Heterogeneity in Neighborhood-Level Price Growth in the U.S., 1993-2009*

Fernando Ferreira and Joseph Gyourko
The Wharton School,
University of Pennsylvania & NBER

December 16, 2011

The development of new and large micro data bases on housing transactions combined with the increased power of geographical mapping systems has opened up new vistas for research in housing and urban economics. We use such tools in this article to document the nature of the heterogeneity found at the very local level for a representative sample of markets in the United States from 1993 to 2009.

Studying local heterogeneity in housing markets is important for a variety of reasons. One is that differences across neighborhoods within a metropolitan area (MSA) affect the measurement and interpretation of MSA price changes, just as MSA-level heterogeneity influences our interpretation of aggregate data at the national level. For example, the extraordinary boom and bust in housing prices recently experienced in the United States was not a single national event, as shown by Sinai's (2011) documentation of wide variation in price movements across metropolitan areas (MSAs), and Ferreira's and Gyourko's (2011) estimates that the timing of the beginning of the housing boom varied by as much as a decade across different U.S. markets. Therefore, understanding the last American housing cycle will require accounting for that local heterogeneity, which immediately suggests that those proposing single cause explanations for the recent boom and bust in housing are on weak ground empirically.

Local heterogeneity may also provide a natural laboratory for studying how housing booms start, how they develop and ultimately bust. In other work, we have shown that local

* We thank the Research Sponsors Program of the Zell/Lurie Real Estate Center at Wharton for financial support.

income growth was an important factor at the start of the boom in many markets and neighborhoods. That begs the question of whether other forces, possibly including some type of contagion impact from neighbors (either geographically close or economically similar) might also have played a role. Cotter, Gabriel and Roll (2011) and Zhu, Fuss and Rottke (2011) have begun to study this question using metropolitan area and regional data. The natural extension will be to use even more local data.

More generally, we know relatively little about the workings of within-metropolitan area housing markets. Basic facts about the level of interdependencies across neighborhoods, whether they are substitutes or complements, and how booms and bust propagate at a very local level have been hidden from us because of a lack of high frequency, detailed information on transactions.

In the remainder of this article we examine the heterogeneity in price growth during the most recent American housing cycle at the neighborhood level. Our estimates rely on a new micro data set which contains over 23 million housing transactions in 94 metropolitan areas from 1993 to 2009, provided by DataQuick. It contains many housing characteristics, such as the date and value of the housing transaction, in addition to the precise location of each house. We first assign each house to a Census tract, then contiguous census tracts were combined into pairs (and sometimes triplets when necessary) using a random process to form tract groups – our definition of neighborhoods - in order to provide sufficient observations to estimate price indexes at the local level.[1]

---

[1] We create neighborhood-level (*m*) constant quality house price series by half-year (*t*) using hedonic regressions. Price (*HP*), in logarithmic form, is a function of the square footage (*Sqft*) of the home entered in quadratic form, the number of bedrooms (*Bed*), the number of bathrooms (*Bath*), and the age of the home (*Age*). The hedonic index values are derived from the coefficients in the vector α_6 on the year-quarter dummies (*YearQtr*) in the following equation: $Log(HP_{m,t}) = \alpha_0 + \alpha_1 * Bed_{m,t} + \alpha_2 * Bath_{m,t} + \alpha_3 * Age_{m,t} + \alpha_4 * Sqft_{m,t} + \alpha_5 * Sqft^2_{m,t} +$
$+ \alpha_6 YearQtr_t + \epsilon_{m,t}$, where $\epsilon_{m,t}$ is an idiosyncratic error term. The estimated indexes are then normalized to 100 in 2000(Q1) for all MSAs.

For each neighborhood in our sample we estimate the beginning of the housing boom by when there was a global structural break in the local area's price appreciation rate (see Ferreira and Gyourko (2011) for details). Subsequently we estimate the length and magnitude of the housing boom across neighborhoods. This done using a randomly drawn split sample approach, as in Card, Mas and Rothstein (2008): one sample of houses is used to estimate the time of the breakpoint, and the other half is used to estimate the magnitude of the price changes over time. This approach mitigates potential specification search bias as described in Leamer (1983).

