What Accounts for Growing House Price and Income Dispersion Across Markets: Productivity, Sorting, or Both?

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

Christopher 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

Draft of November 20, 2007

This paper was written for the NBER Economics of Agglomeration Conference to be held in Cambridge, MA, on November 30, 2007-December 1, 2007. Gyourko and Sinai thank the Research Sponsors Program of the Zell/Lurie Real Estate Center at The Wharton School for financial support. Mayer thanks the Milstein Center for Real Estate at the Columbia Business School for its support.

I. Introduction

One of the most striking changes in the American socio-economic landscape since World War II involves the skewing of income growth, with the right tail of the national income distribution getting much thicker over time (Piketty and Saez (2003)). Less well documented is the fact that families also have become increasingly spatially segregated by income, not only within metropolitan areas, but across them. In addition, real house prices in metropolitan areas (MSAs) such as San Francisco, Boston, and New York have appreciated at rates well above the national average over the post-war period, resulting in these markets being affordable to only the highest income families.

These stylized facts are associated with two very different patterns of urban success. One involves very high house price growth with relatively little population growth, with the other success story pairing strong population expansion with mild house price appreciation. San Francisco and Las Vegas are the canonical examples of each type. As the national population in metropolitan areas more than doubled (by 107 percent) between 1950 and 2000, the San Francisco primary metropolitan statistical area (PMSA) absorbed a disproportionately small share of that growth, expanding its population by less than 60 percent. Moreover, in excess of one-third of its net increase of 635,000 people over the entire post-war era occurred in the 1950s. In addition, when real house prices (measured in constant $2000 dollars throughout the paper) began to spike in San Francisco after 1970, increasing by 271 percent from a 1970 mean of $147,496 to $547,206 in 2000, its population growth slowed. Moreover, there was a tripling from 10 to 31 percent of the share of the richer families earning over $110,000 (also in constant 2000 dollars) in its PMSA between 1970-2000. By contrast, Las Vegas' population grew by an astounding 2,510 percent (or 1.5 million people) over the same time period. While the bulk of the population expansion in this metropolitan area occurred over the past three decades, it has experienced very modest real house price growth that is well below the national average for the same time period. And, its income distribution has skewed right much less than San Francisco's or the nation's.

What accounts for these remarkably different success stories? Basic price theory tells us that consistently high prices require some limits on supply, not merely strong demand. After all, if land were plentiful and homebuilders could supply new units whenever prices rose sufficiently above production costs to provide them a competitive return, San Francisco's consistently high price appreciation could not have occurred. Others have studied supply side constraints, and there is no doubt that many localities have become expert at imposing a myriad of hurdles that raise the cost of developing new housing (Glaeser and Gyourko (2003), Glaeser, Gyourko and Saks (2005a,b); Gyourko, Saiz and Summers (2007); Saks (2005)).

However, our interest is not in accounting for the different supply sides of local housing markets. It is on the growing spatial dispersion of house prices and incomes, both across and within markets. This relationship clearly is endogenous, and the first part of the paper documents the nature of this dispersion. The basic facts themselves are interesting, as suggested above by the stark contrast between San Francisco and Las Vegas. We then ask what can account for this relationship. Two basic explanations are proposed and analyzed. One is differences in urban productivity due to agglomeration economies—a mainstay of urban economics. The other is a preference-based explanation that relies on sorting, with the rich ultimately outbidding others for the scarce slots available in supply-constrained metropolitan areas. The evidence suggests that pure sorting is at least partially responsible for the urban outcomes we see, but it also is clear that much more work is needed to pin down the relative contributions of these two basic factors.

II. Stylized Facts on the Growing Dispersion in House Prices and Incomes

A. House Price and Income Growth

We use and discuss a variety of data from the U.S. decennial censuses, aggregated to the level of the metropolitan area, which corresponds to the local labor market. We use a sample of 280 such areas that had populations of at least 50,000 in 1950 and are in the continental United States.[1] Information on the distribution of house values, family incomes, population, and the number of housing units were collected.

Since the definitions of metro areas change over time, we use one based on 1999 county boundaries to project consistent metro area boundaries forward and backward through time.[2] Data were collected at the county level and aggregated to the metropolitan statistical area (MSA) or primary metropolitan statistical area (PMSA) level in the case of consolidated metropolitan statistical areas. Data for the 1970-2000 period are obtained from GeoLytics, which compiles long-form data from the decennial *Censuses of Housing and Population*. We hand-collected

---

[1] Thirty-six areas with populations under 50,000 in 1950 were excluded from our analysis because of concerns about abnormal house quality changes in markets with so few units at the start of our period of analysis. Those MSAs are: Auburn-Opelika, Barnstable, Bismarck, Boulder, Brazoria, Bryan, Casper, Cheyenne, Columbia, Corvallis, Dover, Flagstaff, Fort Collins, Fort Myers, Fort Pierce, Fort Walton Beach, Grand Junction, Iowa City, Jacksonville, Las Cruces, Lawrence, Melbourne, Missoula, Naples, Ocala, Olympia, Panama City, Pocatello, Punta Gorda, Rapid City, Redding, Rochester, Santa Fe, Victoria, Yolo, and Yuma. That said, none of our key results are materially affected by this paring of the sample. Similar concerns account for our not using data from the first *Census of Housing* in 1940 in the regression results reported below. (All individual housing trait data from the 1940 census were lost, so we cannot track any trait changes over time from that year.) However, we did repeat our MSA-level analysis over the 1940-2000 time period. While the point estimates naturally differ from those reported above, the magnitudes, signs, and statistical significance are essentially unchanged. Finally, the New York PMSA is missing crucial house price data for 1960, and is excluded from the analysis reported below. The census did not report house value data for that year because it did not believe it could accurately assess value for cooperative units, the preponderant unit type in Manhattan at that time.

[2] We use definitions provided by the Office of Management and Budget, available at http://www.census.gov/population/estimates/metro-city/90mfips.txt. One qualification is that in the case of NECMAs, the entire county was included if any part of it was assigned by the OMB.

1950 and 1960 data from hard copy volumes of the *Census of Population and Housing*. Both sources are based on 100 percent population counts. All dollar values are converted into constant 2000 dollars.[3]

In each data set, we divide the distribution of real family incomes into five categories that are consistent over time. The income categories in the original Census data change in each decade. We set the category boundaries equal to 25, 50, 75, and 100 percent of the 1980 family income topcode, and populate the resulting five bins using a weighted average of the actual categories in $2000, assuming a uniform distribution of families within the bins. Since 1980 had amongst the lowest topcode in real terms, using it as an upper bound reduces miscategorization of families into income bins. We call a family "poor" if its income is less than $39,179 in $2000. "Middle-poor" are those families with incomes between $39,179 and $78,358, "middle" income families have incomes between $78,359 and $117,537, "middle-rich" families lie between $117,538 and $156,716, and "rich" families have incomes in excess of the 1980 real topcode of $156,716.

Using these data, we begin by detailing the remarkable dispersion – and even skewness – across MSAs in house price and income growth over the 1950 to 2000 period. Figure 1A plots the kernel density of average annual real house price growth between 1950 and 2000 for our sample of 280 metropolitan areas. The tail of growth rates above 2.6 percent is especially thick and the distribution is right-skewed. Table 1, which lists the average real annual house price growth rate between 1950 and 2000 for the ten fastest and ten slowest appreciating metropolitan areas out of the 50 MSAs with populations of at least 500,000 in 1950, documents that the

[3] We also use some data for 1940. Population and housing unit data for that year area based on 100 percent counts, but housing values 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.

dispersion seen in this figure is not an artifact of a few areas that were small initially and then experienced abnormally rapid price growth.[4]

These annual differences in house price growth rates compound to very large price gaps over time even within the top few markets. For example, San Francisco's 3.5 percent annual house price appreciation implies a 458 percent increase in real house prices between 1950 and 2000, more than twice as large as seventh-ranked Boston at 212 percent, which itself still grew 50 percent more than the sample average of 132 percent for the 50 most populous metropolitan areas.[5] Figure 2A, which plots a kernel density estimate of the 280 metropolitan area average house values in 1950 and 2000, shows that skewness has increased over the last 50 years, with a relative handful of markets ending up commanding enormous price premiums. Figure 2B normalizes the means and standard deviations of the 1950 and 2000 house value distributions so they are equal and plots them against each other. In 2000, the right tail of the MSA house value distribution extends to four times the mean, more than twice the highest MSA from the right tail of the 1950 Census. The left tail ends at about half the mean in both years, although it is slightly more skewed in the 2000 Census.

There also is long-run persistence in the markets which exhibit above-average price growth. Across the two 30-year periods from 1940-1970 and 1970-2000, average annual percentage house price growth has a positive correlation of about 0.3. The root of this latter result can be seen in Table 2, which reports the transition matrix for MSAs ranked by their average real house price growth rates computed over the two 30-year periods of 1940-1970 and

---

[4] A complete list of house price appreciation rates by metropolitan area, along with 1950 and 2000 mean housing prices is reported in Appendix 1.

[5] It is worth emphasizing that the extremely high appreciation seen in the Bay Area, southern California, and Seattle markets is not restricted to the past couple of decades. The top five markets in terms of annual real appreciation rates between 1950-1980 are as follows: (1) San Francisco, 3.65 percent; (2) San Diego, 3.49 percent; (3) Los Angeles, 3.20 percent; (4) Oakland, 2.99 percent; and (5) Seattle, 2.88 percent.

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.[6]

Figure 1B then documents that income growth rates also exhibit wide dispersion across MSAs over the long-run. This kernel density plot of average annual real income growth over the 1950 to 2000 period shows that growth rates range from 0.8 percent per year to 3.1 percent. And, the distribution also evinces some right-skew.

That the right tail of the national income distribution has indeed been getting thicker over time is confirmed in Figure 3, which reports data from Saez (2004) on the share of U.S. income by population percentile over time. The tax return data Saez (2004) uses provides a very clear picture of changes at the high end of the income distribution. The share of income held by the very top percentiles 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 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. While the income data reported in the decennial censuses are not near as fine or detailed that available to Saez (2004), it does a reasonably good job of matching the skewing over time observed in the IRS data.

[6] Over shorter horizons such as a decade, MSAs can experience large price swings. In fact, the correlation in house price appreciation rates across decades often is negative.

B. A Case Study in Contrast: San Francisco and Las Vegas

Over the last 50 years, the U.S. has experienced growth in the absolute number, population share, and income share of high-income households (Autor *et al* (2006), Piketty and Saez (2003), Saez (2004)). The left panel of Figure 4 shows that the aggregate distribution of family income across all MSAs in the U.S. has been shifting to the right in real dollars as the right tail of the income distribution has grown much faster than the mean. The right panel of Figure 4 then displays the evolution of the number of families in each of the income bins. Most of the growth in the number of families was among those earning more than the $78,358 median value for our sample.

