)	-0.004 (0.002)	-0.002 (0.002)
Adj. R^2	0.0568	0.0638	0.1351	0.0670	0.0642
Out-migrants:					
Average House Price Growth, 1960-2000	0.005 (0.010)	-0.012 (0.010)	0.003 (0.008)	0.003 (0.004)	0.001 (0.006)
Average Unit Growth, 1960-2000	0.020 (0.006)	-0.011 (0.006)	-0.003 (0.005)	-0.002 (0.002)	-0.004 (0.003)
Adj. R^2	0.1143	0.730	0.1268	0.0708	0.0712
In-migrants:					
Top quartile price growth	-0.063 (0.018)	-0.006 (0.014)	0.037 (0.010)	0.016 (0.004)	0.016 (0.006)
Top quartile unit growth	0.084 (0.018)	-0.037 (0.014)	-0.033 (0.010)	-0.009 (0.004)	-0.005 (0.006)
Adj. R^2	0.0663	0.0606	0.1386	0.0609	0.0603
Out-migrants:					
Top quartile price growth	0.007 (0.014)	-0.021 (0.014)	0.007 (0.011)	0.004 (0.006)	0.001 (0.008)
Top quartile unit growth	0.040 (0.014)	-0.012 (0.015)	-0.014 (0.012)	-0.008 (0.006)	-0.006 (0.008)
Adj. R^2	0.1013	0.0635	0.1286	0.0725	0.0694

Notes: The left-hand-side variable is the share of the MSA's number of in-migrants (or out-migrants) that are in the income category. The sample is restricted to 1980 and 1990. All specifications include a year dummy. Standard errors in parentheses. To be in the top quartile, average real house price growth must have exceeded 1.75 percent over the 1960-2000 period.

Table 8: The difference in density of rich and poor movers
in high- and low- house price growth MSAs

	In-migrants	Out-migrants
Rich category dummy	-0.317 (0.021)	-0.214 (0.019)
Rich × average house price growth, 1960-2000	0.047 (0.013)	-0.006 (0.012)
Rich × average housing unit growth, 1960-2000	-0.032 (0.008)	-0.021 (0.007)
Adj. R^2	0.6370	0.5910
Rich category dummy	-0.315 (0.009)	-0.254 (0.009)
Rich × top quartile price growth, 1960-2000	0.079 (0.019)	-0.006 (0.017)
Rich × top quartile unit growth, 1960-2000	-0.086 (0.019)	-0.042 (0.017)
Adj. R^2	0.6415	0.5888

Notes: The left-hand-side variable is the share of the MSA's number of in-migrants (or out-migrants) that are in the income category. The sample is restricted to 1980 and 1990 and includes only those families in the topmost (rich) or bottommost (poor) bins. All specifications include a year dummy and a full set of MSA dummies. Standard errors in parentheses. To be in the top quartile, average real house price growth must have exceeded 1.75 percent over the 1960-2000 period.

Table 9: Evolution of the income distribution in superstar places

Period	1970-2000	1970-2000	1970	1980	1990	2000
	Whole Sample	Places in all 4 Periods	Whole Sample	Whole Sample	Whole Sample	Whole Sample
Rich share of MSA						
Price growth	0.014 (0.0004)	0.042 (0.0006)	0.020 (0.0010)	0.011 (0.0005)	0.012 (0.0007)	0.020 (0.0008)
Unit growth	-0.0008 (0.0002)	-0.0045 (0.0003)	-0.0034 (0.0005)	-0.0018 (0.0003)	-0.0007 (0.0003)	0.0005 (0.0004)
Adj. R^2	0.2417	0.3714	0.1854	0.1221	0.2017	0.2723
N	23,663	15,667	4,028	5,885	6,827	6,923
Top quartile price growth	0.043 (0.0010)	0.054 (0.0012)	0.034 (0.0021)	0.027 (0.0013)	0.045 (0.0019)	0.060 (0.0021)
Top quartile unit growth	-0.0038 (0.0010)	-0.0073 (0.0012)	-0.0120 (0.0021)	-0.0062 (0.0013)	-0.0037 (0.0020)	-0.0003 (0.0021)
Adj. R^2	0.2497	0.2893	0.1550	0.1206	0.2198	0.2738
N	24,281	15,667	4,053	5,913	6,928	7,387

All regressions include year and MSA dummies. Growth rates are the annual average over the 1970-2000 period; prices are in real dollars. “Rich” corresponds to income >= 100 percent of the 1980 topcode amount. Sample is restricted to all places with at least 500 families. The results from using the full sample look similar and the coefficients are always statistically significant, albeit the magnitudes are slightly smaller. The results are also little-changed if we include a full set of year, MSA, and year*MSA interactions.

Table 10: The relationship between the MSA income distribution
and the growth in the place income distribution

Change in rich divided by share of families	(Change in all places' rich) / families	**Change in share of families that are rich**	Change in all places' (rich/families)	**Change in rich**	Change in all places' rich
(Change in places' rich) / families		Change in places' (rich / families)		Change in places' rich	
× average annual growth in house values	0.32 (0.0087)	× average annual growth in house values	0.39 (0.0069)	× average annual growth in house values	0.0048 (0.0003)
× average annual growth in housing units	0.091 (0.0069)	× average annual growth in housing units	-0.019 (0.0062)	× average annual growth in housing units	-0.0005 (0.0003)
Adj. R^2	0.5355	Adj. R^2	0.4209	Adj. R^2	-0.3950
N	16,533	N	16,533	N	16,533
× in the top quartile of growth in house values	1.1 (0.036)	× in the top quartile of growth in house values	1.2 (0.031)	× in the top quartile of growth in house values	0.011 (0.0012)
× in the top quartile of growth in housing units	0.63 (0.037)	× in the top quartile of growth in housing units	0.32 (0.030)	× in the top quartile of growth in housing units	0.0001 (0.0012)
Adj. R^2	0.5017	Adj. R^2	0.3367	Adj. R^2	-0.4249
N	16,604	N	16,602	N	16,604

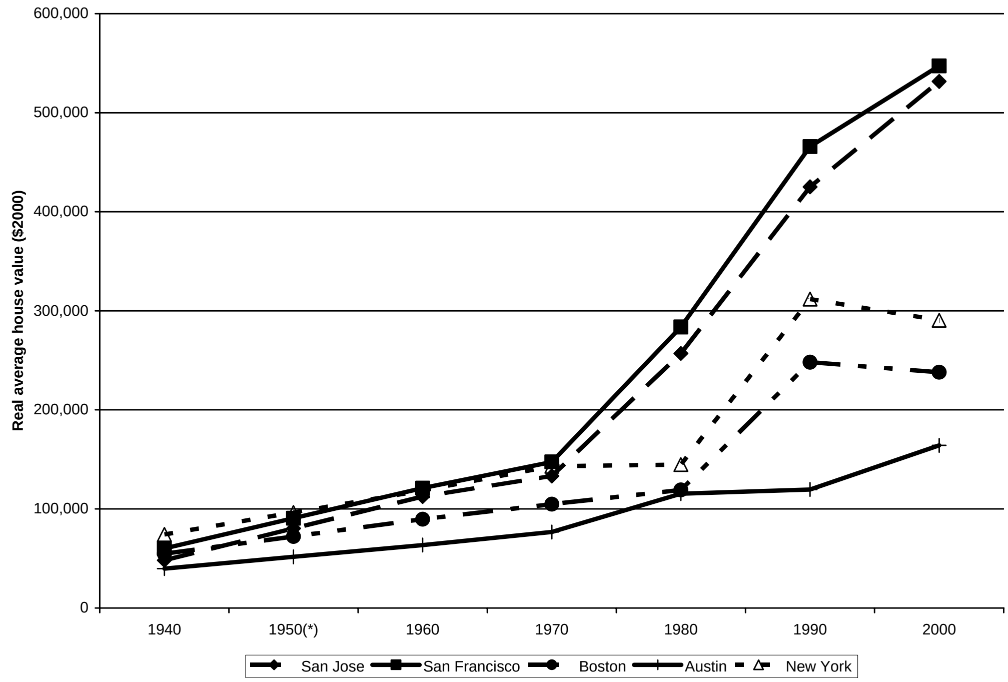
Notes: Standard errors in parentheses. All regressions include year and place dummies. Sample is restricted to all places with at least 500 families. Growth rates are the annual average over the 1970-2000 period; prices are in real dollars. "Rich" corresponds to income >= 100 percent of the 1980 topcode amount.