Below we show how neighborhood price levels evolved from the beginning of the housing boom until the beginning of the bust. We provide stylized facts related to the length of the housing boom by neighborhood, its total magnitude, the heterogeneity in concentration of neighborhood booms by MSA, and finally, whether socio-demographic characteristics are correlated or not with the timing of the housing boom across neighborhoods. We conclude with directions for future research.

I. The Timing and Magnitudes of Booms and Busts Across Neighborhoods

In Ferreira and Gyourko (2011), we found that different neighborhoods began their booms as early as 1995 and as late as 2006. There is a modest concentration in the 2003-2005 period during which the aggregate price indexes reported by FHFA and S&P/Case-Shiller peak, but well over half the neighborhoods in our sample boomed before then, with a few booming very late. The magnitude of the jump in price appreciation rate when the boom began was about 6 percentage points on average.

While there were substantial differences in when the boom began at the neighborhood level, the beginning of the bust was much more concentrated temporally. We define the beginning of the bust by the timing of the peak in local area price levels (not price growth rates). Figure 1 plots the distribution of when house prices peaked by half-year period. Almost two-thirds of the neighborhoods in our sample saw their price levels peak within a 1.5 year period running from the second half of 2005 through the end of 2006. About one-quarter peaked in 2007 or 2008.

We can combine these data with the results on the beginning of the boom to measure how long the booms lasted across neighborhoods. Figure 2 plots the results. The clear negative slope illustrates that booms lasted a long time for early-booming neighborhoods and were much shorter for the later boomers. Neighborhoods that initially boomed in the mid-1990s almost always saw prices grow for another 8-10 years before peaking. In contrast, the late boomers from 2004-on experienced no more than another 2-3 years of price growth before the bust.

However, it is not the case that overall price growth from the beginning of the boom to the peak of prices was substantially higher in early booming neighborhoods, as Figure 3's plot of the aggregate price growth experienced from start of the boom documents. There is a modest inverted-U shape to the plot, so that we see somewhat higher overall price growth during the boom among places that started booming around the year 2000. However, it still is accurate to say that the typical neighborhood experienced nominal price growth of between 100%-150% over the course of the boom, no matter when we estimate its boom started. Essentially, late boomers experienced very large initial jumps in price growth at the start of the boom, as illustrated in Figure 4. And, those very high rates of appreciation tended to be maintained for a short time until the bust began.

II. Variation Across Neighborhoods within Metropolitan Areas

One of the striking features of the last housing cycle is that the starting points of neighborhood booms are much more concentrated among late booming markets. Figure 5 illustrates this in its summary of the relationship between the timing of MSA-level structural break points and the timing of neighborhood-level breaks. The y-axis measures the percentage of neighborhoods that have a break point within 12 months of the beginning of the housing boom in its respective MSA. Among the late-booming metropolitan areas, which include many markets such as Phoenix and Las Vegas in the so-called sand states, at least two-thirds of their neighborhoods boom within a year of the beginning of the metro boom. In early booming markets, which include many coastal metros such as Boston, Los Angeles, and San Francisco, only 40% of their neighborhoods boom within a year of the structural break in the metropolitan area's price growth rate. It is unknown whether this variation is due to differences in the nature of the local neighborhoods (e.g., perhaps neighborhoods in Phoenix are much more alike than those in San Francisco) or in terms of the shocks received by the metropolitan areas themselves. Clearly, this is an area ripe for research.

Figure 6 then illustrates how the timing of the beginning of the neighborhood-level booms correlates with different economic and demographic variables. These plots, which are based on all neighborhoods with statistically significant break points, essentially describe MSA-level variation. The plot in the upper left-hand corner shows that higher income neighborhoods tended to boom earlier. Given this, it is not surprising that the adjacent graph depicts a similar relationship for college graduates. However, the bottom two plots show much weaker relationships of race or population with the timing of the initial boom.

Figure 7 then documents that these national patterns mask a lot of heterogeneity within given metropolitan areas. For example, the first set of plots for the Boston metropolitan area shows that it was the relatively low income neighborhoods of that market which boomed first. Markets such as San Francisco and Fresno present opposite patterns. And, the concentration of neighborhood booms within a very short time span in Las Vegas means there is little temporal relationship with sociodemographics in late-booming markets.