These changes in the national high-income share were accompanied by very disparate patterns at the metropolitan area level. San Francisco experienced low levels of new construction and high house price growth (Figure 5). Between 1950 and 1960, the San Francisco primary metropolitan statistical area (PMSA) expanded its population by about 48,000 families. Over the subsequent *four* decades, San Francisco grew by only 44,000 families, with two-thirds of that growth taking place between 1960 and 1970. Real house prices spiked in San Francisco after 1970, growing between 3 and 4 percent per year between 1970 and 1990, about 1.5 percentage points above the average across all MSAs, and 1.4 percent per year between 1990 and 2000, almost one percentage point above the all-MSA average. By contrast, over the same time period, Las Vegas saw explosive population growth, expanding from fewer than 50,000 families in 1960 to the size of the San Francisco PMSA by 2000 (Figure 6). Yet, it experienced modest real house price growth that was well below the national average.

Note that San Francisco's high-income share grew disproportionately. San Francisco, which always had relatively more rich families and fewer poor families than Las Vegas, became even more skewed toward high income families between 1960 and 2000. Since the number of families in the San Francisco MSA did not grow by much, the MSA actually experienced an increase in the number of rich families and a reduction in the number of lower income ones. In fact, only the richest groups, with incomes of $78,358 and above, increased their share of the number of families in the San Francisco MSA.

In stark contrast, the overall income distribution in Las Vegas did not keep up with the nation (left panel of Figure 6), leaving that metropolitan area progressively more poor relative to both San Francisco and the U.S. metropolitan area aggregate. The large numbers of new families in Las Vegas were both rich and poor, leading to substantial growth in the number of families across Las Vegas' income distribution. Relative to the national income distribution, however, the growth in Las Vegas was skewed towards poorer families.

C. A Broader Look at the Data Across All Metropolitan Areas

C.1. Categorizing Markets

Of course, the evidence for two specially chosen markets need not generalize, so we now turn to data across all metropolitan areas. We begin by documenting that there has been a marked change in the relationship between housing price and housing unit growth for the high price appreciation markets. Table 3 reports the results from regressing at the MSA level the decadal growth in the number of housing units on the long-run growth in house price, allowing a different intercept and slope for those areas in the top quartile of the price appreciation distribution. More specifically, we estimate

(1) $\%\Delta H_{i,t} = \alpha + \beta^{*}\%\Delta P_i + \gamma^{*}(TopQuartile_i) + \delta^{*}(\%\Delta P_i{*}\ TopQuartile) + \varepsilon_{i,t}$,

where $\%\Delta H_{i,t}$ is the percentage change in housing units in metropolitan *i* during decade *t*, $\%\Delta P_{i,}$ is the percentage house price growth in metropolitan area *i* during between 1960-2000, and TopQuartile is a dummy indicator for whether the metropolitan area is among the top quartile of areas in terms of house price appreciation over the 1960-2000 period.

These results show that the price growth/unit growth relationship for the top quartile of the price appreciation distribution essentially has disappeared between the 1960s and the 1990s. 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 the 1980s, flat over the decades. The MSAs in the top quartile in terms of price appreciation start out in 1970 with a slightly less positive correlation than for the lower 75 percent (11.12 – 3.12 = 8.0 correlation). By the 1970s, however, the highest price growth markets already are in negative territory (17.18 – 18.14 = -0.96), while there still is a large positive relationship between long-run price growth and housing unit production for the other metropolitan areas. The negative correlation for the top quartile increases over time, to -3.62 in the 1980s and -3.89 in the 1990s. Thus, it is not just in San Francisco that that housing unit growth is no longer strongly positively correlated with price appreciation. It is generally true for the markets with the highest long-run price appreciation.

We now turn to other work we have done (Gyourko, Mayer and Sinai (2006)) to classify markets as more like San Francisco or Las Vegas. San Francisco is termed a 'superstar' market in that research. For our purposes here, it is a market which is in high demand and in which something prevents the development of many new homes.[7] Thus, house price growth is very high, but housing unit growth is not.

[7] That something could be a natural constraint such as an ocean or a man-made constraint in the form of binding growth controls on housing development.

Because we do not observe the true state of demand and the literature does not provide high quality estimates of the elasticity of supply, the following two measures are combined to determine whether a market is a 'superstar'. First, a market is classified as in high demand if the sum of its housing unit and housing price growth is above the sample median for the relevant period of analysis. Second, a metropolitan area is defined to have a low elasticity of supply if its ratio of housing price growth to housing unit growth is at or above the 90th percentile of the distribution for all metropolitan areas over the relevant period of analysis.

In the regression results reported below, each of these measures is constructed using data from the two decades prior to the year for which a superstar designation is made. Thus, the status of each metropolitan area is classified from 1970-2000, with 1970 being the first year because the underlying data begin in 1950.[8] Figure 7 documents the outcome of this methodology for the most recent period—using 1980-2000 data to determine superstar status in 2000. Average real annual house price growth between 1980-2000 is on the y-axis, with housing unit growth over the same two decades on the x-axis. The single downward-sloping line reflects the boundary between markets with a sum of price and unit growth above the sample median across all our MSAs for 1980-2000. Any metro area lying below that line is a relatively low demand place by definition. The left-most and steepest positively-sloped line from the origin captures the elasticity of supply at the 90th percentile of the distribution of the ratio of price growth to unit growth. For this twenty year period, the MSA at the 90th percentile has a ratio of real annual house price growth to unit growth above 1.7. The right-most and flattest positively-

[8] Because the empirical task here is to document whether equilibrium relationships implied by our model exist in the data rather than to identify causal mechanisms for why a place becomes a superstar, the use of lagged data is not driven by endogeneity concerns (which these lags would not deal with effectively in any event). Rather, we wish to be able to classify superstar status in the most recent census data from the year 2000, and we suspect that any relationship between income segregation and house price effects occur after the superstar market has 'filled up'.

sloped line from the origin reflects the inverse of the 90th percentile ratio value (i.e., 1/1.7, or 0.59).

Cities in the region marked ‘A’, which is both above the boundary determining low demand status and above the boundary marking significant inelasticity of supply, include many coastal markets including San Francisco, New York, and Boston that have experienced very strong house price appreciation (indicating high latent demand), but little supply response in terms of new construction over the past two decades. The other markets in relatively high demand areas are divided into two groups for the purposes of the empirical analysis below. What we term ‘Non-superstars’ are the metropolitan areas in the ‘C’ range, which include markets with relatively high housing unit production and relatively low housing price growth. These high demand markets, which include Las Vegas and Phoenix, build sufficient new housing to satisfy demand so that real price growth is low. The remaining high demand markets are in-between the Superstars and Non-Superstars and lay in the ‘B’ range in Figure 7. They have experienced relatively high demand, and have both built at least a modest amount of new units and experienced a moderate amount of real house price appreciation. The final set of metropolitan areas are in low demand, and lay in the ‘D’ region below the negatively-sloped line in Figure 7.

This nonlinear categorization is useful because it allows us to observe how MSAs evolve over time. It seems natural that metropolitan areas could become more inelastically supplied as they grow and begin to “fill up” in the face of geographic constraints or politically-imposed restrictions on development. This would appear as a market moving over time from area C to B to A in Figure 7. We do observe such an evolution over time. In 1980, only San Francisco and Los Angeles clearly qualified as superstars in 1980, with the other markets filling up over time.

C.2. Income Distributions Across MSAs

Table 4 reports regression results on the link between income distributions and house prices using our categorization of cities into superstar versus non-superstar status. We start with the cross sectional relationship and then examine the data over time. Specification (2) investigates whether a typical superstar market's household income is skewed to the right of the U.S. income distribution, as we saw was the case for San Francisco. More specifically, we estimate the following regression for MSA *i* in year *t*:

$$(2)\ \frac{\#\text{ in Income Bin}_{\text{yit}}}{\#\text{ of Households}_{\text{yit}}} = \beta_1(\text{Superstar}_\text{i}) + \beta_2(\text{Non - superstar}_\text{i})$$
$$+\beta_3(\text{Superstar}_\text{it}) + \beta_4(\text{Non - superstar}_\text{it}) + \gamma_1(\text{Low Demand}_\text{i}) + \gamma_1(\text{Low Demand}_\text{it}) + \delta_t + \varepsilon_{it}$$

Essentially, this regression relates the share of a MSA's families that are in each income bin to its superstar status, and controls for total demand.[9]

The first column of the top panel of Table 4 is based on a pooled cross section of 1,116 MSA × year observations.[10] This regression treats superstar status as a (non-exclusive) fixed MSA characteristic, including indicator variables for whether the MSA ever was a superstar over the 1970-2000 period, whether it is ever in the non-superstar range, whether the MSA ever moved inside the low-demand area, and time dummies. The intermediate, high demand MSAs from region *B* are the excluded category in all the regressions reported in Table 4.

The difference in income distribution between superstars and all other MSAs is pronounced. MSAs that ever were superstars have a 2.5 percentage point greater share of their families that are in the rich category relative to the excluded high-demand cities (row 1, column 1). This effect is largest at the high end of the income distribution and declines in magnitude as incomes fall. For example, as reported in square brackets in row 1, the high income share of

---

[9] See Appendix Table 2 for summary statistics on all variables used in these regressions.
[10] This represents 279 MSAs in each census year from 1970-on.

superstar MSAs is about 83 percent more than the 3 percent share rich for the average MSA that is not a superstar. The share of the next-highest income category is 69 percent greater in superstars relative to the average of other MSAs, and 34 percent higher in the middle category. Markets that have ever been superstars also have a nearly 9 percentage point lower share of poor families (row 1, column 5), almost 21 percent less than the other MSAs.

Non-superstar cities appear similar to the in-between group (row 2). Those coefficients are relatively small and do not exhibit a clear pattern. Low-demand MSAs are less high-income and poorer relative to all of the high demand categories of MSAs, although the magnitudes are modest (row 3).

The second panel of Table 4 adds time-varying superstar, non-superstar, and low demand indicator variables to the previous specifications. Prior to becoming superstars, MSAs that eventually will become a superstar are richer on average, with a 1.3 percentage point greater share rich and a 7.1 percentage point lower share poor (row 1 of panel 2). When these areas are actually in the superstar region, their share rich goes up by an additional 2.8 percentage points and their share poor declines by a further 4.1 percentage points (row 4 of panel 2). As a baseline, superstar cities have a 43 percent higher share rich, declining monotonically to 17 percent lower share poor, than other MSAs. After their transition to superstar status, these MSAs have an additional 80 to 90 percent greater share of the top two income groups and an 8 to 10 percent lower share of the bottom two income categories. As before, this pattern of results is robust to adding a host of controls for potential unobservables, such as MSA fixed effects, differential time trends for superstars vs. not, or separate year dummies for superstars/non-superstars/low-demand MSAs.

In sum, we have established that both incomes and house prices have become increasingly skewed not just at the national level, but across metropolitan areas. While it is clear that restrictions on supply must be present for high prices to result in the first place, we take supply conditions as given and now consider the underlying forces that could account for the relationship between the growing spatial income and price dispersion just documented. We first draw on urban economics to propose dynamic agglomeration economies as an explanation and then consider a less traditional story that relies on preference-driven sorting in the presence of housing supply constraints in attractive labor market areas.