Table 11: The Relationship Between High Price Growth Places and the Change in the Number of Housing Units, by Decade

	1980	**1990**	**2000**
Average House Price	0.053	0.018	0.019
Growth, 1970-2000	(0.017)	(0.009)	(0.005)
In Top Quartile of	0.017	0.058	0.034
Average Price Growth	(0.05)	(0.035)	(0.026)
Average Price Growth x	-0.076	-0.015	0.005
In Top Quartile	(0.019)	(0.013)	(0.009)
N	3994	5885	6654
Adj. R^2	0.1657	0.1090	0.1301

Notes: The left-hand-side variable is the decadal percent change in the number of housing units. Standard errors in parentheses. Sample is restricted to all places with at least 500 families. All regressions include MSA dummies.

Figure 1: Real average house values, 1940-2000, selected MSAs



*Notes: 1950 data is linearly interpolated.

Figure 2: The distribution of average real annual house price growth, 1940-2000, For MSAs with 1940 population above 170,000

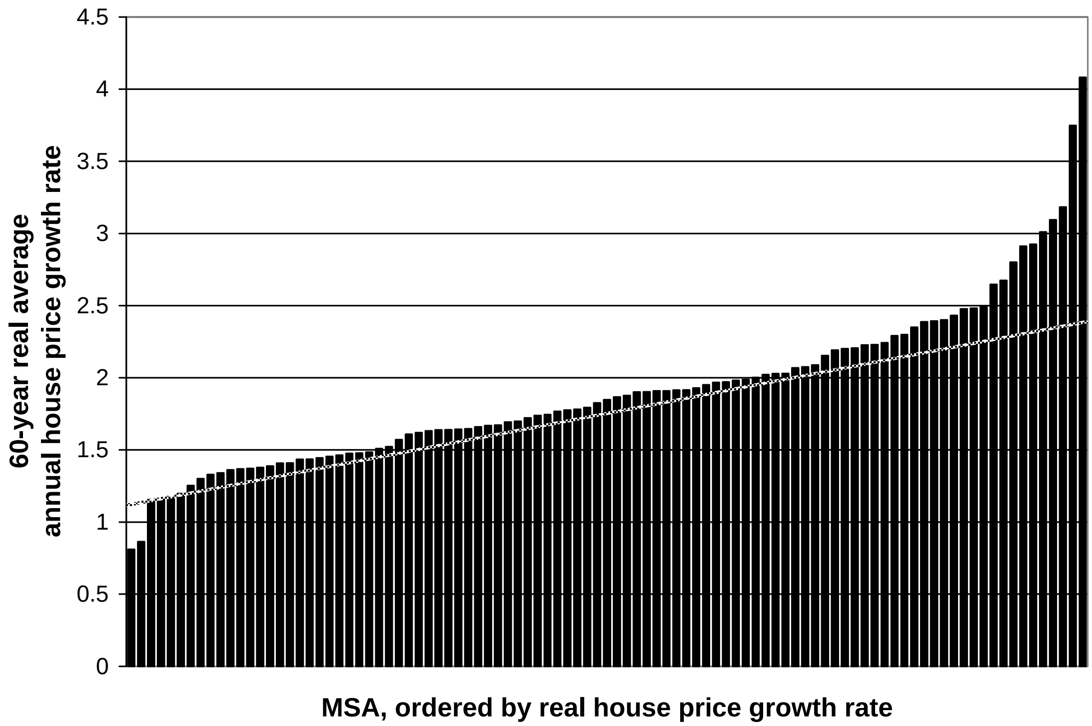
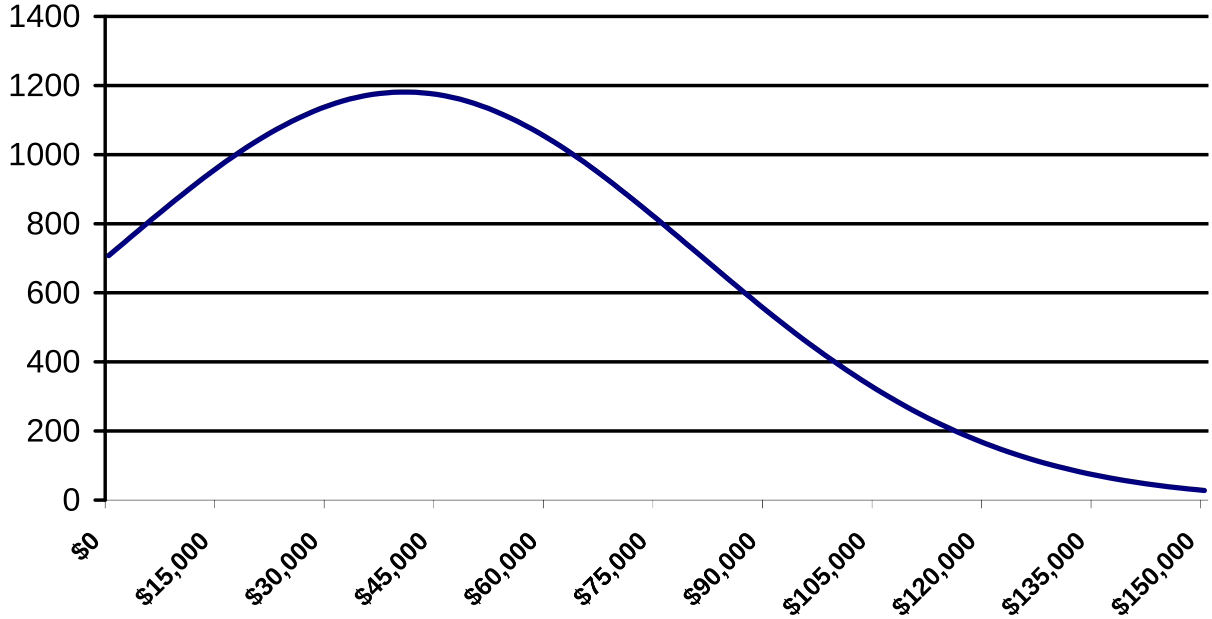
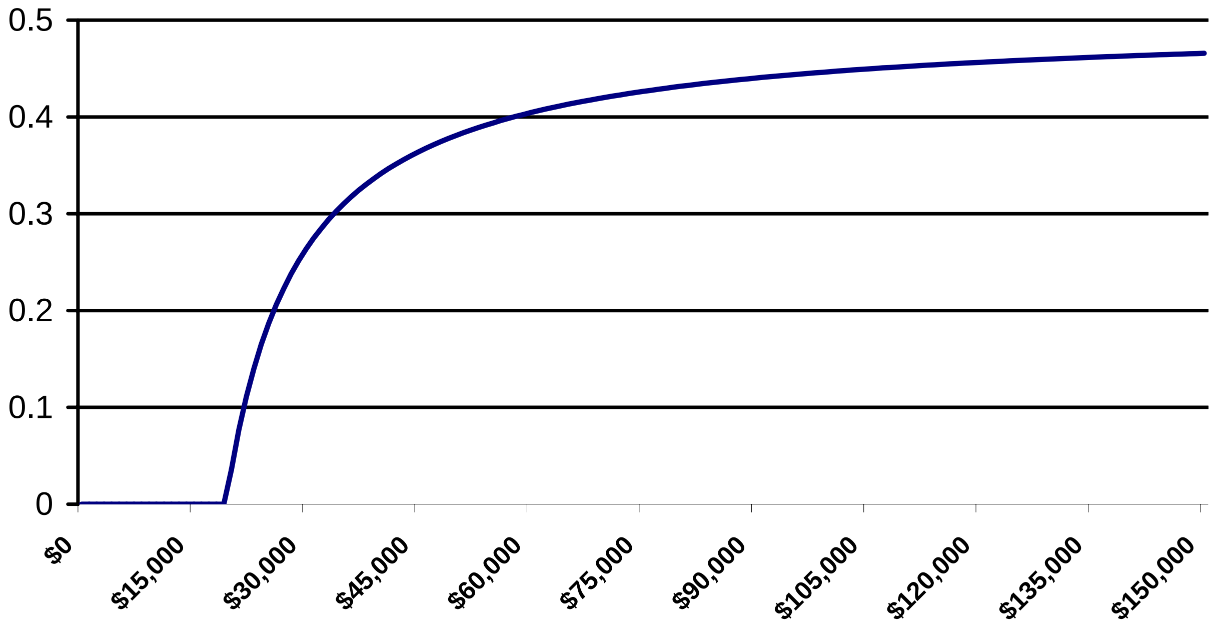


