Gentrification and Rising Returns to Skill [*,†]

Lena Edlund,[‡] Cecilia Machado,[§] and Maria Sviatschi[¶]

November 13, 2017

Abstract

In 1980, housing prices in large US cities rose with distance from the city center. By 2010, that relationship had reversed. We propose that the inversion can be traced to more hours worked by the skilled. Scarce non-market time downgrades the importance of residential space and upgrades that of proximity to work, factors favoring the central-city location. Geo-coded census micro data covering the 27 largest US cities and the period 1980-2010 support our hypothesis: full-time skilled workers are more likely to locate in the city center and their growth can account for the observed price changes.

Keywords: Gentrification; suburbanization; returns to skill; labor supply; location choice.
JEL: J13, J24

[*]Formely circulated under the title: "Bright Minds, Big Rent: Gentrification and the Rising Returns to Skill." We have benefitted from discussions with Joe Altonji, Nathan Baum-Snow, V. V. Chari, Julie Cullen, David Deming, Jonathan Fisher, Jonathan Heathcote, Matthew Kahn, Jeanne Lafortune, Shirley Liu, Douglas Lucius, Chris Mayer, Enrico Moretti, Marcelo Moreira, Jordan Rappaport, Bernard Salanie, Aloysius Siow, seminar participants at the Minneapolis Federal Reserve, the Barcelona Graduate Summer School 2015, The Econometric Society World Congress August 2015, 2016 Research Symposium on Gentrification and Neighborhood Change at the Federal Reserve Bank of Philadelphia, 2016 Canadian Women Economists Network Luncheon, 31st European Economic Association Annual Congress 2016, Lacea 2016, 2016 mini conference on family/labor/development economics at CUHK, Simon Fraser University and seminar participants in USP, PUC-Rio and Insper.

[†]We thank the New York Federal Statistical Research Data Centers, Baruch, Center for Economic Studies, U.S. Census Bureau. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

[‡]le93@columbia.edu, Columbia University

[§]cecilia.machado@fgv.br, Getulio Vargas Foundation (FGV EPGE)

[¶]msviatschi@princeton.edu, Princeton University

I know things will get better
You'll find work and I'll get promoted
We'll move out of the shelter
Buy a bigger house and live in the suburbs

Tracy Chapman, Fast Car, 1988

The truth is that we are living at a moment in which the massive outward migration of the affluent that characterized the second half of the twentieth is coming to an end.

Ehrenhalt, The Great Inversion, 2013

1 Introduction

In 1980, 2-3 bedroom, single-family residences in major US cities were more expensive outside than inside the 10-mile radius ring, reflecting the then dominant urban pattern – affluent suburbs alongside struggling inner cities. Fast forward to 2010 and the price-distance relationship is the reverse (Figure 1). The ascendancy of housing prices in the city center forms the set piece of what is loosely referred to as gentrification.

At the same time, educated Americans work more. Among the college educated, full-time employment has risen, as has the share working a 50+ hour week (or "long hours" to use the terminology of Kuhn and Lozano [2006]). The increase held for both men and women, mirroring the rise of dual earner households.[1]

Greater labor supply by the skilled, we propose, has revived the city center for the following three reasons. First, skilled jobs remain concentrated in the city center. Second, more hours at work means more precious non-work hours, a fact that boosts the attractiveness of the city center: the commute can be shorter, the high density environment allows for the easy availability of services demanded by the time starved (restaurants, bars, etc.), the suburban space advantage may be less important with little time spent at home.[2] Third, the skilled are important housing-market actors because they can typically outbid the competition – a generalization strengthened by several decades of rising returns to skill.

Longer hours by the skilled is not an inevitability. Higher wages paired with shorter hours

[1] In 1970, about 40 percent of college women living with children under age 18 were in traditional housewife-breadwinner households. By 2010, that number had halved [DVera Cohn and Wang, 2014].