III. Urban Productivity Differences and the Skewing of House Prices and Incomes

The standard spatial equilibrium model in urban economics developed by Rosen (1979) and Roback (1982) suggests that house price differences across markets are a function of amenity and wage (productivity) differentials. Much has been written in urban economics and the broader growth literature about agglomeration effects and the potential for increasing returns in some markets that conceivably could causally link the endogenous relationship between house price and income growth documented above. Indeed, Lucas (1988) explicitly notes that cities are a natural laboratory in which to test growth models involving some type of productivity spillover. Glaeser, et. al. (1992, 1995) and Henderson, et. al. (1995) soon followed with analyses of dynamic agglomeration economies that extend across time. While there is much debate about the precise nature of the spillovers involved, there is widespread agreement that there are long-run effects from urban agglomerations.[11]

Much of the more recent agglomeration research starts with the basic fact that skilled cities growth more quickly, where growth is measured in terms of quantities such as population

[11] See Rosenthal and Strange (2003) for an extensive and excellent review of the urban agglomeration literature.

or employment. For example, Glaeser and Saiz (2003) document that at the metropolitan area level, a 1 percentage point higher population share for college graduates is associated with about a 0.5 percentage point higher decadal population growth rate. Similarly, Shapiro (2006) shows that from 1940-1990, a ten percent higher concentration of college graduates is associated with a 0.8 increase in future employment growth (also at the metropolitan area level). Within the spatial equilibrium framework pioneered by Rosen (1979) and Roback (1982), the relevant question is whether this empirical link between education and population (or employment) growth is the result of rising productivity or amenities.

Ever since Rauch (1993), we have known wages in a market rise with the skill level of that market, holding constant individual worker skills. Moretti (2004) recently confirmed Rauch's basic correlation, identifying human capital externalities via an instrumental variables estimation that uses the presence of land grant universities as an instrument that proxies for human capital in the area, but is plausibly exogenous to wages.[12]

Urban wage premia do appear to be relatively large. Glaeser and Mare (2001) estimate them to be on the order of 20-35 percent for workers in larger cities. Those authors also find that long-term residents in bigger cities earn a premium over new arrivals, and that when long-term workers leave their city for another, their wages in the new location are higher the larger the size of their previous market.[13]

Glaeser and Saiz (2003) and Shapiro (2006) both conclude that the link between education and metro area population/employment growth largely is due to productivity, with amenities playing a smaller role. Going beyond the reduced form OLS estimation standard in the

[12] That said, there is some debate about the strength of such externalities, with Acemoglu and Angrist (2000) finding small effects, but at the state level. See Moretti (2003) for a recent review of the literature on human capital externalities in cities.
[13] There is research on the firm side, too. For example, Henderson (1997) shows that concentrations of own-industry employment have measureable impacts on growth many years into the future.

literature, Shapiro (2006) calibrates a neoclassical urban growth model and estimates that about 60 percent of the impact of a higher local population share of college graduates on metropolitan area employment growth is due to productivity, as reflected in wage growth. This does leave room for improvements in the quality of life to play a role, too, and they appear related to 'consumer city'-type attributes as reflected in various local cultural traits (Glaeser, Kolko and Saiz (2001)).

While there is much evidence consistent with the presence of dynamic spillovers, the agglomeration literature has not focused on the relationship between house price and income dispersion. However, it is not hard to see a natural link. If productivity differences across markets are growing, then the higher wages that result in the most productive agglomerations should be capitalized into land values (and thus, house price) in a market such as San Francisco which constrains the supply of housing.

This story requires a very high rate of value growth, consistency in the location of agglomeration benefits in areas with inelastic supply slides to their housing markets, and that firms will not move to cheaper places. It certainly is not hard to understand how difficult it would be to recreate somewhere else the production or consumption externalities that lead to increasing returns. In the short-run, this probably is impossible, although it seems more open to debate whether we should expect mobility of people and firms to be high over half-century long periods. In addition, it is not immediately clear why such productivity would tend to occur in supply-constrained markets.

Leaving aside those questions, we now turn our focus to whether urban productivity differentials are growing sufficiently to account for the increases in house price and income dispersion that we observe in the data. The most direct effort to answer that question is by Van

Nieuwerburgh and Weill (2006). They do so in a dynamic, general equilibrium version of the Rosen-Roback model, in which they then run calibration exercises to see whether there has been enough growth in wage dispersion across labor markets to account for the growth in house price dispersion. More specifically, they investigate whether their model can match the increase in the coefficient of variation in house prices across markets between two steady states.

The model itself constitutes a meaningful contribution to urban economics in that it provides one of the first truly dynamic frameworks in which to analyze issues that are inherently dynamic.[14] However, it does lack certain features of urban environments, including local amenities, that urban economists long have thought important for understanding spatial variation. Essentially, homogeneous physical markets receive exogenous productivity shocks.

Van Nieuwerburgh and Weill (2006) chose parameter values to match an assumed steady state in 1975, and then allow for wage shocks that mimic the increase in wage dispersion between 1975 and 2004, holding constant the stringency of supply side constraints across markets. This exercise yields a very good match of the mean annual increase in house prices (e.g., 1.67% in their simulation versus 1.71% in the real data), as well as a tight fit of the increase in the coefficient of variation in house prices across markets. This statistic is too high in 1975, but they can replicate the increase (of 0.08) that occurred by 2004. This simulation also results in a good match of the growth of population in the productive places with higher wages. This latter result is a key implication of their model, as the movement of workers to high productivity markets to obtain the high wages therein is what helps drive up housing costs. Although the framework is dynamic, the essential insight of Rosen and Roback still holds—housing costs are

[14] Glaeser and Gyourko (2006) also have produced a dynamic model, but it is designed to investigate higher frequency movements in house prices.

the price one has to pay to access the productivity (and amenities in a more complete model) of a given labor market area.[15]

These results certainly are consistent with growing urban productivity differentials being the cause of the growing house price dispersion across labor market areas. However, they are not conclusive in proving causality. A careful review of the data indicates that the productive/high wage markets to which Van Nieuwerburgh and Weill's (2006) model predicts people should move include both San Francisco and Las Vegas. More generally, there is a mixture of both types of markets in the top wage quintile. We know from Section II.B that there was almost no net population growth in San Francisco during the time period covered by this calibration exercise. Thus, it appears that their model's ability to match the data is the result of it picking up much of the growing price dispersion from very high price (and price appreciation) coastal markets that have very little home building and population growth; analogously, it looks to be picking up much of the housing unit/population growth from large Sunbelt markets that have relatively low house price levels and that have experienced relatively little price appreciation. This suggests that it remains an open question whether the growing dispersion in house prices is being driven exclusively by underlying productivity shocks.

IV. Sorting and Supply Constraints As Explanations for the Spatial Skewing of House Prices and Incomes

While it makes good sense to presume that wage differences at least partially reflect productivity differences, Waldfogel's (2003) preference externality research raises the possibility that productive people agglomerate, not just that agglomerations make people more productive.

[15] There are a host of other results, ranging from the role of supply-side constraints to the change in the ratio of house prices to construction costs. We do not review those findings here, so as to stay focused on the relationship between the skewing of incomes and house prices across markets.

Importantly, in Waldfogel's framework, people may group together for various reasons that are not necessarily socially value-enhancing. Gyourko, Mayer and Sinai (2006) provide another preference-related explanation for such grouping, and it is to that work that we now turn.

That paper tries to account for the stylized facts documented in Section II with a model that assumes households vary by their incomes and tastes for two locations which differ by whether they have available land. Using the structure first established by Epple and Platt (1998), Gyourko, Mayer and Sinai (2006) further simplify the analysis by having only two cities and assuming that there is elastic supply in only one market. Their comparative statics do not depend upon the reasons for location preferences or the inelasticity of supply in the one market. All that is required is that some households prefer one city over the other and that there be some binding limit (natural or regulatory) on the supply of new housing units.

Sorting behavior, not urban productivity, is what accounts for the skewing of house prices and incomes across markets in this framework. Ultimately, the relatively rich with a preference for the market with an inelastic housing supply outbid the poor for the scarce slots. The model in Gyourko, Mayer and Sinai (2006) is not dynamic, but the comparative static results yield some interesting predictions that could help distinguish between productivity and preference reasons for the skewing of house prices and incomes across metropolitan areas. One is that it is the number of rich people nationally, not just locally, that should be correlated with the spatial skewing of prices and incomes. While an important task for this strand of research is to introduce realistic dynamics, the intuition is that skewing can continue and increase as long as the country keeps producing rich people, at least some of whom have a preference for the supply-constrained market. The urban productivity model does not predict any such relationship with national aggregates.

Gyourko, Mayer and Sinai (2006) report regression results consistent with this prediction. More specifically, they regress a proxy for the entry price of a home (they use the 10th percentile house value in each metropolitan area) on a set of indicators for superstar/non-superstar/low growth status that are also interacted with the national number of rich families. The specification also includes metro and year fixed effects, and is estimated over the 1970-2000 time period. Their findings imply that when the national number of rich families is ten percent higher, the gap in the 10th percentile house value between MSAs that are ever superstars and those 'in-between' markets is 1.1 percent greater (see column 1 of their Table 4).

Importantly, a similar pattern is observed when they use data on Census-designated places within metropolitan areas. In this case, each of the nearly 280 MSAs has its own pattern of growth in the number of rich families over four decades. The entry house price is 0.98 percent higher in superstar places relative to the 'in-between' places in the same MSA in the years when the metropolitan area's number of rich families is ten percent greater (see column 2 of their Table 4).[16]

Using similar regressions with the share of rich families as the dependent variable, Gyourko, Mayer and Sinai (2006) also confirm that the income distribution in superstar markets shifts to the right most when the number of rich families is largest. At the MSA level, a ten percent rise in the number of rich nationally is associated with a 0.31 percent greater rise in the share rich for superstar markets relative to 'in-between' markets in the same year. At the place level, a 10 percent rise in the MSA's number rich is associated with a 0.22 percentage point greater increase in the share rich for its superstar places.

[16] The underlying regression is similar to that run at the MSA level. That is, the 10th percentile house price for each place is regressed on the interaction of place superstar status and the number of rich families in the MSA. However, place fixed effects and MSA x year effects also are included to control for MSA-level dynamics

The findings of similar patterns for places within metropolitan areas is perhaps the most convincing evidence in Gyourko, Mayer and Sinai (2006) that something beyond urban productivity differentials is involved. The effect of local productivity growth on house prices should occur throughout the labor market area, even when agglomeration benefits accrue in the workplace. The correlations just discussed suggest that preference-based sorting can and does exist independently of agglomeration effects.

Another potentially useful way to discriminate between productivity and preference-based stories is to examine migrants. The sorting model suggests that rising incomes in a superstar market are due to an influx of highly productive, high-income workers, not gains in existing workers' productivity. Ideally, one would compare workers' wages before and after an exogenous move to a superstar city (Glaeser and Mare (2001)). Unfortunately, the Census is a cross section so preexisting wages are not observed, and truly exogenous moves are very hard to convincingly identify.