Figure 3
Hypothetical Wage Distribution



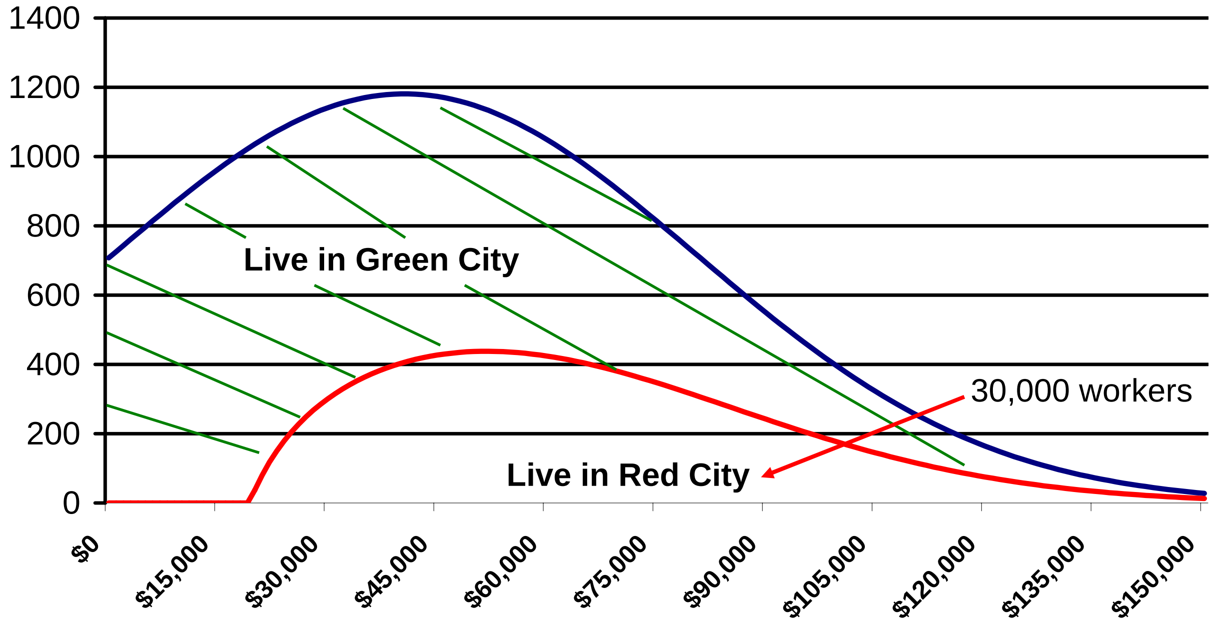
Note: 100,000 workers; truncated normal distribution; mean=40,000; std.=40,000, w_L=$0, w_U =$150,000

Figure 4
Equilibrium Cutoff Values, c*(w)



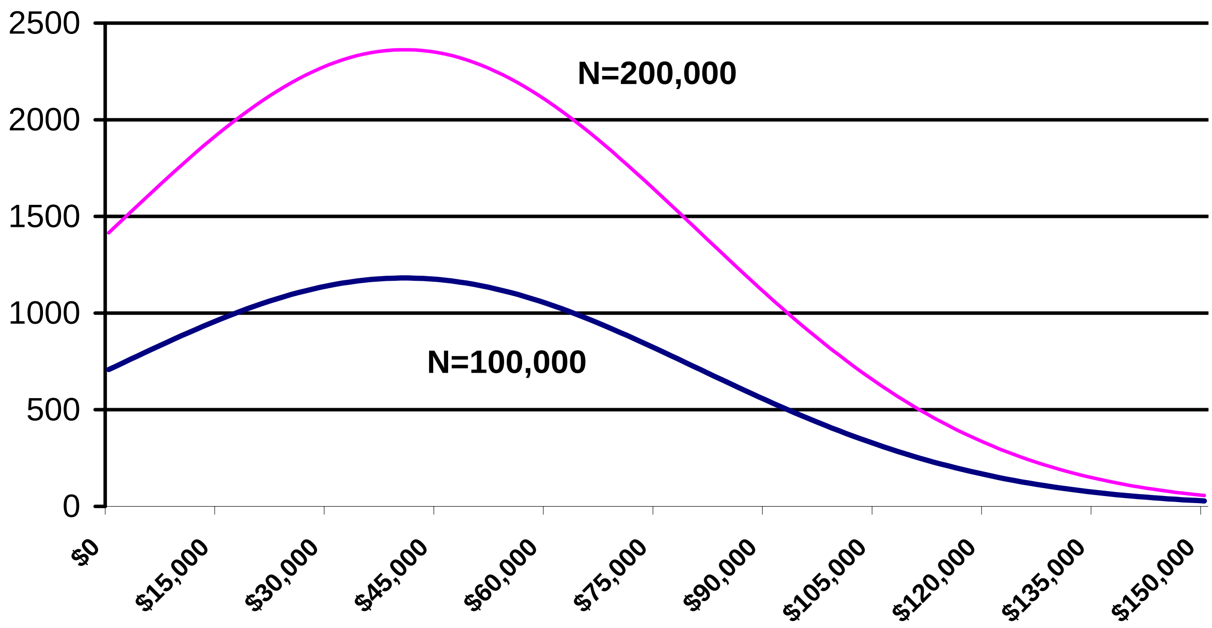
Note: 100,000 workers; truncated normal distribution; mean=40,000; std.=40,000, w_L=$0, w_U =$150,000