[2] The suburban space advantage extends beyond the home but with longer work hours and less leisure [Aguiar and Hurst, 2009] there may be less time to enjoy space-intensive activities such as golf (which incidentally is also time consuming, "Why golf is in decline in America," April 2nd 2015, *The Economist*).

for the better part of the 20th century [Costa, 2000]. Rapid technological change may lie behind this sea change [Katz and Murphy, 1992, Juhn et al., 1993, Autor et al., 2008]. For instance, skills may depreciate faster in skilled jobs and especially so at times of technological change; there may be a high training component to work in skilled jobs; and/or, the returns to even more skills may be progressively higher. Interestingly, long hours appear disproportionately rewarded [Goldin, 2014].

Our hypothesis relates housing prices to the presence of full-time skilled workers but causal inference based on the correlation of the two is open to criticism. Full-time skilled workers saw healthy income growth which could have made central-city amenities more attractive [Glaeser et al., 2001, Gyourko et al., 2013, Couture and Handbury, 2015]. City real estate could have reached the end of its life cycle, making replacement economical [Brueckner and Rosenthal, 2009]. It could be a cultural reaction against the ubiquity of subdivisions and cul-de-sacs (e.g., Gallagher [2014]). Reverse causality is also a possibility; higher rents may elicit greater labor supply [Johnson, 2012].

Therefore, we turn to a Bartik-type demand shifter [Bartik, 1991] to generate arguably exogenous variation in labor supply (following a number of recent papers, e.g., Diamond [2016], Moretti [2013]). This demand shifter uses a location's employment composition in a base year and national employment trends excluding the locality in question to generate a predicted labor demand.

Our principal data set uses restricted-use micro data from the decennial censuses and the American Community Surveys (ACS), and covers the period 1980-2010. With 1970 as our base year, 1980 is the first decennial census year for which we have a Bartik shifter. The last year, 2010, is dictated by data availability. To obtain sufficient sample size, we use the pooled 5-year ACS sample 2008-2012. We include top-20 cities (by population size) in either 1970 or 2010, 27 cities in total.

We aggregate the micro data to the census tract, and our key dependent variable is the median (owner reported) price for 2-to-3 bedroom homes (single family houses or apartments) in the tract. To measure the presence of full-time skilled workers, we consider two work-hour cut-offs: 40 or 50 hours per week; and two skill cut-offs: four-year college or advanced degree. Further, we limit our sample to the prime working age population (25-55).

Recognizing the role of the IT-revolution in driving spatial variation in demand for skilled labor [Beaudry et al., 2010], we construct the demand shifter from the 1970 city-level employment composition and national growth trends. We interact the city-year demand shifter with a flexible function of distance from the central business district (CBD), thus allowing the shock to propagate differentially through space.

We start by showing that the Bartik shifter correlates positively with the fraction of the population that is skilled and works full time. Other tract demographics such as race, income, or marital status, moved with the Bartik shifter and in the expected direction. However, we view these movements as incidental, derived from the primary relationship linking the Bartik shifter to the full-time and skilled demographics.

We then turn to the relationship between the Bartik demand shifter and housing prices, using a specification that includes city-year and city-distance fixed effects. City-year fixed effects absorb any city-wide changes, for instance city-level labor demand shifts, public safety policies, or credit expansion,[3] as well as any city fixed effects, such as topography, climate, or historically fixed factors such a history of racial tension or the quantity and quality of pre-existing infrastructure, monuments, or cultural institutions. City-distance fixed effects proxies for tract fixed effects. (As a robustness check, we also cross walk tracts to allow for tract fixed effects.) Our most demanding specification includes distance-year fixed effects, which absorb any changes over time to the overall price-distance relationship.

Results indicate that the Bartik demand shifter is associated with an increase in housing prices, and this relationship is stronger closer to the CBD. As the demand shifter prompts an exogenous shift in labor supply, both the reduced form and IV regressions support our hypothesis that the rise in full-time skilled workers has driven gentrification. Further, the IV predictions are in line with observed price changes.

We then turn to address alternative interpretations to the found results. First, we argue that the Bartik demand shifter operates through an increase in hours worked of the skilled, and does not represent a demand shock that uniformly affects the local labor market. We find no effects of the Bartik shifter on the labor supply of the unskilled, as well as do not find evidence of increased presence of skilled workers that do not work long hours in the CBD.

Second, we divide our sample to examine the role of various factors that arguably contributed to central city revival. Violent crime has fallen markedly, a development that benefited the central city disproportionately. However, it is not until the early 1990s that crime falls, crime was high in the 1980s. Therefore we separate the years 1980 and 1990 from 2000 and 2010. We also separate the cities that had large declines in crime from those that saw modest or even increases in crime.