Thus, Gyourko, Mayer and Sinai (2006) compare the wage distributions of recent in-migrants and out-migrants in different markets. Their results confirm that the income distribution of recent movers into superstar markets is heavily weighted towards richer families. [See their Table 5.] Superstar MSAs also have a smaller fraction of rich out-migrants and a higher fraction of poor out-migrants, but the pattern generally is not as strong for out-migrants as for in-migrants. Still, the difference in the income distributions of in-migrants and out-migrants does vary considerably between superstar and non-superstar cities, with the difference in net in-migration being statistically significantly different in four out of five census years examined. While this certainly cannot be considered proof of the sorting model, there is no reason to believe that the income benefits of productivity growth should accrue disproportionately to recent in-migrants of superstar markets. As such, these results also suggest that something beyond urban productivity is occurring to generate the relationship between income and house price dispersion across markets.

Of course, this is not to say that there is no role for productivity. It certainly is the case that differential agglomeration effects exist across urban areas. The issue is whether urban productivity is the sole explanation for the variation in the data. We believe the answer to this question is 'no', but the empirical importance of the different explanations remains unresolved. Parsing this out is an essential task for future research that will not be easy, but is important for our understanding of urban markets.[17] The answer also has important implications for views on policy issues such as tax-based subsidies to homeownership. While economists can justify subsidies based on positive externalities involving better citizenship or improved outcomes for children (DiPasquale and Glaeser (1999) and Green and White (1997)), the case becomes harder if one believes that the high prices in America's coastal markets are due more to preference-based sorting combined with binding local regulation on home building than to productivity. Hence, this issue is much more than an intellectual curiosity for urban economists.

V. Conclusion

The growing dispersion in house price and income growth rates across locations is one of the most important stylized facts about metropolitan areas in America. The spatial sorting by income that it necessarily involves goes to the heart of how we live and organize ourselves socially. Whether these phenomena are due primarily to increasing value from amenities and productivity benefits or are the result of a growing number of high income families willing to

[17] Even if preference-based sorting explains the moves of the rich into markets like San Francisco, it is possible that once the rich agglomerate in that market, productivity then increases. Hence, the two forces may interact in various ways.

pay increasingly large amounts to live in a few supply-constrained markets is likely to have much to say about how many of us view this on-going situation.

This paper has documented the basic facts about the spatial distribution of house prices and incomes, and has outlined two possible explanations for the patterns we see in the data. Our review concludes that it is unlikely that urban productivity differentials are the sole causal force involved, and that the skewing of the income distribution nationally is interacting with binding supply-side constraints in certain (primarily coastal) markets to help generate the variation observed. Much work remains to be done to parse out the empirical contributions of the productivity and sorting explanations, but a roadmap for future progress is evident.

More generally, these changes in the nature of metropolitan America have profound implications for the evolution of urban areas. If the skewing and dispersion continues to grow, even large metropolitan areas could evolve into markets that are affordable only by the rich. In effect, an entire labor market area could have the income distribution of an exclusive resort. We do not know whether such a MSA is sustainable. Moreover, should public policy ensure that living in a particular city is available to all, or since superstar cities are like luxury goods, should we not care whether lower income households can buy into those markets any more than we care they can buy a Mercedes? These and other questions will provide fertile ground for thought and research by economists interested in urban agglomerations.

References

Acemoglu, Daron and Joshua Angrist (2000). "How Large Are Human Capital Externalies? Evidence from Compulsory Schooling Laws", *NBER Macroeconomics Annual*: 9-59.

Autor, David, Lawrence Katz, and Melissa Kearney (2006). "The Polarization of the U.S. Labor Market", *American Economic Review*, 96: 189-194.

DiPasquale, Denise and Edward Glaeser (1999). "Incentives and Social Capital: Are Homeowners Better Citizens?", *Journal of Urban Economics*, 45(2): 354-384.

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

Glaeser, Edward and Joseph Gyourko (2003). "The Impact of Zoning on Housing Affordability", *Economic Policy Review*, 9(2): 21-39.

_______________________________ (2006). "Housing Dynamics", NBER Working Paper No. 12787, December 2006.

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

Glaeser, Edward, Joseph Gyourko, and Raven Saks (2005b) "Why Is Manhattan So Expensive? Regulation and the Rise in House Prices", *Journal of Law and Economics*, 48(2): 331-370.

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 and D. Mare (2001). "Cities and Skills", *Journal of Labor Economics*, 19(2): 316-2001.

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

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

Green, Richard and Michelle White (1997). "Measuring the Benefits of Homeowning: Effects on Children", *Journal of Urban Economics*, 41(3): 441-461.

Gyourko, Joseph, Christopher Mayer and Todd Sinai (2006). "Superstar Cities", NBER Woring Paper No. 12355, June, 2006.

Gyourko, Joseph, Albert Saiz and Anita Summers (2007). "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index", *Urban Studies*, forthcoming.

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

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

Moretti, Enrico (2003). "Human Capital Externalities in Cities" in J.V. Henderson and J-F. Thisse (eds) *Handbook of Urban and Regional Economics*, Volume 4.

____________ (2004). "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross Section Data", *Journal of Econometrics*, 121(1-2): 175-212.

Piketty, Thomas and Emmanuel Saez (2003). "Income Inequality in the United States, 1913-1998", *Quarterly Journal of Economics* 118(1), 1-39.

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

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

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.

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

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

Saks, Raven (2005). "Job Creation and Housing Construction: Constraints on Metropolitan Area Growth", Finance and Economics Discussion Series 2005-49. Washington: Board of Governors of the Federal Reserve System.

Shapiro, Jesse (2006). "Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital", *Review of Economics and Statistics*, 88(2): 324-335.

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

Table 1: Real annualized house price growth, 1950-2000, Top and Bottom 10 MSAs with 1950 population>500,000

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

Population-weighted average of the 50 MSAs in this sample: 1.70

Table 2: 30-Year House Price Appreciation 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

Note: The underlying sample for this table includes only 129 metropolitan areas due to limitations on data available back to 1940.

Table 3: The Relationship Between High Long-Run Price Growth MSAs and the Change in the Number of Housing Units, by Decade

	1960s	***1970s***	***1980s***	***1990s***
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 x In Top Quartile	-3.12	-18.14	-15.35	-13.26
	(7.91)	(6.26)	(3.64)	(2.51)
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 4: The income distribution in superstar MSAs

	Left-hand-side variable: Share of MSA's families in income bin:				
	Rich	Middle-rich	Middle	Middle-poor	Poor
Cross-section:					
$Superstar_i$ [Relative to mean share]	0.025 (0.001) [0.833]	0.022 (0.001) [0.688]	0.042 (0.003) [0.339]	-0.004 (0.004) [-0.010]	-0.086 (0.007) [-0.208]
$Non\text{-}superstar_i$	0.005 (0.001)	0.003 (0.001)	0.002 (0.002)	-0.023 (0.003)	0.013 (0.006)
$Low\ Demand_i$	-0.008 (0.001)	-0.007 (0.001)	-0.010 (0.003)	0.007 (0.004)	0.017 (0.007)
Adj. R^2	0.442	0.621	0.377	0.178	0.214
Time-varying superstar/non-superstar status					
$Superstar_i$ [Relative to mean share]	0.013 (0.002) [0.433]	0.011 (0.002) [0.344]	0.035 (0.004) [0.282]	0.013 (0.005) [0.0325]	-0.071 (0.009) [-0.171]
$Non\text{-}superstar_i$	0.005 (0.001)	0.005 (0.001)	0.002 (0.003)	-0.022 (0.004)	0.010 (0.007)
$Low\ Demand_i$	-0.006 (0.001)	-0.006 (0.001)	-0.009 (0.003)	0.000 (0.004)	0.021 (0.007)
$Superstar_{it}$ [Relative to mean share]	0.028 (0.003) [0.903]	0.027 (0.002) [0.818]	0.017 (0.006) [0.135]	-0.030 (0.008) [-0.075]	-0.041 (0.015) [-0.100]
$Non\text{-}superstar_{it}$	-0.003 (0.002)	-0.006 (0.001)	-0.004 (0.004)	0.010 (0.005)	0.003 (0.009)
$Low\ Demand_{it}$	-0.003 (0.001)	-0.003 (0.001)	-0.003 (0.003)	0.015 (0.004)	-0.006 (0.007)
Adj. R^2	0.504	0.669	0.383	0.207	0.219
Mean of LHS: $Superstar_i$=0 [$Superstar_{it}$=0]	0.030 [0.031]	0.032 [0.033]	0.124 [0.126]	0.400 [0.402]	0.414 [0.409]

Notes: Number of observations is 1,116, for four decades (1970-2000) and 279 MSAs. Standard errors are in parentheses. All specifications include year dummies. $Superstar_{it}$ is equal to 1 when an MSA's ratio of real annual price growth over the previous two decades to its annual housing unit growth over the same period exceeds 1.7 (the 90th percentile) and the sum of price and unit growth over that period exceeds the median. $Superstar_i$ is equal to 1 for an MSA if $superstar_{it}$ is ever equal to 1. $Non\text{-}superstar_{it}$ is equal to 1 when the price growth/unit growth ratio is below 1/1.7, and $non\text{-}superstar_i$ is an indicator whether $non\text{-}superstar_{it}$ is ever 1. To control for MSA-demand, the top panel includes an indicator variable for whether the MSA's sum of annual price growth and unit growth over any 20-year period fell below the median in that period. The bottom panel includes that variable plus a time-varying variable for whether the sum of the growth rates over the preceding 20 years was below the median.

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

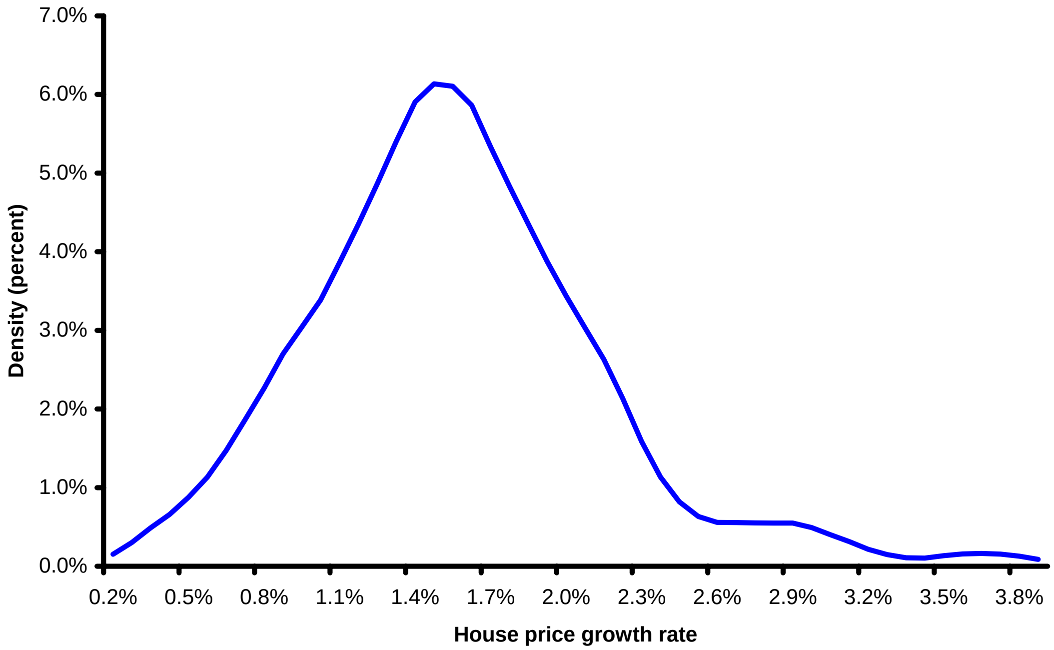


Figure 1B: Density of 1950-2000 Annualized Real Income Growth Rates
Across MSAs with 1950 population > 50,000