Figure 5
Wage Distribution in Red and Green Cities



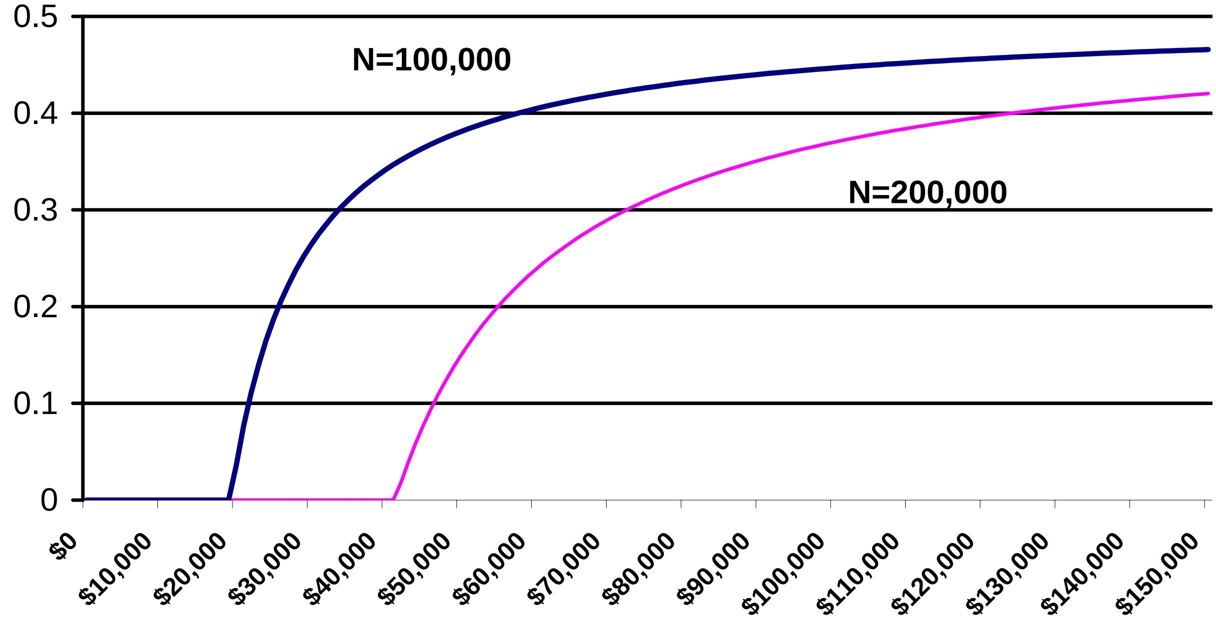
Note: 100,000 workers; truncated normal distribution; mean=40,000; std.=40,000, w_L=$0, w_U =$150,000

Figure 6
Hypothetical Wage Distribution: g=100%



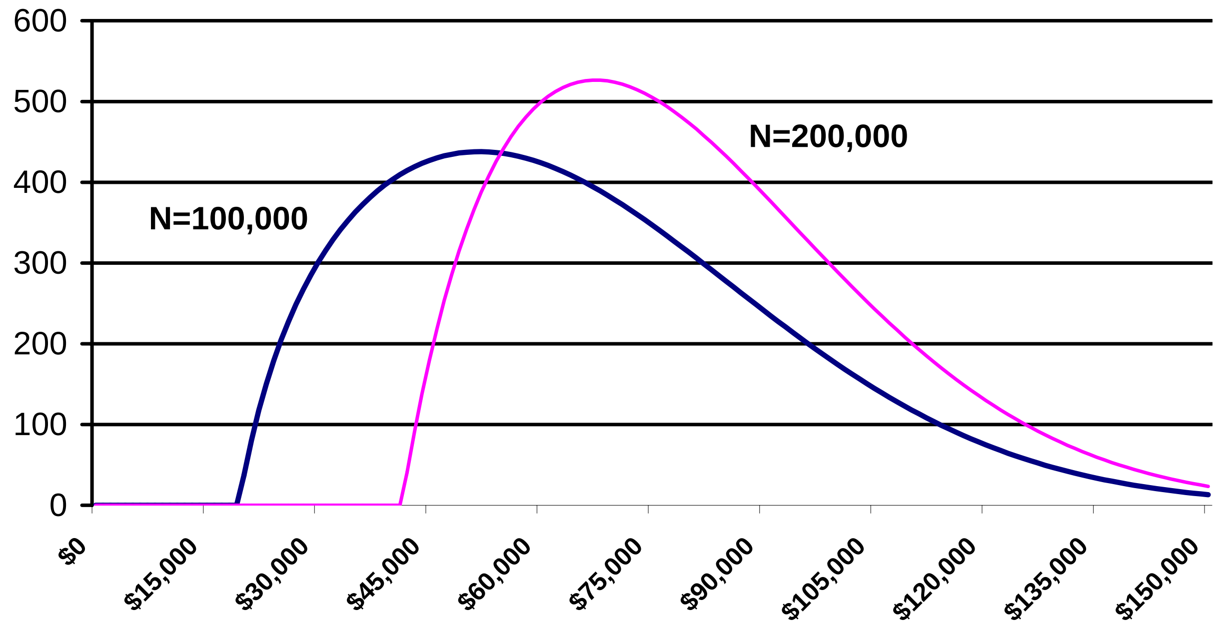
Note: truncated normal distribution; mean=40,000; std.=40,000, w_L=$0, w_U =$150,000

Figure 7
Equilibrium Cutoff Values, c*(w): g= 100%



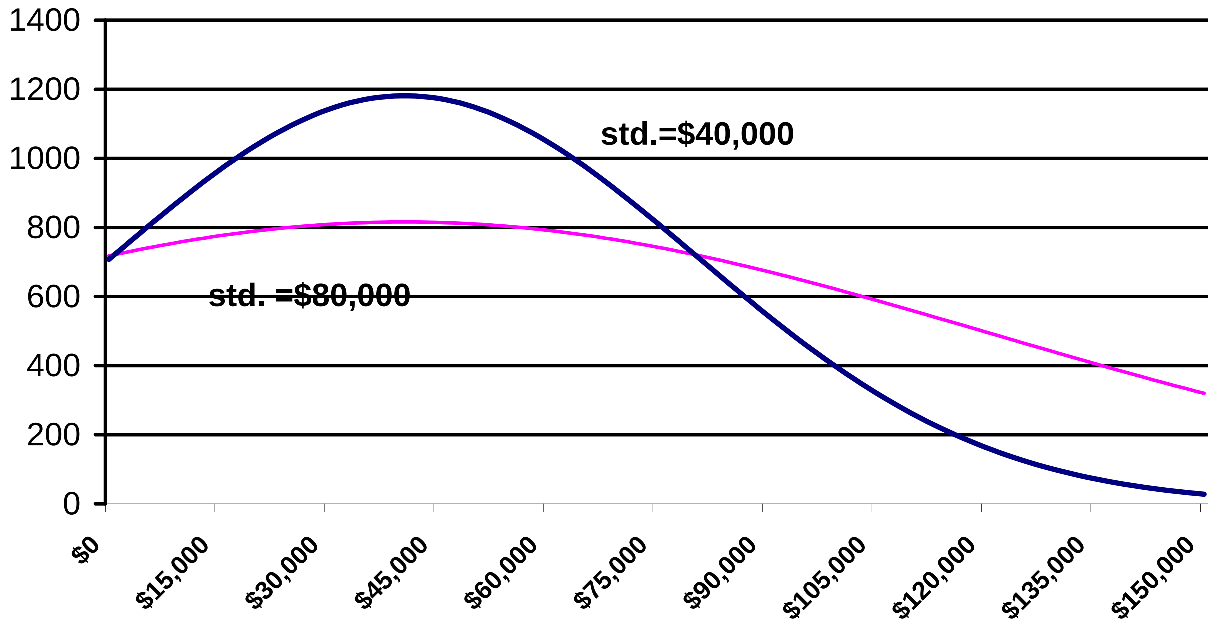
Note: 100,000 workers; truncated normal distribution; mean=40,000; std.=40,000, w_L=$0, w_U =$150,000