III. Conclusions and Implications

There is much heterogeneity in how the American housing boom and bust played out at the neighborhood level. This is evident in when the booms began, how big the initial jumps in price growth were, how long the booms lasted, how concentrated were booms across neighborhoods within the same metropolitan area, and what types of neighborhoods boomed first or last. The two metrics on which there is more similarity across neighborhoods is in the timing of the bust and in the aggregate price appreciation experienced before the bust began. Most neighborhoods saw prices peak within a two-year window and the vast majority experienced nominal price growth of between 100% and 150% over the course of the boom.

If we are to ever truly understand the last great housing cycle, we will have to understand the great heterogeneity and limited homogeneity regarding what went on at the very local level. Clearly, these data should caution anyone who professes a single explanation for the boom. There simply is too much local variation for that. We also need to be cognizant that heterogeneity at the neighborhood-level affects measurement of the boom at the metropolitan area level, just as cross-MSA variation impacts our interpretation of aggregate, national data. We also need to understand what drove the differences in neighborhood patters across metropolitan areas. Just what makes Las Vegas neighborhoods so different from those in San Francisco? The same holds for why they correlations with sociodemographics are so different across markets. Finally, the neighborhood-level data provide a natural laboratory for examination of how market fundamentals and factors such as psychological contagion may have influenced how the boom developed over time.

References

Cotter, John, Stuart Gabriel and Richard Roll. "Integration and Contagion in U.S. Housing Markets", working paper, Ziman Center for Real Estate, UCLA, October 12, 2011.

Card, David, Alexandre Mas and Jesse Rothstein. "Tipping and the Dynamics of Segregation", *Quarterly Journal of Economics*, Vol. 123, no. 1 (2008): 177-218.

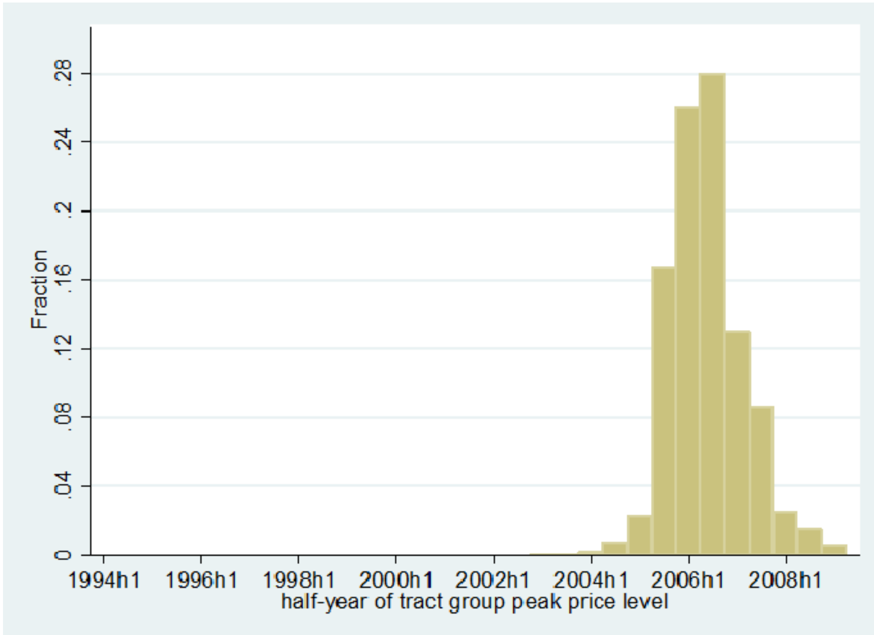
Ferreira, Fernando, and Joseph Gyourko. "Anatomy of the Beginning of the Housing Boom: U.S. Neighborhoods and Metropolitan Areas, 1993-2009", NBER working paper No. w17374.

Leamer, Edward. "Let's Take the Con Out of Econometrics", *American Economic Review*, Vol. 73, no. 1 (1983): 31-43.

Sinai, Todd. "House Price Moments in Boom-Bust Cycles", mimeo, The Wharton School, University of Pennsylvania, 2011.