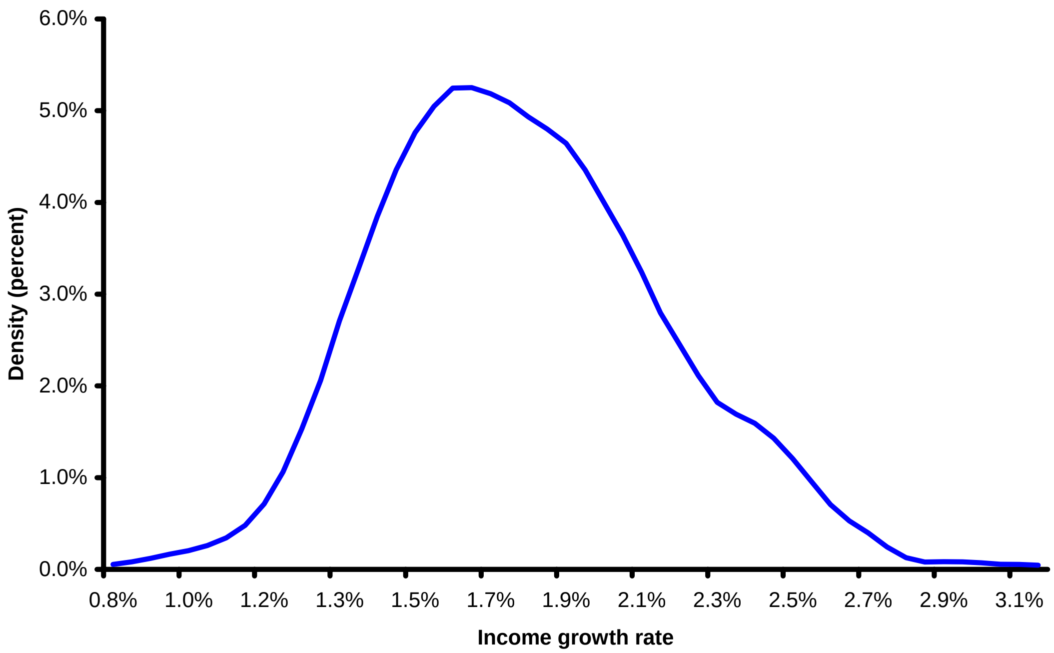


Figure 2A
Density of Mean House Values Across MSA's
1950 versus 2000

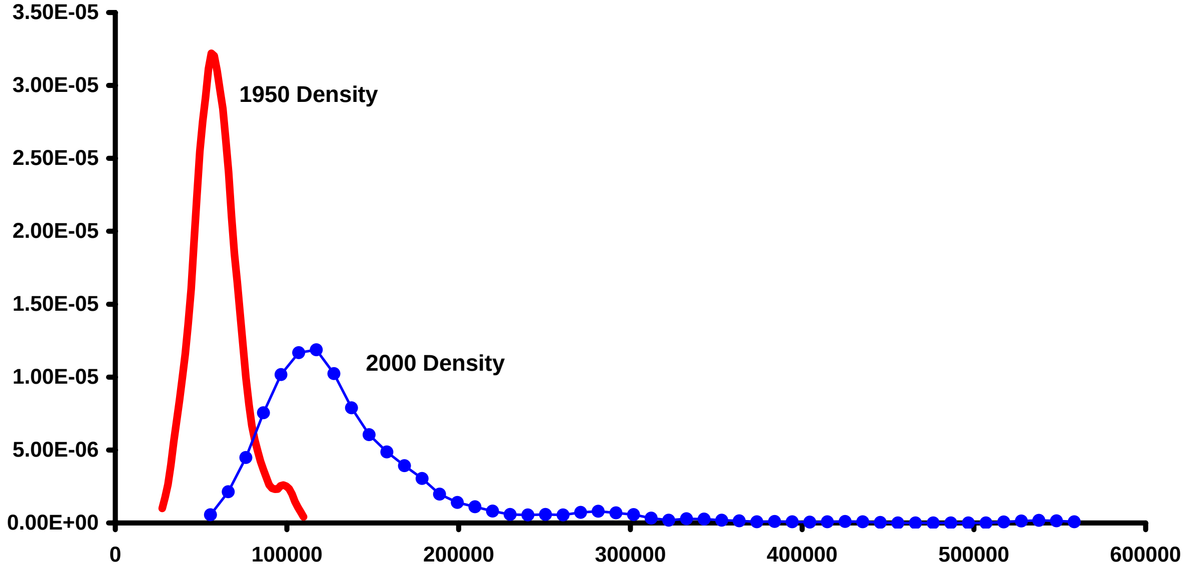


Figure 2B
Skewness in Mean House Values Across MSA's
1950 versus 2000

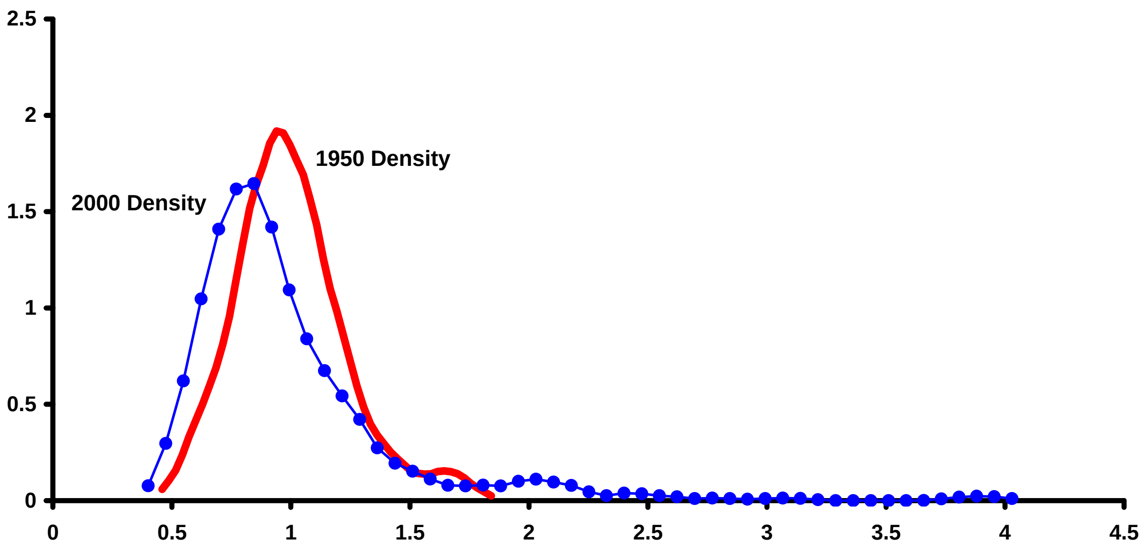


Figure 3 (from Saez (2004))