Figure 8
Wage Distribution in Red City: g=100%



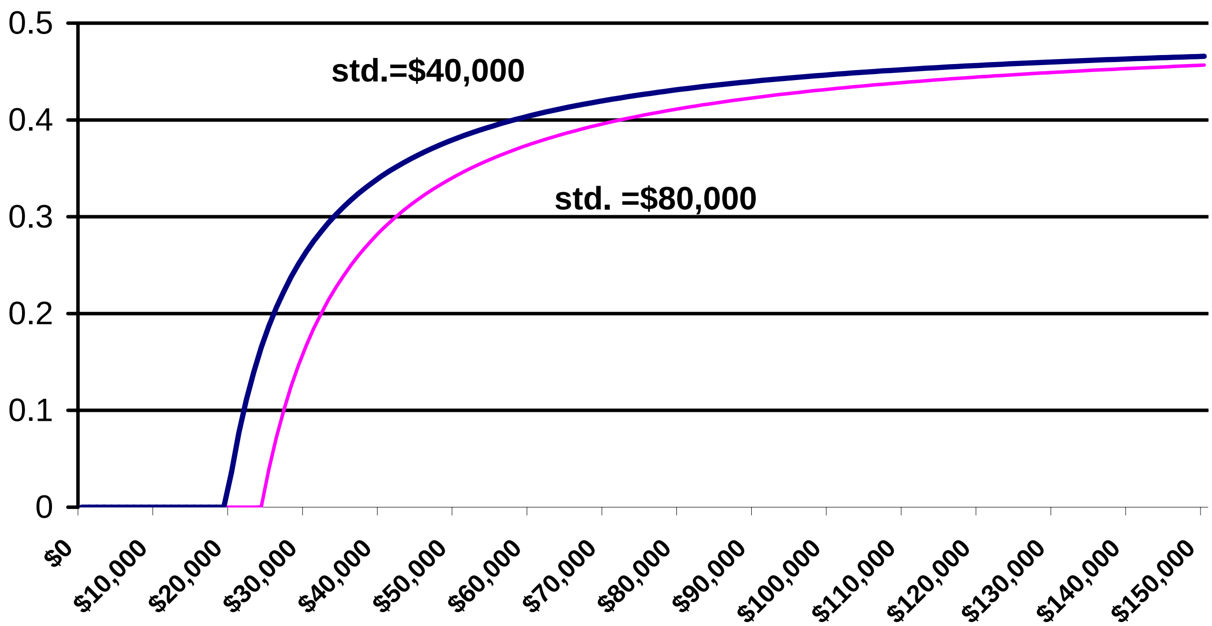
Note: truncated normal distribution; mean=40,000; std.=40,000, w_L=$0, w_U =$150,000

Figure 9
Hypothetical Wage Distribution: Fatter Tails



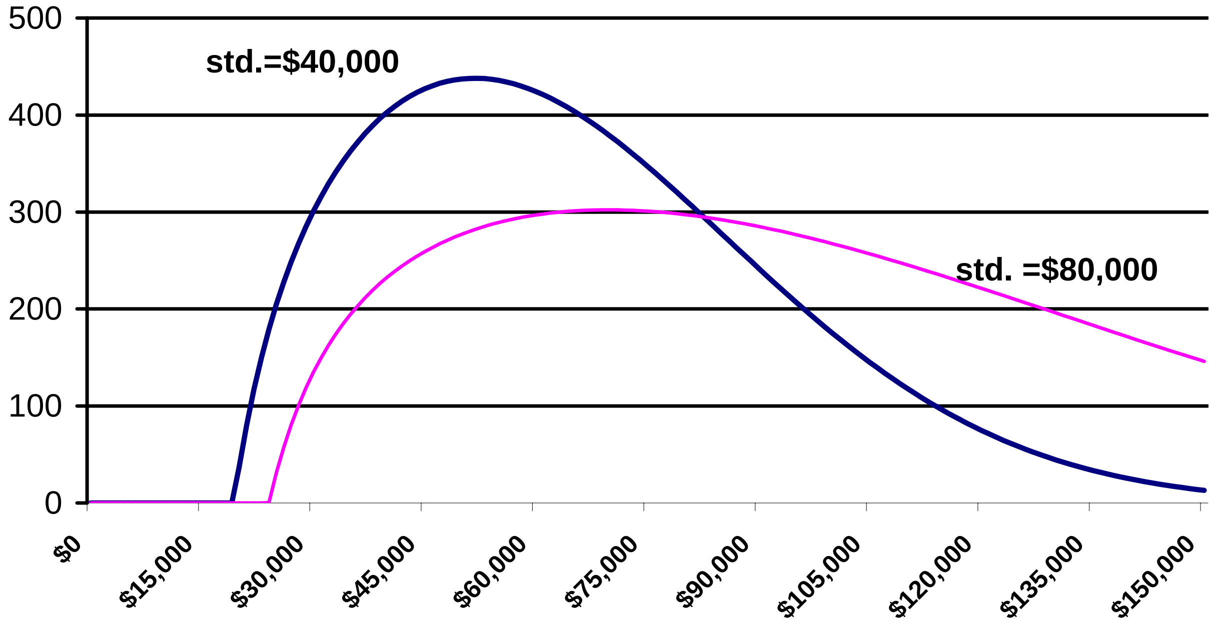
Note: 100,000 workers; truncated normal distribution; mean=40,000; w_L=$0, w_U =$150,000

Figure 10
Equilibrium Cutoff Values, c*(w): Fatter Tails



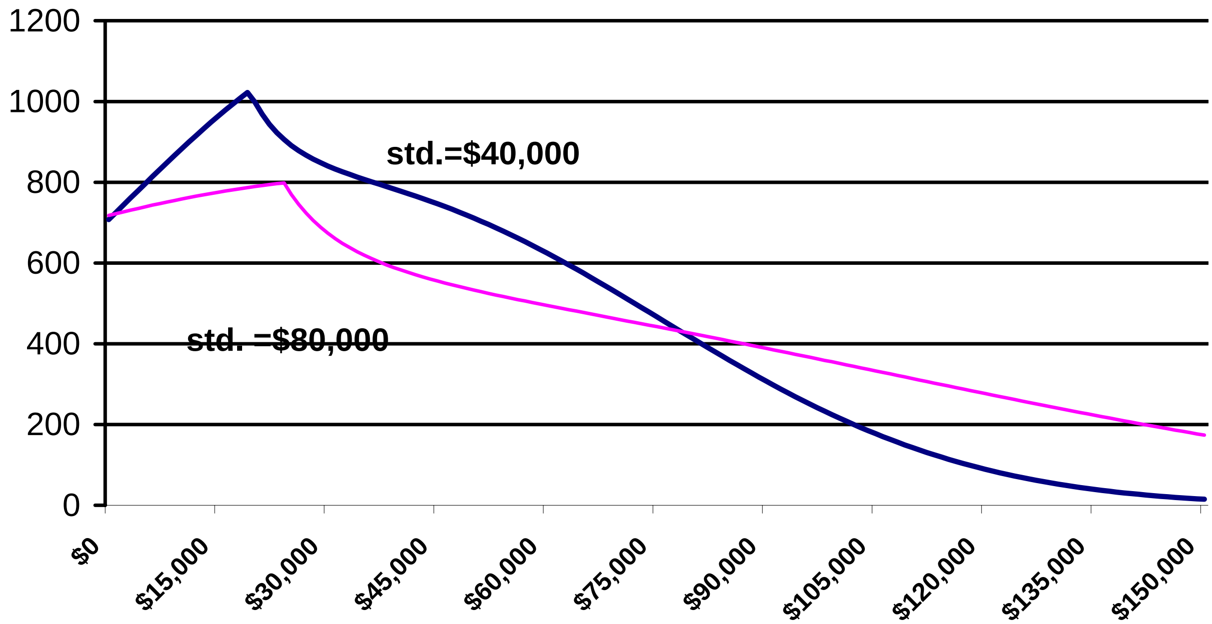
Note: 100,000 workers; truncated normal distribution; mean=40,000; w_L=$0, w_U =$150,000