Population growth is another dimension of interest. While the sample cities grew on average, about 1/3rd shrank. The role of road congestion, foreign demand, or limitations on developable land is presumably limited in cities with shrinking populations, such as Detroit, St. Louis, Cleveland, or Baltimore.

[3]For recent papers on the role of the latter, see Glaeser et al. [2010], Favara and Imbs [2015].

It may also be that results are driven by the singular experience of one large city, viz. New York City. Therefore, as a robustness check, we drop New York City.

Throughout above subsamples, we find the Bartik demand shifter to tilt the housing-price profile in favor of the center, suggesting that found effects are not confounded.

Third, yet another possibility is that gentrification is driven by the growing number of high-income households [Glaeser et al., 2001, Gyourko et al., 2013, Couture and Handbury, 2015], the central city being particularly attractive to such households.[4] To separate our hours-based explanation from an income focussed one is challenging considering the high overlap between the two demographics of interest.

While we would like to control for income in our analysis, we lack independent exogenous variation of a second instrument. Alternatively, we can gauge the importance of this competing hypothesis by looking at residential location patterns according to fine income bands. We find that, holding income constant, households in which the head and spouse (if applicable) were skilled, full-time workers clearly favored centrality. This was true both in 1980 and 2010. What changed was the prevalence of such households. In 1980, at all income points, they were a minority. In 2010, beyond the 80th percentile, such households constituted the majority (Figures 8-14).

To sum up, this paper points to gentrification stemming from high-income households working more and therefore favoring living close to work. The city center, long the home of skilled jobs, has thus come to also house its workers. The fact that the city can offer a plethora of other amenities within easy reach enhances its appeal, but is second order to the jobs-location and hours explanation. The distinction may seem academic but a local-amenities based explanation leaves unanswered why country clubs or high-end restaurants could not anchor suburbs.

The remainder of the paper is organized as follows. The balance of this section shows how longer work hours shift the balance in favor of centrality in the standard land-use model. Section 2 describes our empirical strategy and the data. Section 3 presents our results. Section 4 concludes.

1.1 Conceptual framework

This section adapts the canonical land-use model [Mieszkowski and Mills, 1993, Rappaport, 2014, 2016] to examine the role of hours worked, an exercise which ties us closely to an earlier literature which anticipated the centralizing power of a broadside entry of women into labor force [Oi, 1976, White, 1977, Madden, 1980].[5] It also relates us to LeRoy and Sonstelie [1983].

[4]The very top of the income distribution never entirely left for the suburbs [Baum-Snow and Hartley, 2016].

[5]Women's greater familial responsibilities may make them more adverse to commuting. Another margin of adjustment is of course labor supply [Black et al., 2014].

Their model turns on rising income making commuting by car increasingly affordable. Once even the poor can afford a car, gentrification results. This is because added demand by the poor raises suburban housing prices.[6] By contrast, our model focuses on the location choices of the rich once leisure is more scarce. Readers familiar with the above body of literature can skip this section without loss of context.

We assume two types of workers, skilled and unskilled; the skilled earn higher wages than the unskilled. Further, wages and hours are assumed such that the skilled have higher income than the unskilled so that for residential location patterns, we can limit our attention to the preferred location choice of the skilled.

There are two locations city and suburb, indicated by subscripts $i = c$(ity) and $i = s$(uburb), respectively.

The upshot is that if the skilled are required to work long hours, then they prefer to live in the city, whereas if hours are shorter, then the suburb may be preferred. Housing in the location preferred by the skilled commands the higher price – when the skilled prefer the suburb, prices rise with distance to the city center, and vice versa.

All jobs are located in the city whereas residential housing can be in either location. Since all jobs are in the city, suburban location comes with a commute. For simplicity we assume that

$$t = \begin{cases} t_s > 0 & \text{if } i = s, \\ 0 & \text{otherwise.} \end{cases}$$

For concreteness, we let half the population be skilled and command a high wage w, and the remainder unskilled and command a low wage (we do not need a notation for their wage). Workers have 16 hours per day at their disposal. Time can be used in three ways: work, h, commuting, t, and leisure, l. In order to focus on the role of longer hours on location choice, hours are exogenously determined and the same for everybody.