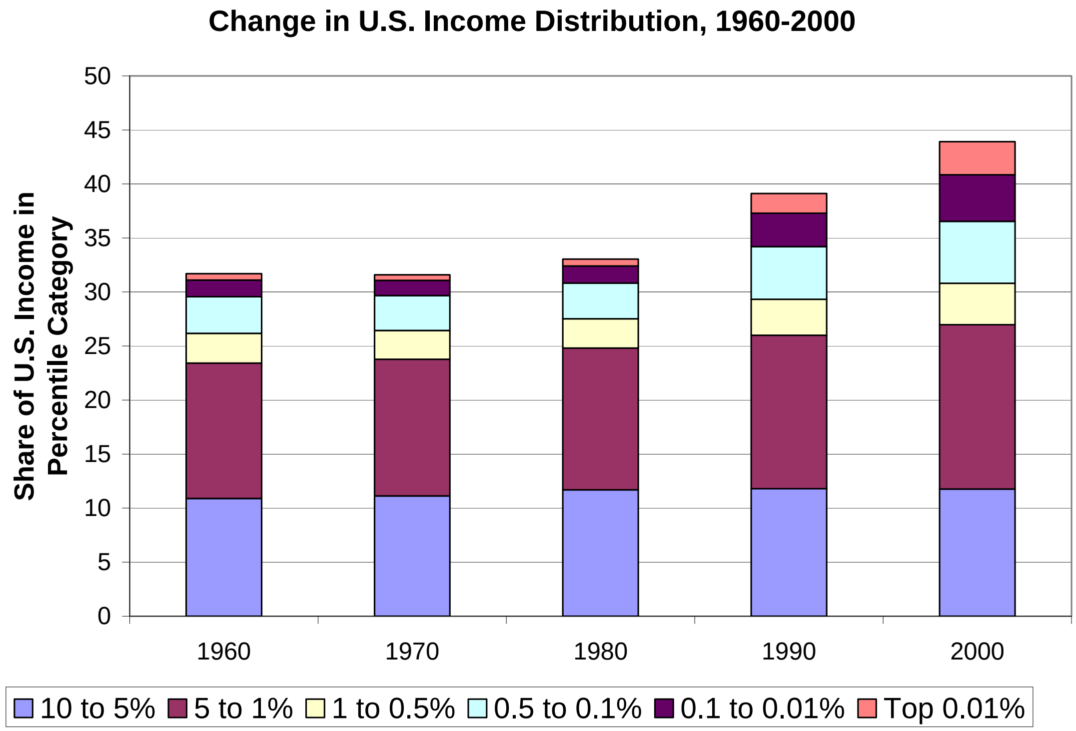


Figure 4: The Evolution of the National Income Distribution

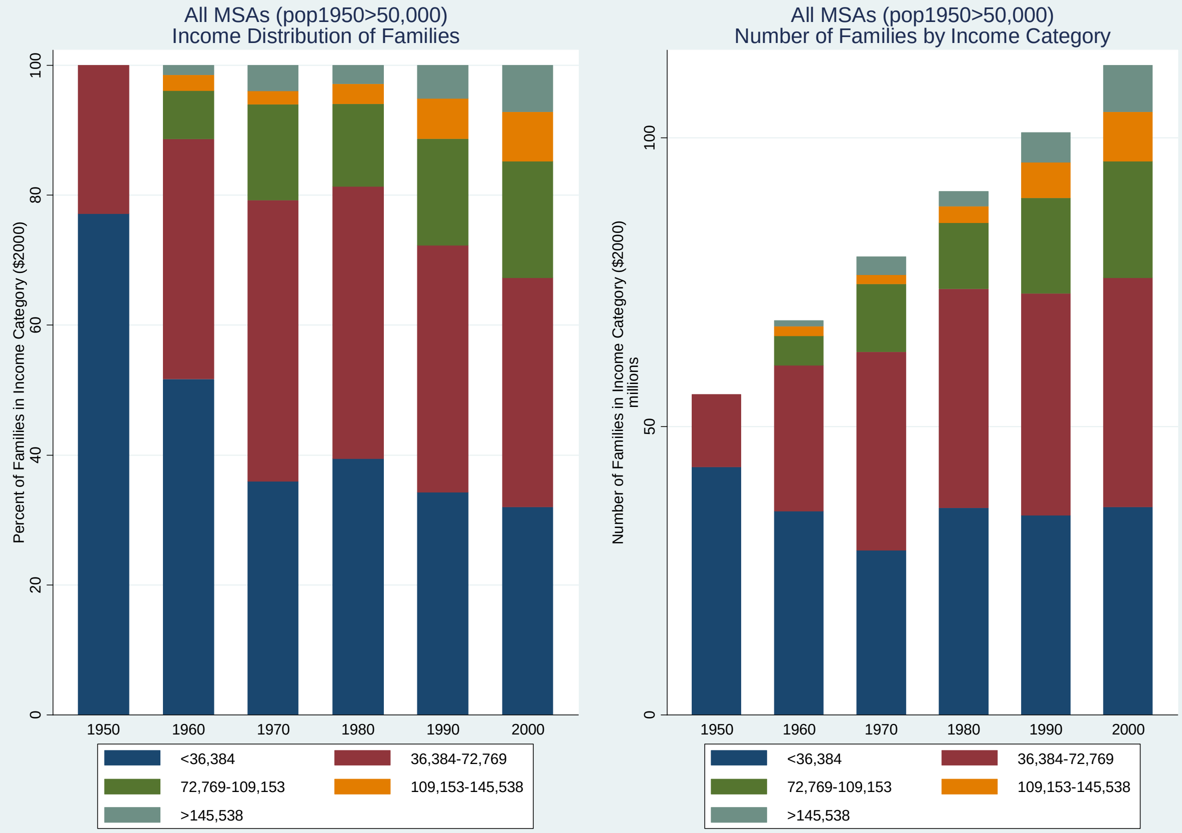


Figure 5: San Francisco (big price growth) gains rich, loses poor

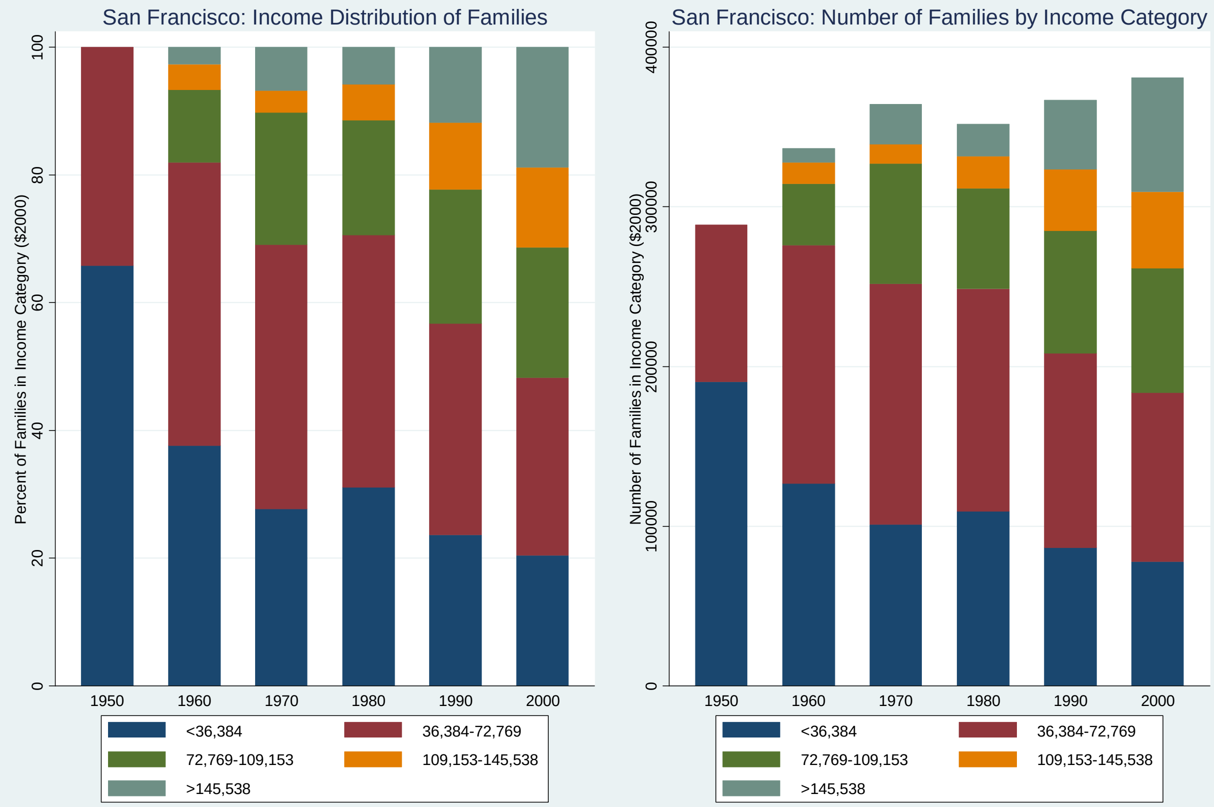


Figure 6: Las Vegas (big unit growth) gains rich and poor, shares stay constant

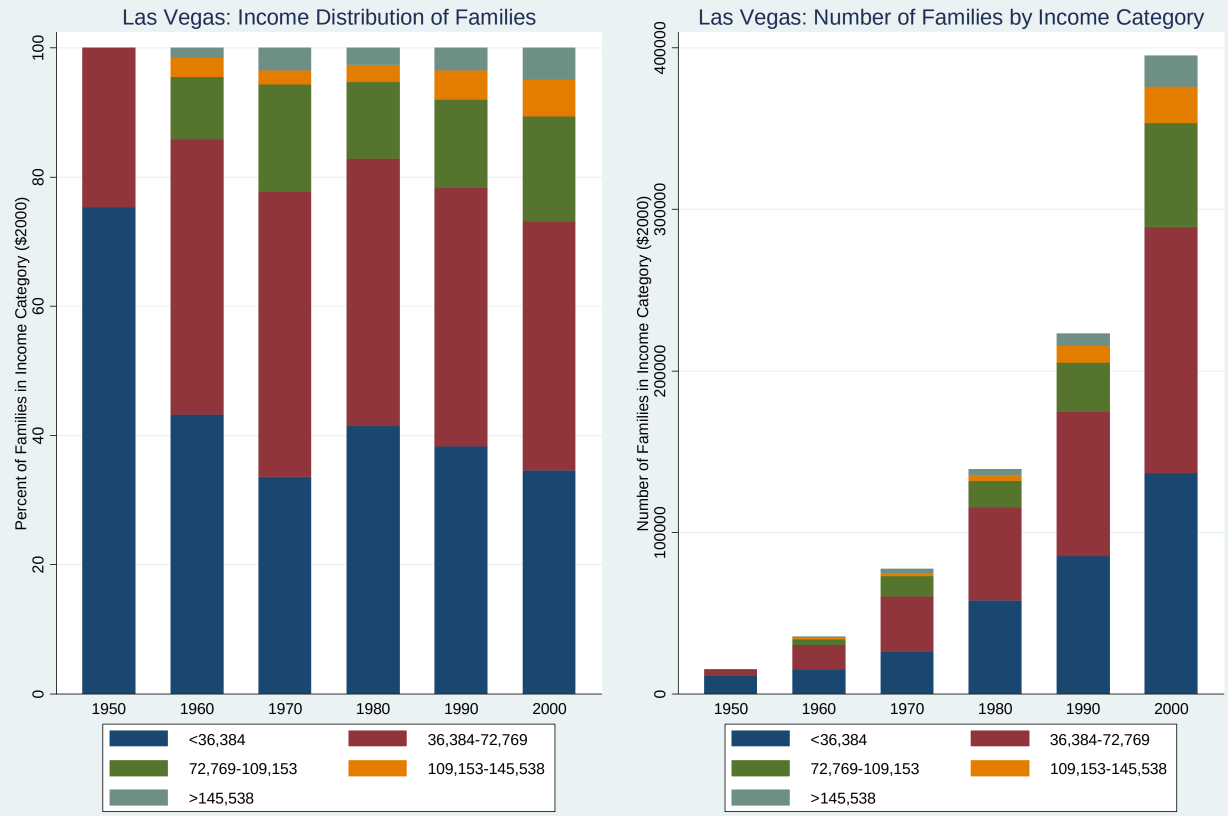
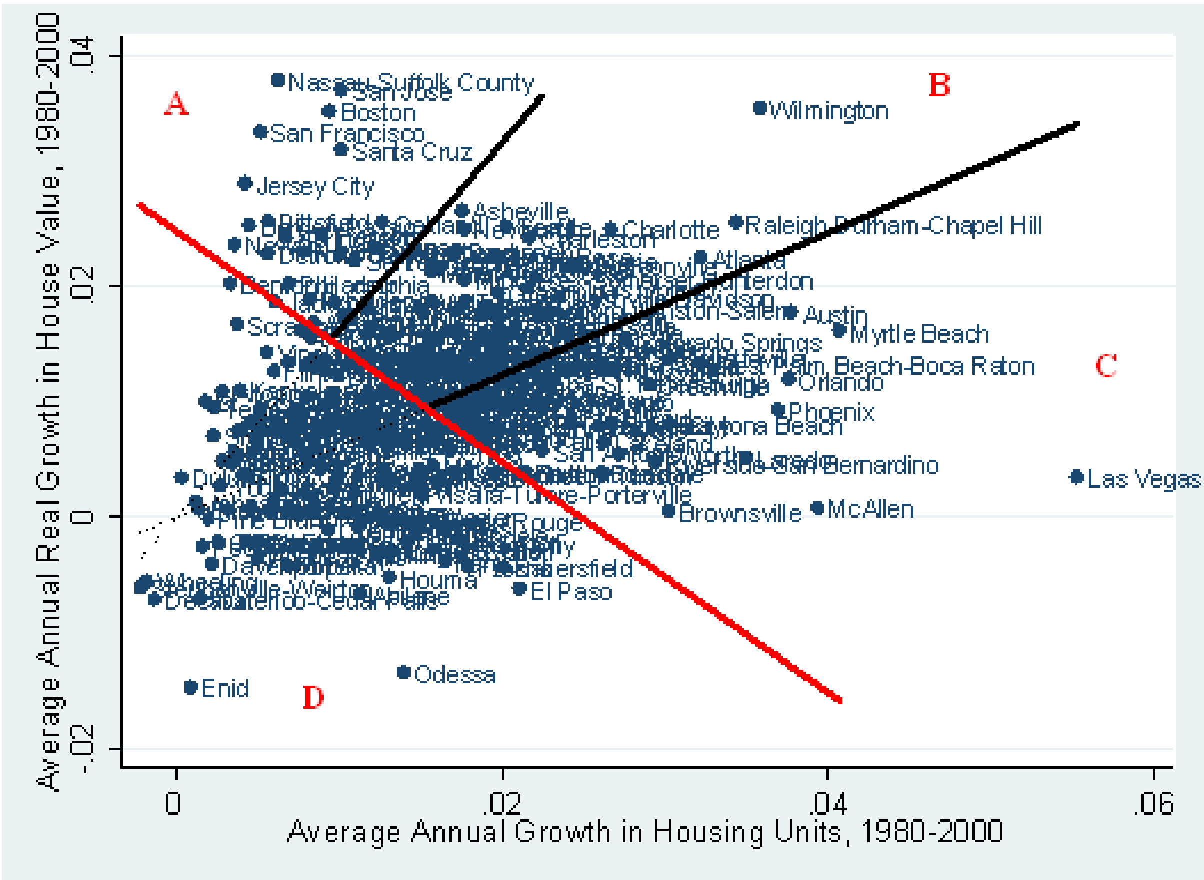