Figure 11
Wage Distribution in Red City: Fatter Tails



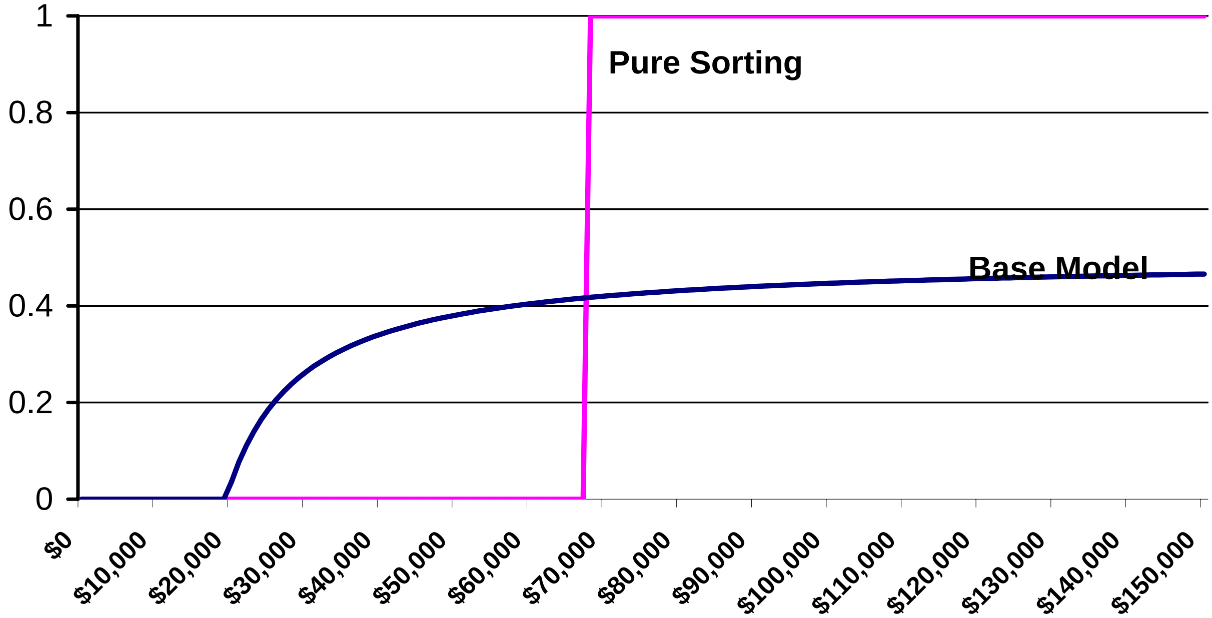
Note: 100,000 workers; truncated normal distribution; mean=40,000; w_L=$0, w_U =$150,000

Figure 12
Wage Distribution in Green City: Fatter Tails



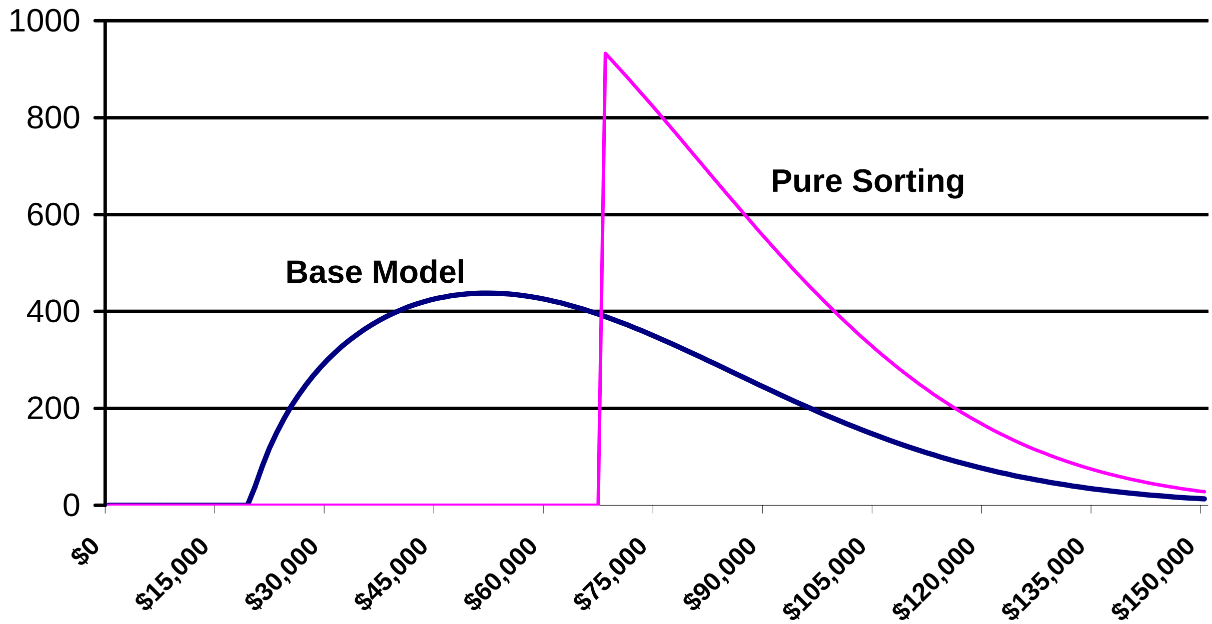
Note: 100,000 workers; truncated normal distribution; mean=40,000; w_L=$0, w_U =$150,000