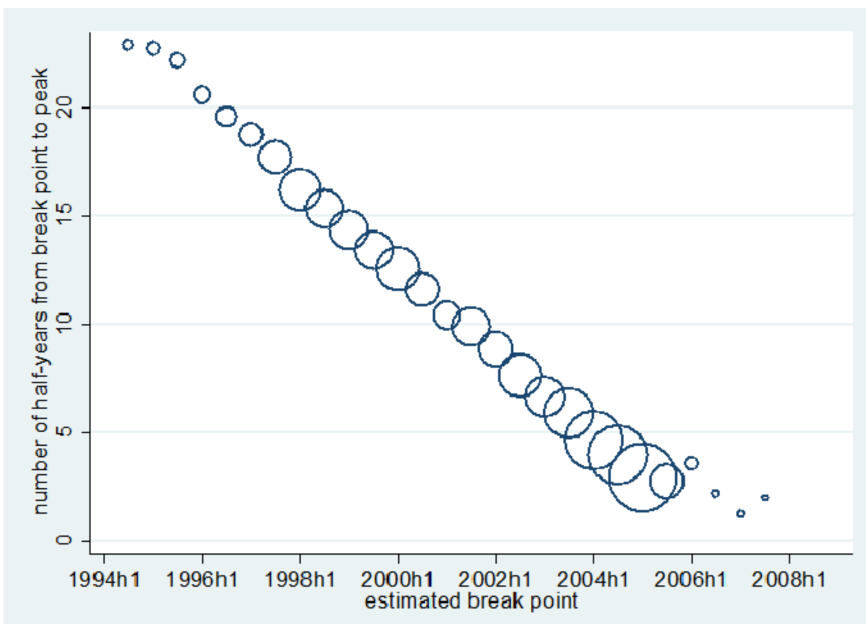
Zhu, Bing, Roland Fuss, and Nico Rottke. "Correlated Risk and Volatility Spillovers across U.S. Regional Housing Markets", working paper, January 2011.

Figure 1: Distribution of Price Level Peaks, Tract Groups with Statistically Significant Breakpoints, by Half Year



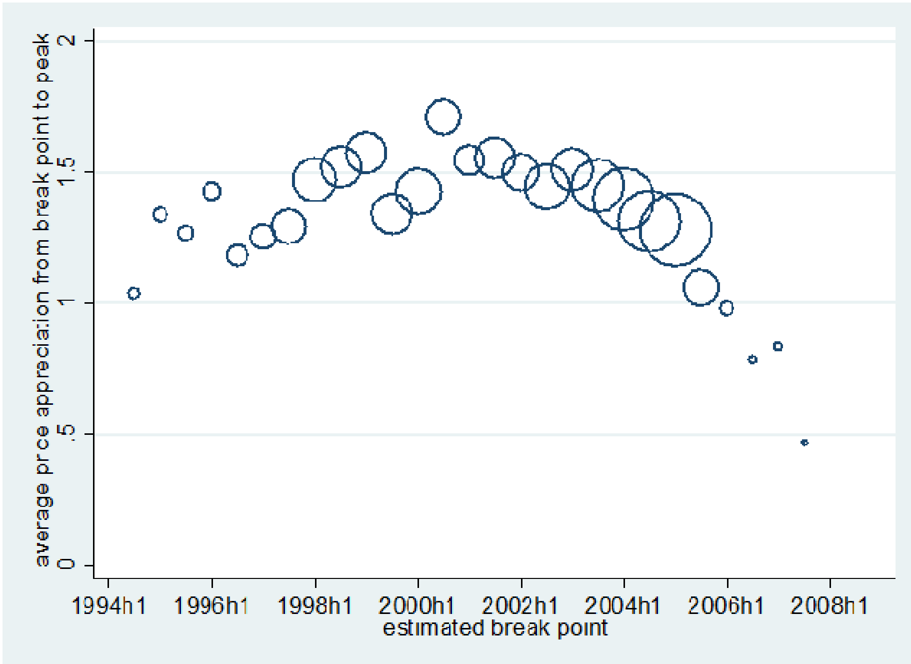
Notes: The histogram plots the fraction of tract groups with price peak levels in each half-year. It uses the sample of tract groups with statistically significant break points, and only considers half-years after the beginning of the boom.