Figure 7: Real annual house price growth versus unit growth, 1980-2000



Appendix 1: House Prices and Appreciation

Decadal Census, All Values in $2000

MSA	% House Appreciation 1950-2000	1950 Mean Value ($2000)	2000 Mean Value ($2000)
Abilene, TX	34.6%	$54,917	$73,918
Akron, OH	93.9%	$69,720	$135,174
Albany, GA	61.7%	$60,388	$97,630
Albany, NY	62.3%	$75,522	$122,604
Albuquerque, NM	132.6%	$64,411	$149,835
Alexandria, LA	98.9%	$46,114	$91,722
Allentown, PA	110.3%	$61,811	$129,981
Altoona, PA	99.2%	$43,163	$85,966
Amarillo, TX	54.4%	$61,713	$95,299
Ann Arbor, MI	177.2%	$70,125	$194,421
Anniston, AL	132.5%	$37,222	$86,527
Appleton, WI	93.3%	$63,152	$122,098
Asheville, NC	192.3%	$50,005	$146,159
Athens, GA	191.4%	$49,138	$143,184
Atlanta, GA	178.8%	$61,933	$172,667
Atlantic City, NC	128.3%	$68,581	$156,590
Auburn-Opelika, AL	151.5%	$50,779	$127,708
Augusta, GA	138.9%	$45,543	$108,814
Austin, TX	193.8%	$55,895	$164,223
Bakersfield, CA	94.7%	$57,461	$111,850
Baltimore, MD	148.6%	$65,817	$163,594
Bangor, ME	113.3%	$43,328	$92,403
Barnstable, MA	205.5%	$76,239	$232,912
Baton Rouge, LA	115.3%	$56,276	$121,178
Beaumont, TX	47.6%	$51,200	$75,580
Bellingham, WA	276.4%	$49,780	$187,380
Benton Harbor, MI	103.9%	$59,222	$120,727
Bergen-Passic, NJ	196.0%	$98,065	$290,265
Billings, MT	48.4%	$79,117	$117,401
Biloxi, MS	170.4%	$39,205	$106,029
Binghamton, NY	24.4%	$70,626	$87,873
Birmingham, AL	178.1%	$47,949	$133,362
Bismarck, ND	72.0%	$61,250	$105,354
Bloomington, IN	112.1%	$61,691	$130,870
Bloomington, IL	176.3%	$47,973	$132,556
Boise City, ID	142.6%	$58,231	$141,275
Boston, MA	212.4%	$76,168	$237,974
Boulder, CO	377.2%	$61,206	$292,063
Brazoria, TX	123.5%	$46,086	$103,025
Bremerton, WA	280.4%	$51,233	$194,886
Brownsville, TX	73.8%	$39,569	$68,775
Bryan, TX	137.9%	$48,788	$116,046
Buffalo, NY	31.1%	$79,254	$103,880
Burlington, VT	131.9%	$65,502	$151,915
Canton, OH	78.4%	$65,215	$116,324
Casper, WY	37.8%	$72,285	$99,579
Cedar Rapids, IA	76.4%	$69,121	$121,942
Champaign, IL	49.6%	$75,056	$112,277
Charleston, SC	236.8%	$47,790	$160,960
Charleston, WV	56.0%	$67,951	$105,994
Charlotte, NC	194.1%	$53,454	$157,233
Charlottesville, VA	158.7%	$66,377	$171,734
Chattanooga, TN	154.3%	$45,327	$115,264
Cheyenne, WY	75.5%	$68,901	$120,934

Appendix 1: House Prices and Appreciation (continued)

Decadal Census, All Values in $2000

MSA	% House Appreciation 1950-2000	1950 Mean Value ($2000)	2000 Mean Value ($2000)
Chicago, IL	113.7%	$97,920	$209,302
Chico, CA	173.8%	$53,621	$146,827
Cincinnati, OH	76.2%	$82,734	$145,774
Clarksville, TN	146.1%	$39,349	$96,846
Cleveland, OH	57.0%	$91,687	$143,988
Colorado Springs, CO	162.7%	$67,264	$176,709
Columbia, MO	106.2%	$64,039	$132,067
Columbia, SC	109.0%	$62,560	$130,741
Columbus, GA	97.8%	$52,647	$104,113
Columbus, GA	112.5%	$68,152	$144,797
Corpus Christi, TX	60.8%	$52,261	$84,055
Corvallis, OR	190.3%	$65,383	$189,834
Cumberland, MD	78.8%	$45,269	$80,950
Dallas, TX	138.4%	$60,875	$145,125
Danville, VA	79.1%	$49,789	$89,160
Davenport, IA	46.4%	$69,396	$101,616
Dayton, OH	63.9%	$72,429	$118,740
Daytona Beach, FL	100.2%	$56,285	$112,670
Decatur, AL	39.7%	$59,324	$82,878
Decatur, IL	162.9%	$39,426	$103,651
Denver, CO	184.3%	$75,357	$214,261
Des Moines, IA	104.8%	$59,610	$122,069
Detroit, MI	123.8%	$72,666	$162,595
Dothan, AL	132.9%	$41,834	$97,447
Dover, DE	142.0%	$52,372	$126,746
Dubuque, IA	55.7%	$71,399	$111,178
Duluth, MN	77.0%	$50,214	$88,899
Dutchess County, NY	103.9%	$84,876	$173,021
Eau Claire, WI	106.0%	$53,068	$109,346
El Paso, TX	21.9%	$68,651	$83,652
Elkhart, IN	124.8%	$51,894	$116,662
Elmira, NY	19.8%	$65,681	$78,693
Enid, OK	35.4%	$52,425	$70,985
Erie, PA	60.8%	$63,623	$102,287
Eugene, OR	169.8%	$60,521	$163,308
Evansville, IN	104.6%	$51,168	$104,673
Fargo, SD	65.2%	$64,995	$107,401
Fayetteville, NC	163.0%	$44,821	$117,882
Fayetteville, AR	131.1%	$46,057	$106,439
Flagstaff, AZ	226.7%	$50,500	$164,989
Flint, MI	108.4%	$52,717	$109,844
Florence, AL	105.7%	$53,411	$109,874
Florence, SC	143.6%	$41,008	$99,881
Fort Collins, CO	246.9%	$58,103	$201,557
Fort Lauderdale, FL	112.5%	$76,577	$162,733
Fort Myers, FL	224.3%	$47,951	$155,498

Appendix 1: House Prices and Appreciation (continued)

Decadal Census, All Values in $2000

MSA	% House Appreciation 1950-2000	1950 Mean Value ($2000)	2000 Mean Value ($2000)
Fort Pierce, FL	164.5%	$55,601	$147,065
Fort Smith, AR	123.3%	$38,849	$86,732
Fort Walton Beach, FL	310.2%	$32,220	$132,178
Fort Wayne, IN	81.9%	$58,417	$106,245
Fort Worth, TX	125.2%	$51,794	$116,627
Fresno, CA	110.9%	$61,792	$130,339
Gadsden, AL	95.6%	$43,564	$85,218
Gainesville, FL	131.6%	$52,261	$121,013
Galveston, TX	73.9%	$62,502	$108,689
Gary, IN	79.6%	$68,478	$123,004
Glens Falls, NY	111.5%	$52,596	$111,252
Goldsboro, NC	117.0%	$48,770	$105,809
Grand Forks, ND	84.0%	$52,702	$96,954
Grand Junction, CO	182.4%	$50,121	$141,565
Grand Rapids, MI	122.4%	$61,120	$135,937
Great Falls, MT	60.5%	$66,267	$106,331
Greeley, CO	240.5%	$47,601	$162,079
Green Bay, WI	92.0%	$69,589	$133,603
Greensboro-Winston-Salem, NC	160.4%	$51,382	$133,785
Greenville, NC	126.8%	$53,496	$121,353
Greenville, SC	136.4%	$51,358	$121,431
Hagerstown, MD	128.9%	$56,392	$129,058
Hamilton, OH	101.9%	$67,859	$136,985
Harrisburg, PA	104.5%	$60,176	$123,036
Hartford, CT	85.9%	$94,780	$176,237
Hattiesburg, MS	157.9%	$37,870	$97,658
Hickory, NC	169.4%	$43,043	$115,939
Houma, LA	161.1%	$36,392	$95,011
Houston, TX	100.2%	$63,203	$126,516
Huntington, WV	60.5%	$52,196	$83,751
Huntsville, AL	204.2%	$41,005	$124,754
Indianapolis, IN	123.2%	$60,474	$134,977
Iowa City, IA	101.6%	$77,367	$155,995
Jackson, MI	112.1%	$47,567	$100,887
Jackson, MS	123.8%	$51,349	$114,931
Jackson, TN	77.8%	$61,374	$109,126
Jacksonville, FL	134.7%	$56,494	$132,578
Jacksonville, NC	226.7%	$31,850	$104,044
Jamestown, NY	31.3%	$58,609	$76,940
Janesville, WI	76.8%	$62,627	$110,704
Jersey City, NJ	136.8%	$72,622	$171,946
Johnson City, TN	121.3%	$46,771	$103,517
Johnstown, PA	65.9%	$45,873	$76,127
Jonesboro, AR	128.9%	$43,218	$98,938
Joplin, MO	143.5%	$34,162	$83,176
Kalamazoo, MI	97.9%	$58,856	$116,504

Appendix 1: House Prices and Appreciation (continued)