Figure 13
Equilibrium Cutoff Values, c*(w): Various Models



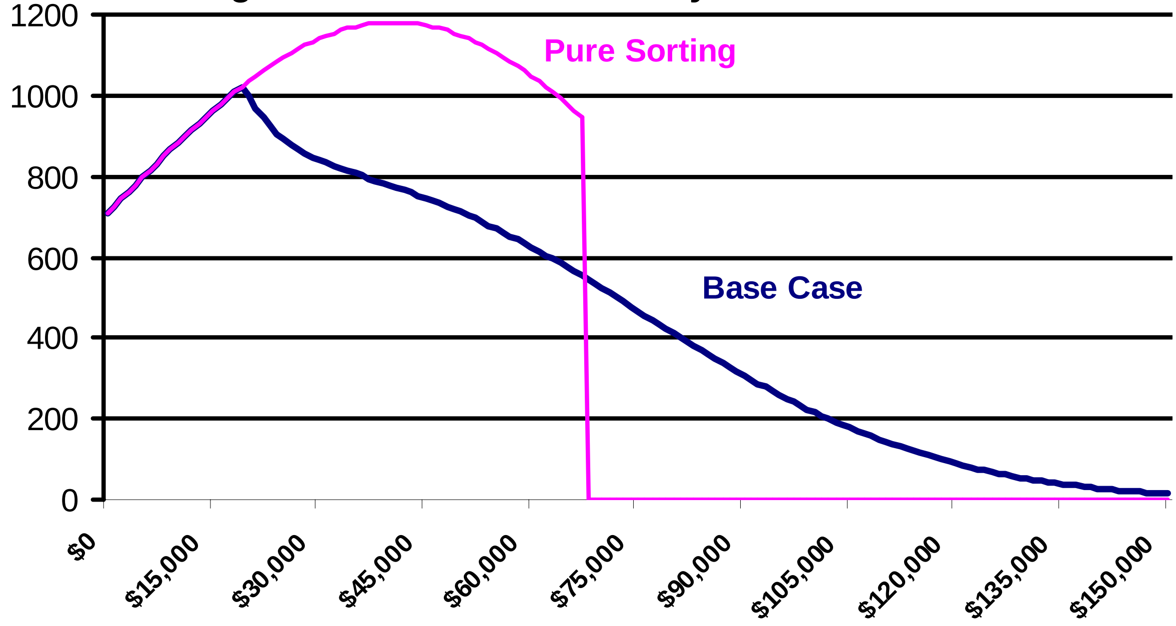
Note: 100,000 workers; truncated normal distribution; mean=40,000; w_L=$0, w_U =$150,000

Figure 14
Wage Distribution in Red City: Various Models



Note: 100,000 workers; truncated normal distribution; mean=40,000; w_L=$0, w_U =$150,000

Figure 15
Wage Distribution in Green City: Various Models



Note: 100,000 workers; truncated normal distribution; mean=40,000; w_L=$0, w_U =$150,000