Figure 2: Average Number of Half-Years from Breakpoint to Price Level Peak, Tract Groups with Statistically Significant Breakpoints, Weighted by Population



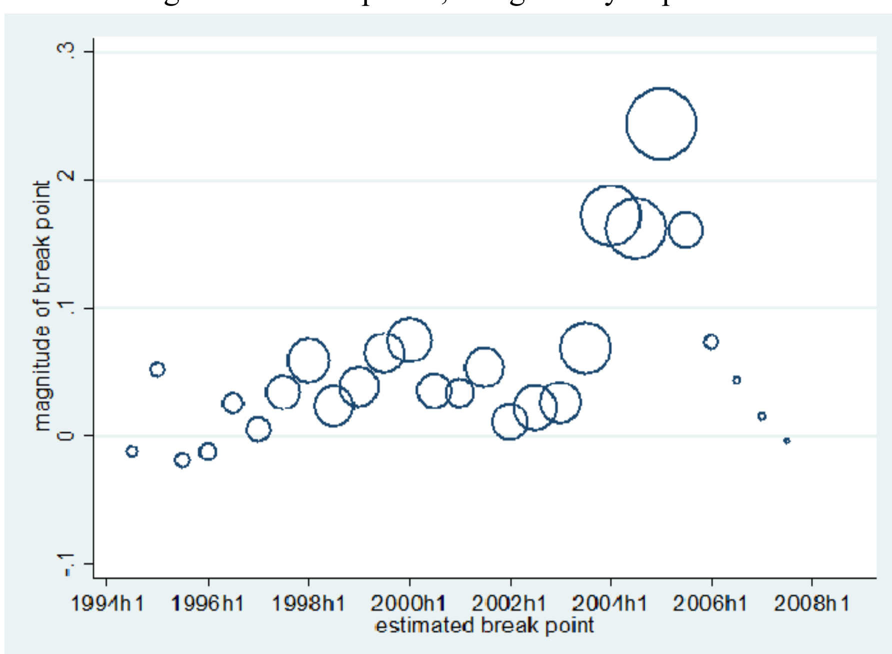
Notes: Each dot represents the average number of quarters from breakpoint to price level peak. Only tract groups with statistically significant breakpoints are included. Results are averaged by half-year, and weighted by the Census 2000 population.

Figure 3: Average Total Price Appreciation from Breakpoint to Price Level Peak, Tract Groups with Statistically Significant Breakpoints, Weighted by Population



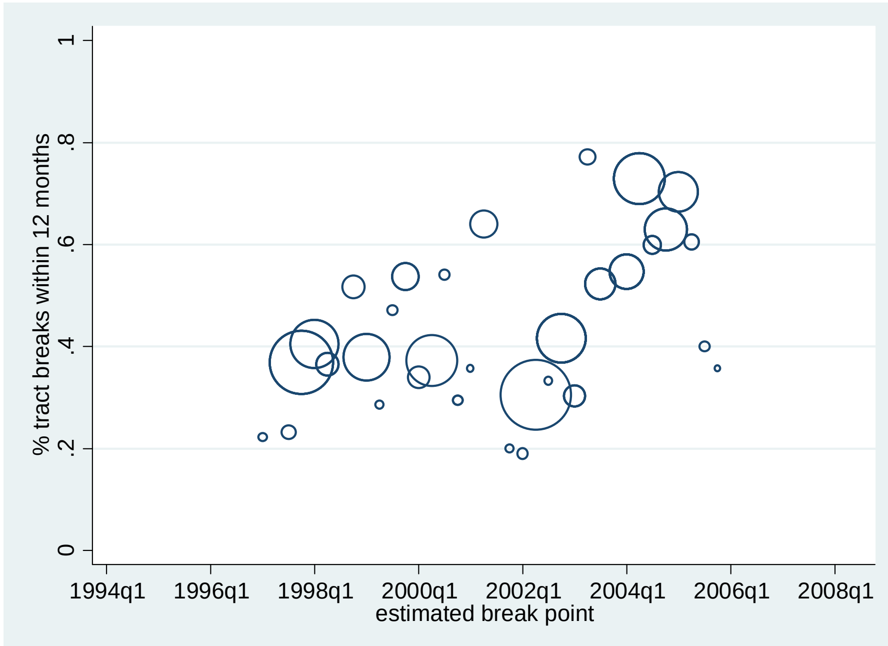
Notes: Each dot represents the average total price appreciation from breakpoint to price level peak. Only tract groups with statistically significant breakpoints are included. Results are averaged by half-year, and weighted by the Census 2000 population.