Decadal Census, All Values in $2000

MSA	% House Appreciation 1950-2000	1950 Mean Value ($2000)	2000 Mean Value ($2000)
Kankakee, IL	70.3%	$68,181	$116,145
Kansas City, MO	118.4%	$58,259	$127,225
Kenosha, WI	93.3%	$71,148	$137,515
Killeen, TX	100.3%	$44,527	$89,207
Knoxville, TN	179.7%	$44,710	$125,053
Kokomo, IN	129.7%	$45,759	$105,114
La Crosse, WI	99.0%	$56,323	$112,078
Lafayette, LA	155.4%	$39,681	$101,363
Lafayette, IN	108.3%	$59,286	$123,521
Lake Charles, LA	95.2%	$50,583	$98,730
Lakeland, FL	101.7%	$49,523	$99,883
Lancaster, PA	100.4%	$67,637	$135,567
Lansing, MI	118.0%	$56,559	$123,283
Laredo, TX	181.2%	$30,869	$86,801
Las Cruces, NM	157.1%	$43,025	$110,607
Las Vegas, NV	147.5%	$65,114	$161,166
Lawrence, KS	187.3%	$49,050	$140,902
Lawton, OK	72.7%	$48,036	$82,946
Lewiston, ME	92.2%	$52,248	$100,434
Lexington-Fayette, KY	113.7%	$60,367	$129,025
Lima, OH	72.7%	$56,382	$97,381
Lincoln, NE	105.5%	$61,336	$126,018
Little Rock, AR	117.1%	$50,879	$110,443
Longview, TX	123.2%	$37,678	$84,102
Los Angeles, CA	236.6%	$85,150	$286,633
Louisville, KY	113.4%	$60,413	$128,893
Lubbock, TX	36.1%	$62,442	$84,999
Lynchburg, VA	124.4%	$52,348	$117,452
Macon, GA	133.0%	$44,416	$103,502
Madison, WI	99.0%	$86,136	$171,383
Mansfield, OH	49.3%	$64,370	$96,099
McAllen, TX	99.9%	$33,393	$66,759
Medford, OR	199.0%	$56,647	$169,383
Melbourne, FL	114.9%	$55,488	$119,262
Memphis, TN	100.7%	$61,886	$124,183
Merced, CA	121.8%	$58,295	$129,318
Miami, FL	95.2%	$83,286	$162,594
Middlesex-Somerset-Hunterdon, NJ	185.6%	$80,437	$229,739
Milwaukee, WI	69.3%	$92,698	$156,918
Minneapolis-St. Paul, MN	117.6%	$77,421	$168,496
Missoula, MT	162.5%	$59,653	$156,573
Mobile, AL	184.0%	$41,465	$117,766
Modesto, CA	162.2%	$55,669	$145,969
Monmouth-Ocean, NJ	160.2%	$77,938	$202,758
Monroe, LA	101.7%	$49,470	$99,781
Montgomery, AL	107.2%	$55,648	$115,307

Appendix 1: House Prices and Appreciation (continued)

Decadal Census, All Values in $2000

MSA	% House Appreciation 1950-2000	1950 Mean Value ($2000)	2000 Mean Value ($2000)
Muncie, IN	66.8%	$51,851	$86,505
Myrtle Beach, SC	176.3%	$52,277	$144,456
Naples, FL	406.7%	$51,144	$259,155
Nashville-Davidson, TN	178.8%	$56,363	$157,166
Nassau-Suffolk County, NY	167.6%	$99,692	$266,806
New Haven, CT	185.4%	$103,118	$294,297
New London, CT	132.5%	$74,479	$173,185
New Orleans, LA	81.2%	$71,836	$130,140
New York, NY	181.4%	$103,209	$290,412
Newark, NJ	155.2%	$101,549	$259,115
Newburgh, NY	125.2%	$70,748	$159,289
Norfolk, VA	150.2%	$54,670	$136,783
Oakland, CA	300.8%	$86,596	$347,050
Ocala, FL	146.8%	$40,186	$99,169
Odessa, TX	27.4%	$59,116	$75,294
Oklahoma City, OK	65.8%	$58,078	$96,278
Olympia, WA	194.8%	$57,586	$169,788
Omaha, NE	104.3%	$59,470	$121,483
Orange County, CA	356.1%	$72,185	$329,206
Orlando, FL	122.8%	$61,908	$137,919
Owensboro, KY	72.7%	$55,968	$96,648
Panama City, FL	238.7%	$34,908	$118,233
Parkersburg, WV	64.7%	$56,158	$92,516
Pensacola, FL	185.6%	$40,422	$115,431
Peoria-Pekin, IL	59.8%	$66,167	$105,723
Philadelphia, PA	121.6%	$66,426	$147,186
Phoenix, AZ	209.2%	$53,106	$164,191
Pine Bluff, AR	113.6%	$33,106	$70,724
Pittsburgh, PA	66.1%	$64,015	$106,345
Pittsfield, MA	96.9%	$73,066	$143,854
Pocatello, ID	71.7%	$60,819	$104,417
Portland, OR	221.4%	$63,337	$203,578
Portland, ME	169.3%	$60,377	$162,576
Providence, RI	94.1%	$81,189	$157,574
Provo-Orem, UT	207.3%	$61,174	$187,982
Pueblo, CO	120.8%	$50,635	$111,798
Punta Gorda, FL	215.2%	$39,342	$124,010
Racine, WI	72.1%	$74,706	$128,537
Raleigh-Durham-Chapel Hill, NC	205.7%	$58,153	$177,794
Rapid City, SD	89.5%	$59,458	$112,668
Reading, PA	96.3%	$59,750	$117,313
Redding, CA	168.0%	$52,416	$140,465
Reno, NV	115.4%	$96,874	$208,650
Richland, WA	117.2%	$60,700	$131,811
Richmond-Petersburgh, VA	116.5%	$64,964	$140,677
Riverside-San Bernardino, CA	173.7%	$59,725	$163,483

Appendix 1: House Prices and Appreciation (continued)

Decadal Census, All Values in $2000

MSA	% House Appreciation 1950-2000	1950 Mean Value ($2000)	2000 Mean Value ($2000)
Roanoke, VA	103.8%	$60,679	$123,680
Rochester, MN	68.1%	$81,995	$137,822
Rochester, NY	56.1%	$72,348	$112,926
Rockford, IL	51.2%	$73,216	$110,727
Rocky Mount, NC	109.4%	$50,538	$105,837
Sacramento, CA	167.9%	$71,504	$191,567
Saginaw, MI	90.4%	$54,865	$104,471
Salem, OR	159.8%	$59,484	$154,551
Salinas, CA	316.6%	$83,456	$347,705
Salt Lake City, UT	157.1%	$70,810	$182,029
San Angelo, TX	52.8%	$50,539	$77,215
San Antonio, TX	75.2%	$56,397	$98,829
San Diego, CA	262.4%	$78,640	$284,952
San Francisco, CA	465.9%	$96,703	$547,206
San Jose, CA	513.3%	$86,667	$531,562
San Luis Obispo, CA	346.0%	$59,995	$267,605
Santa Barbara-Santa Maria, CA	328.4%	$89,559	$383,707
Santa Cruz, CA	522.0%	$68,494	$426,041
Santa Fe, NM	284.9%	$66,127	$254,503
Santa Rosa, CA	362.5%	$69,007	$319,124
Sarasota, FL	166.7%	$62,131	$165,729
Savannah, GA	153.5%	$53,867	$136,552
Scranton, PA	111.5%	$49,142	$103,948
Seattle, WA	285.7%	$70,684	$272,603
Sharon, PA	58.4%	$56,123	$88,901
Sheboygan, WI	84.6%	$67,042	$123,742
Shermon-Denison, TX	119.4%	$38,321	$84,065
Shreveport-Bossier, LA	61.6%	$57,812	$93,411
Sioux City, IA	55.1%	$57,815	$89,660
Sioux Falls, SD	87.5%	$64,197	$120,400
South Bend, IN	66.4%	$62,322	$103,678
Spokane, WA	119.0%	$60,147	$131,739
Springfield, IL	83.1%	$60,736	$111,198
Springfield, MA	93.7%	$72,294	$140,063
Springfield, MO	128.5%	$47,932	$109,543
St. Cloud, MN	135.0%	$48,134	$113,132
St. Joseph, MO	126.5%	$39,063	$88,484
St. Louis, MO	78.6%	$72,973	$130,348
State College, PA	145.6%	$54,367	$133,541
Steubenville-Weirton, OH	34.4%	$57,706	$77,550
Stockton, CA	171.8%	$60,531	$164,517
Sumter, SC	93.4%	$47,929	$92,696
Syracuse, NY	39.8%	$69,624	$97,341
Tacoma, WA	201.6%	$58,269	$175,746
Tallahassee, FL	137.0%	$53,971	$127,889
Tampa-St. Petersburgh, FL	109.4%	$58,714	$122,967

Appendix 1: House Prices and Appreciation (continued)

Decadal Census, All Values in $2000

MSA	% House Appreciation 1950-2000	1950 Mean Value ($2000)	2000 Mean Value ($2000)
Terre Haute, IN	134.0%	$36,094	$84,467
Texarkana, TX	123.4%	$35,200	$78,620
Toledo, OH	80.4%	$65,783	$118,705
Topeka, KS	72.1%	$54,593	$93,969
Trenton, NJ	189.2%	$67,916	$196,431
Tucson, AZ	130.5%	$63,094	$145,417
Tulsa, OK	99.0%	$53,533	$106,510
Tuscaloosa, AL	178.6%	$46,197	$128,691
Tyler, TX	97.4%	$52,262	$103,168
Utica-Rome, NY	30.6%	$64,791	$84,587
Vallejo, CA	233.4%	$69,620	$232,145
Ventura, CA	319.6%	$70,971	$297,826
Victoria, TX	57.2%	$55,147	$86,680
Vineland-Millville-Bridgeton, NJ	91.2%	$53,459	$102,201
Visalia-Tulare-Porterville, CA	159.7%	$46,174	$119,908
Waco, TX	70.1%	$48,552	$82,577
Washington, DC	112.7%	$106,235	$225,914
Waterloo-Cedar Falls, IA	40.2%	$64,682	$90,685
Wausau, WI	114.3%	$51,753	$110,908
West Palm Beach-Boca Raton, FL	159.7%	$73,275	$190,261
Wheeling, WV	46.3%	$53,928	$78,871
Wichita, KS	62.9%	$60,499	$98,554
Wichita Falls, TX	57.4%	$47,826	$75,266
Williamsport, PA	82.3%	$53,625	$97,759
Wilmington, DE	90.8%	$82,087	$156,661
Wilmington, NC	310.6%	$42,865	$176,011
Yakima, WA	140.7%	$54,809	$131,944
Yolo, CA	205.7%	$65,842	$201,293
York, PA	104.8%	$60,915	$124,730
Youngstown-Warren, OH	49.8%	$63,044	$94,470
Yuba, CA	146.4%	$51,463	$126,793
Yuma, AZ	156.6%	$44,473	$114,101

Appendix 2: MSA Summary Statistics

Variable	Mean	Standard deviation
MSA time-invariant characteristics:		
Average Annual Real House Price Growth, 1950-2000 (N=279)	1.57	0.56
Average Annual Housing Unit Growth, 1950-2000 (N=279)	2.10	0.98
Average Annual Real Income Growth, 1950-2000 (N=279)	1.82	0.35
Ever a “superstar”	0.165 [46]	0.372
Ever a “non-superstar”	0.337 [94]	0.474
Ever “low demand”	0.821 [229]	0.384
MSA time-varying characteristics:		
Average 20-year Real House Price Growth	1.50	1.04
Average 20-year Housing Unit Growth	2.10	1.20
Average 20-year house price growth + housing unit growth	3.60	1.86
Average ratio of 20-year price growth to 20-year unit growth	0.936	0.642
Real house value	111,329	54,889
Average price/average annual rent	17.00	3.99
Year	# “superstars”	# “non-superstars”
1970	3	55
1980	3	34
1990	30	43
2000	21	36
Income Distribution	Mean	Standard deviation
Share of an MSA’s population that is “rich”	0.033	0.021
Share “middle-rich”	0.035	0.024
Share “middle”	0.129	0.043
Share “middle-poor”	0.400	0.050
Share “poor”	0.402	0.095
National number “rich”		
1970	1,571,136	
1980	1,312,103	
1990	2,611,178	
2000	4,098,324	