Figure 16: Identifying Superstar Cities – house price growth versus unit growth

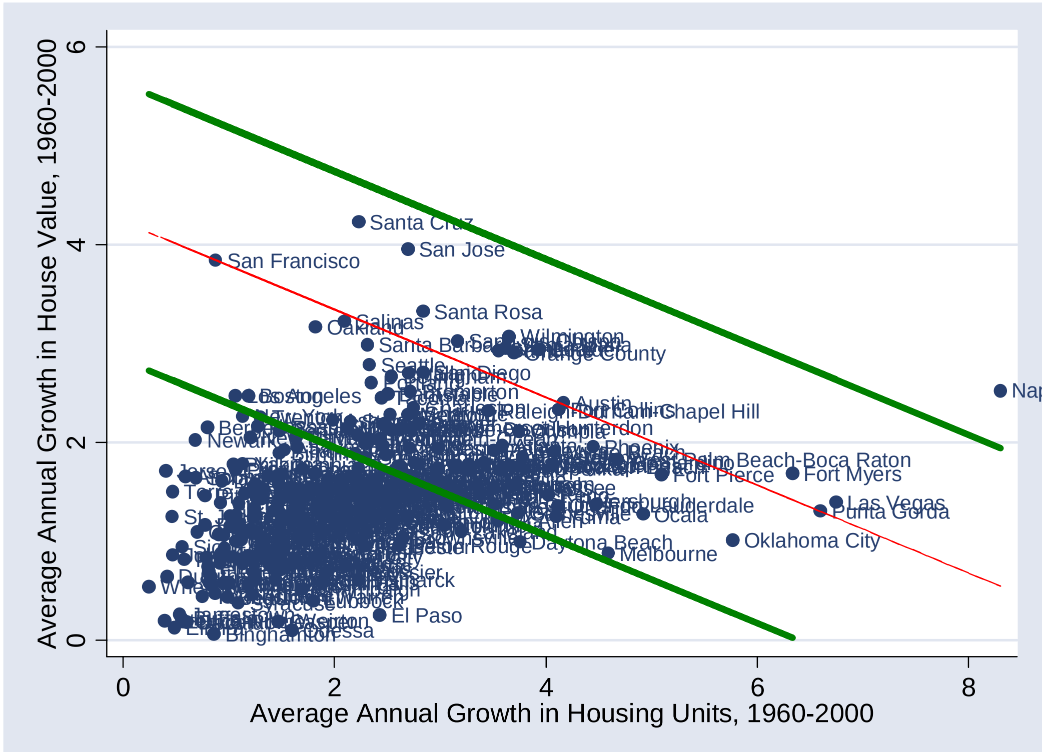


Figure 17

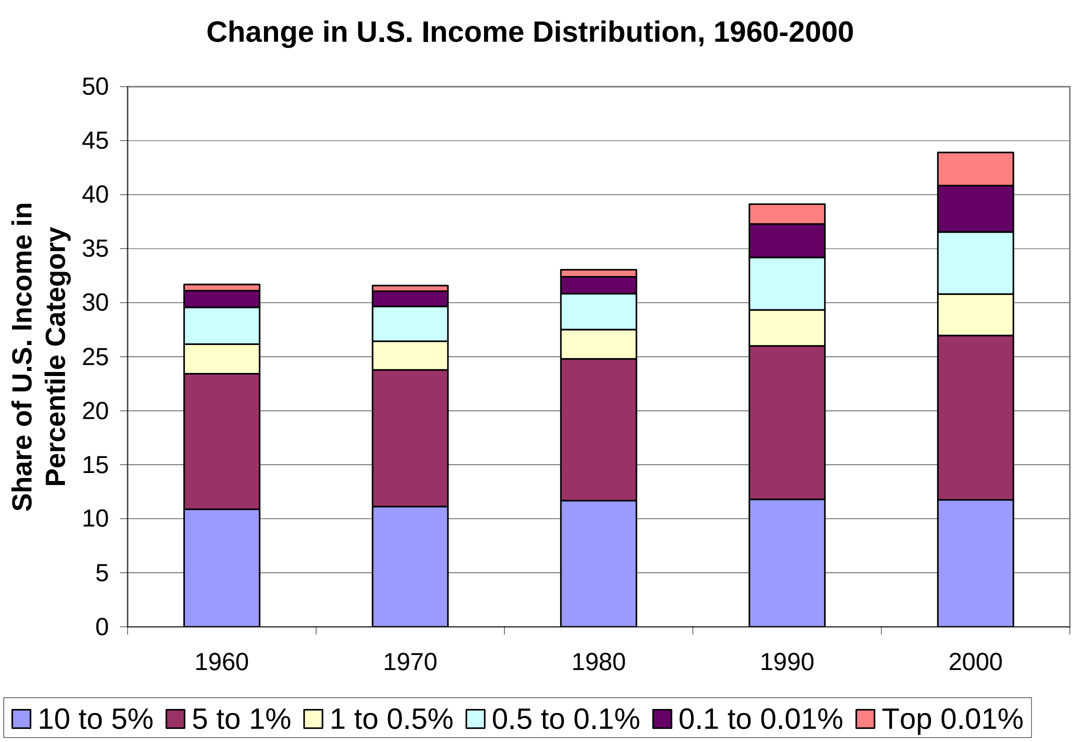


Figure 18: The Evolution of the National Income Distribution

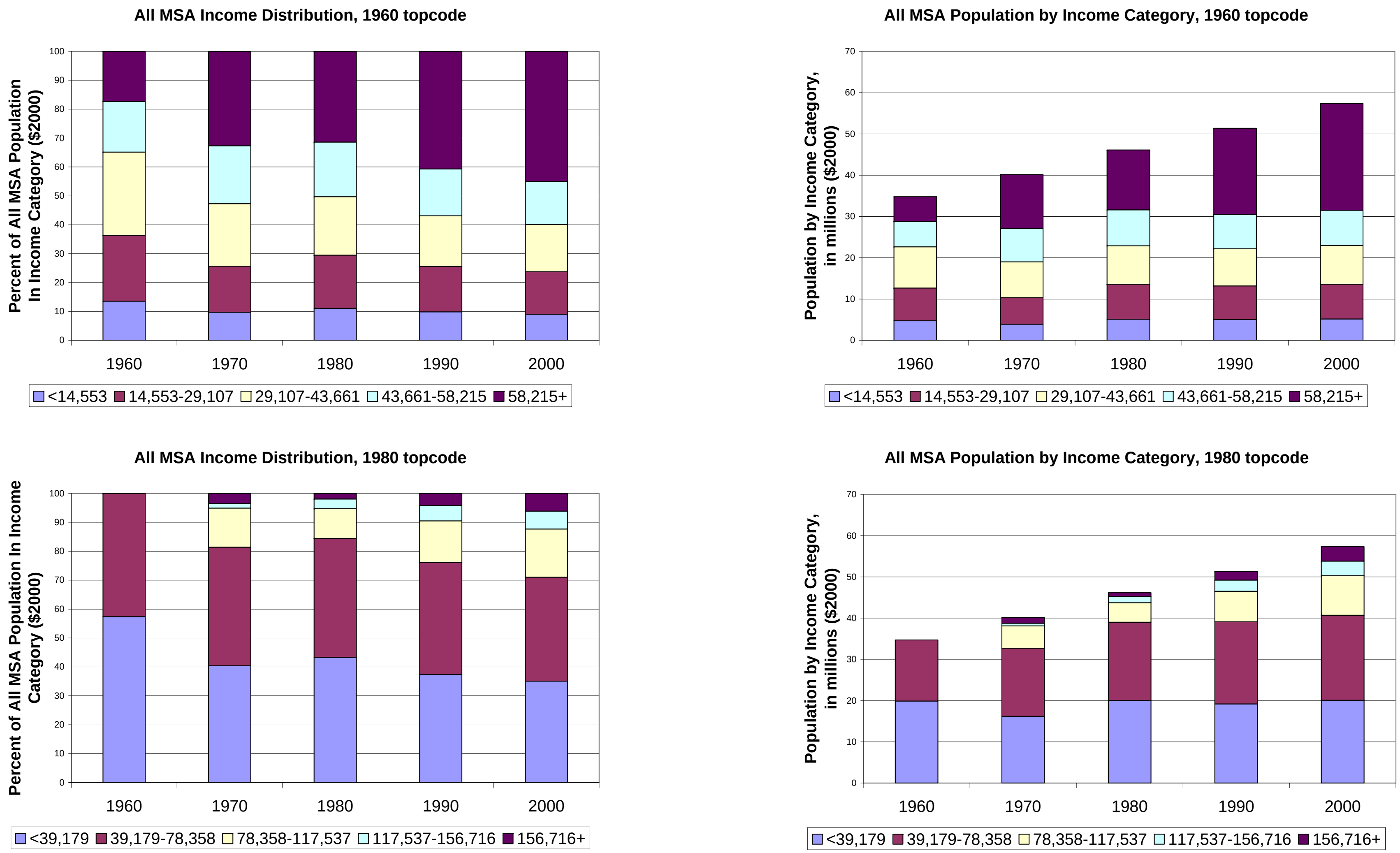


Figure 19: San Francisco (big price growth) gains rich, loses poor

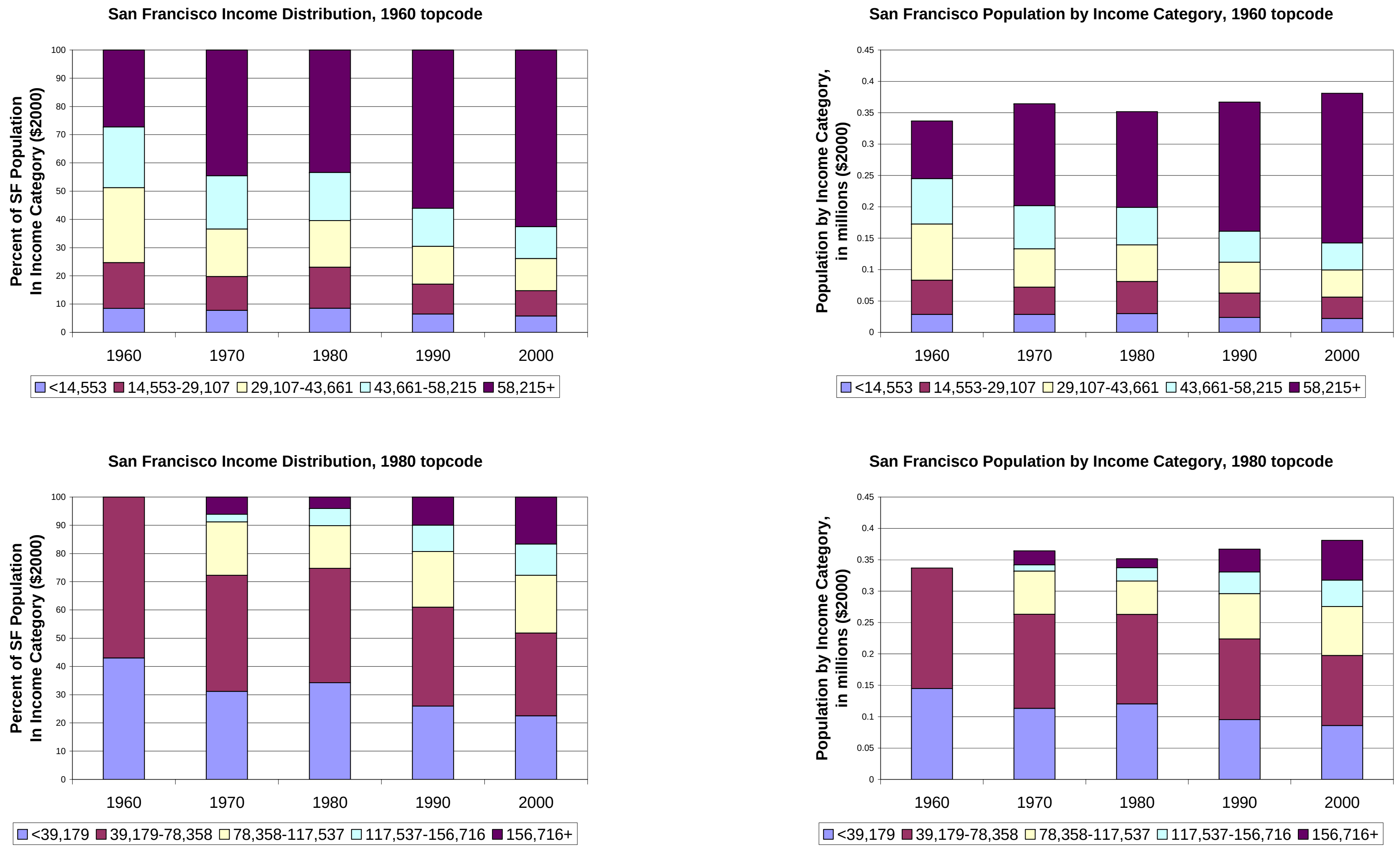


Figure 20: Las Vegas (big unit growth) gains rich and poor, shares stay constant