Figure 4: Average Estimated Magnitude at the Breakpoint, All Tract Groups with Statistically Significant Breakpoints, Weighted by Population



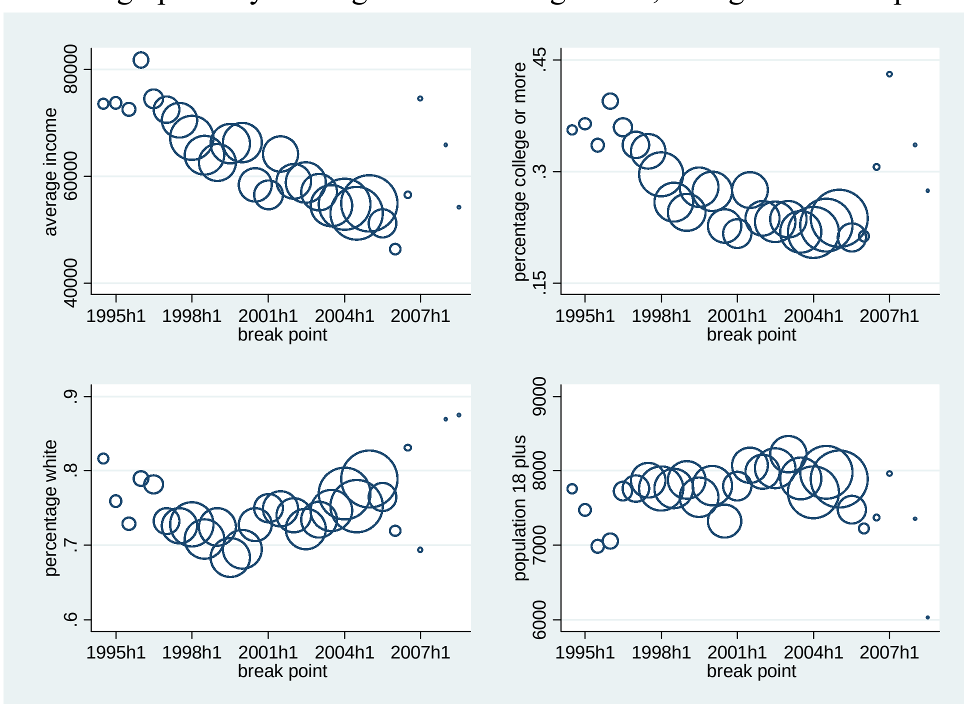
Notes: Each dot represents the average magnitude of the change in growth rates at the estimated tract group breakpoint. Only MSAs with statistically significant breakpoints are included. Results are averaged by quarter, and weighted by the Census 2000 population.

Figure 5. Percentage of Tract Group Breakpoints Within 12 Months of the MSA Breakpoint



Notes: Each dot represents the average percentage of tract group breakpoints within 12 months of the MSA breakpoint. Results are averaged by quarter, and weighted by the Census 2000 population.

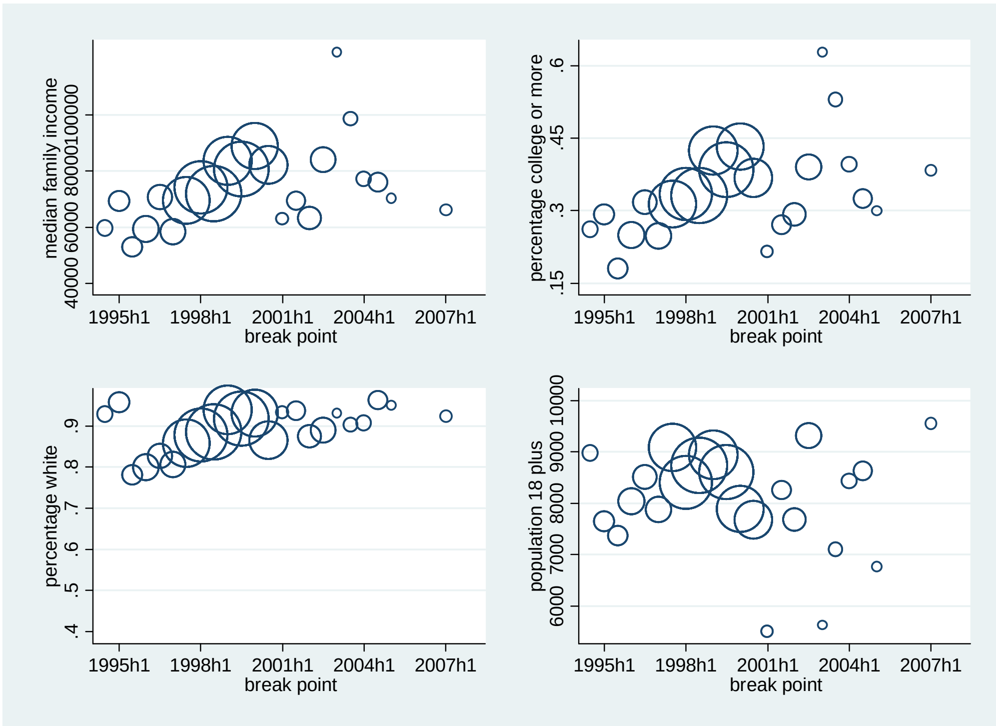
Figure 6: Demographics by Timing of the Housing Boom, Using Tract Group Breakpoints



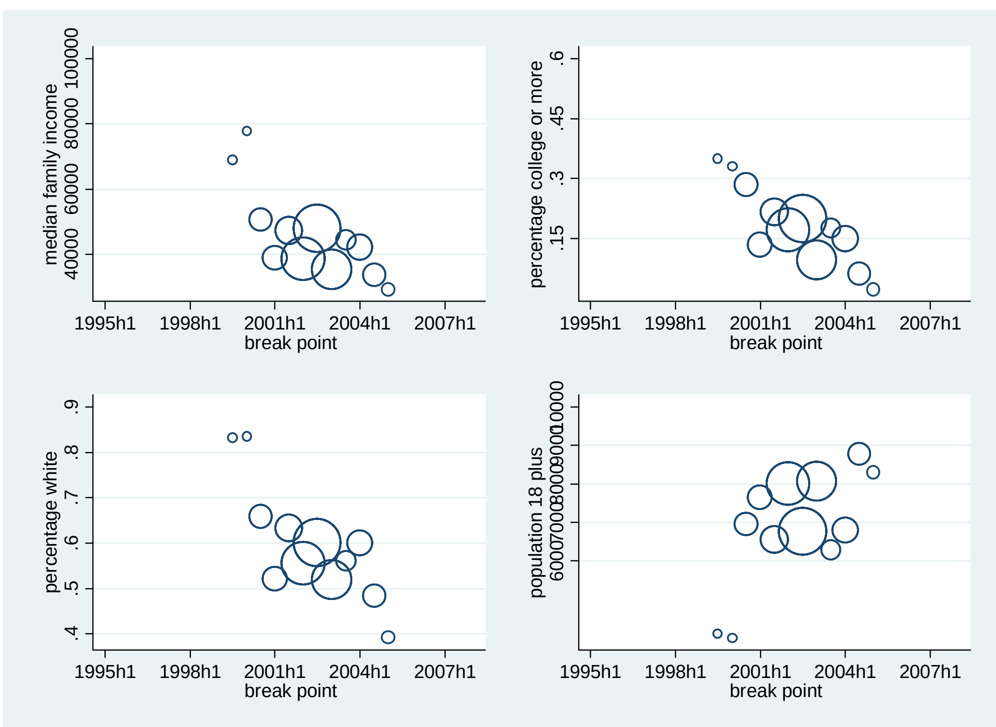
Notes: Each dot represents the average of the sociodemographic variable for all tract groups that had a statistically significant breakpoint. Results are averaged by quarter, and weighted by the Census 2000 population. All sociodemographic variables are based on the Census 2000.

Figure 7: Demographics by Timing of the Housing Boom for Selected MSAs, Using Tract Group Breakpoints

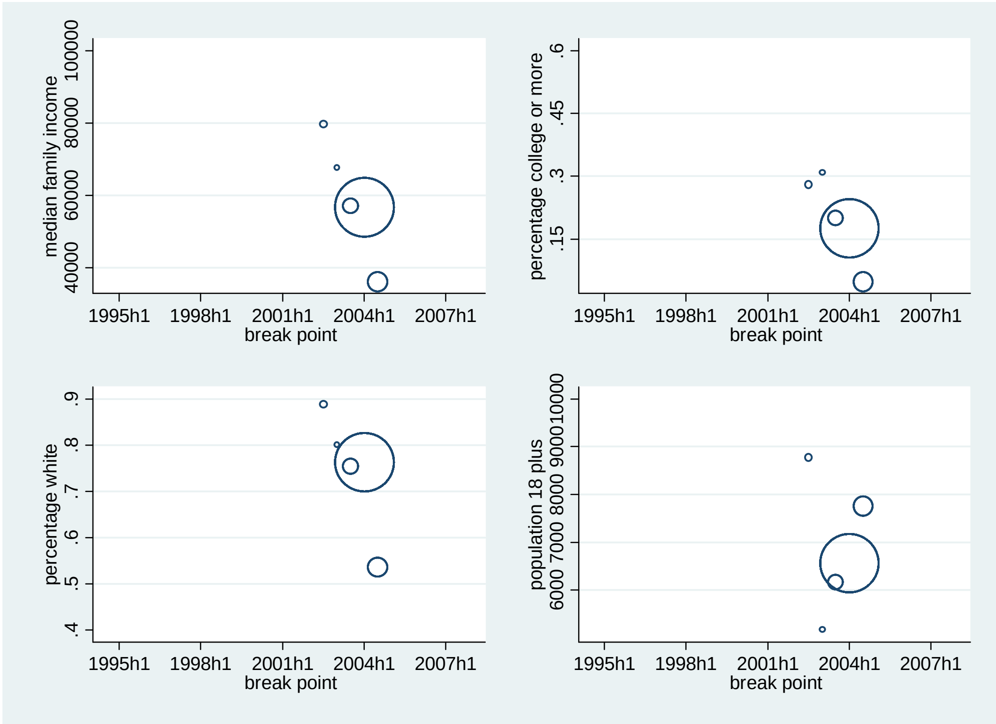
A. Boston



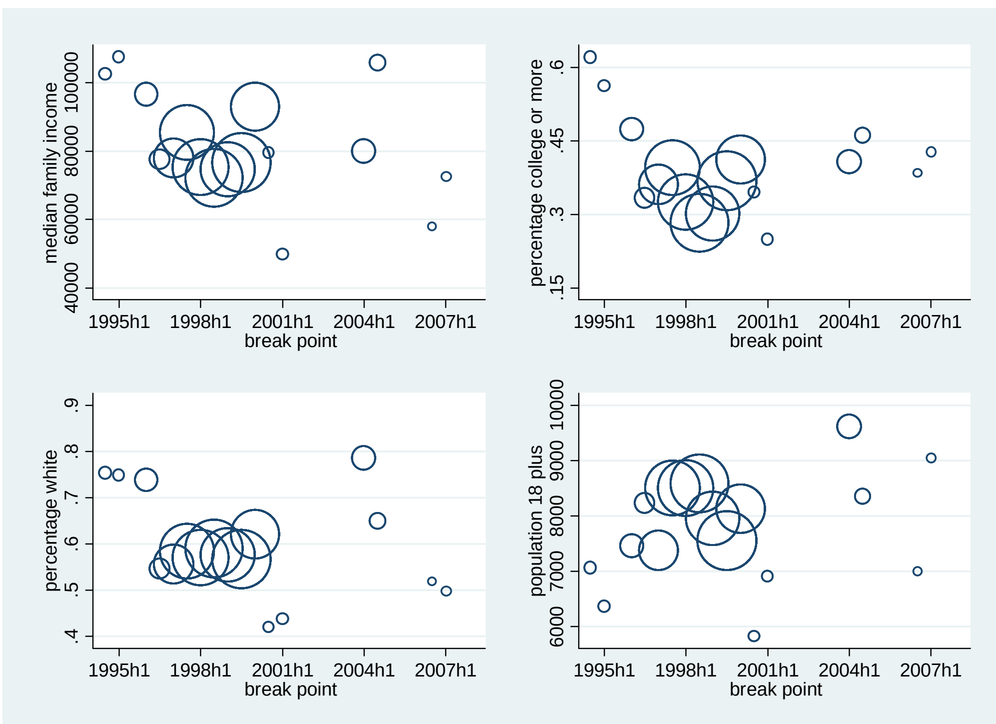
B. Fresno



C. Las Vegas



D. San Francisco



Notes: Each dot represents the average of the sociodemographic variable for all tract groups that had a statistically significant breakpoint. Results are averaged by quarter, and weighted by the Census 2000 population. All sociodemographic variables are based on the Census 2000