Construction costs are such that only the skilled can afford to build durable housing. The unskilled either live in pre-existing housing or in cheap non-durable housing. Since the latter can be ignored in the analysis, for the purpose of this section, housing will refer to the durable type. Housing accumulate as location specific housing stock, $s_i \geq 0, i = c, s$ which rents at $r_i, i = c, s$ per square foot. Cost of maintenance serves as a floor on rents and for simplicity, it is assumed the same in the city and the suburb, $m_c = m_s = m > 0$.[7]

Initially, only the city is developed. If aggregate demand by the skilled, $d_i(r_i)$, at rent $r_i = m$, exceeds the available housing stock, then housing is built to meet demand.

[6] The affordability of the suburb has if anything increased, Figure 1.

[7] Qualitatively similar, but amplified, results are obtained if the housing stock is fixed.

In the suburb, land is assumed in infinite supply. Thus, the marginal cost of new construction consists of labor and material, both which are elastically supplied. Thus, the marginal cost of construction is a constant $c_s > m$. In the city, land is scarce and we model this by letting the marginal cost of construction depend positively on the existing housing stock $c_c(s_c), c_c(0) = c_s > m, c_c'() > 0$.[8]

Skilled workers rent a_i square feet at the rental price r_i, in one and one location only. The rent is

$$r_i = \begin{cases} m & \text{if } d_i(m_i) \leq s_i, \\ c_i & \text{otherwise, } i = c, s; \end{cases}$$

where the suburban construction cost, c_s, is a constant whereas the city construction cost, c_c, varies positively with the existing housing stock.[9]

Workers derive utility from living space, a, leisure, l, and a numeraire consumption good, x. We let utility be of the following form:

$$u = x^\alpha a^\beta l^\gamma, \ \alpha, \beta, \gamma > 0.$$

The budget constraints are:

$$x + r_i a_i \leq w\,h,$$

and

$$l = 16 - h - t.$$

The optimization problem is a two-stage process: first determine maximal utility in each location, home size chosen optimally at the implied rental price; then pick location. In other words, to pin down the urban location (price) pattern, we need to know where the skilled obtain the highest utility: the city or the suburb? Let a_i^* denote the utility maximizing home size in locality i at the implied rental price r_i. The city gives higher utility if

$$\left(\frac{wh - r_c a_c^*}{wh - r_s a_s^*}\right)^\alpha \left(\frac{a_c^*}{a_s^*}\right)^\beta > \left(\frac{16 - h - t_s}{16 - h}\right)^\gamma. \tag{1}$$

Note that if suburban homes are relatively large (low a_c^*/a_s^*), for instance from abundance of undeveloped land and high wages for the skilled, then the choice tilts in favor of the suburbs

[8] Land scarcity and regulations [Glaeser and Gyourko, 2005, Ortalo-Magne and Prat, 2014], as well as higher cost of vertical construction contribute to making construction costlier in more densely built areas.

[9] Strictly speaking, there is an intermediate case where demand at the cost of maintenance exceeds the existing housing stock but demand is not strong enough to elicit new construction. In that case, rent is indeterminate in the range (m, c_i). This possibility does not change our qualitative results.

(possibly contributing reasons for why the US post-WWII suburbanization pattern did not extend to Western Europe).

Our key question is how longer hours affect the location decisions. The right hand side of inequality 1 goes to 0 as $h \to 16 - t_s$. The left hand side has two components: the city-suburb consumption ratio, which under our functional form assumption is a constant $(\frac{wh - r_c a_c^*}{wh - r_s a_s^*} = \frac{x_c^*}{x_s^*}) =$ 1, and the city-suburb optimal housing size ratio, $\frac{a_c^*}{a_s^*}$.

The ratio of city to suburban home sizes decreases with higher hours because hours raise income and with higher income, more money is spent on housing and the suburbs have more elastic housing supply.[10] But, for the suburb to be favored at hours close to $16 - t_s$, the left hand side $\frac{a_c^*}{a_s^*}$ would have to be close to zero, an unrealistic scenario. Thus, higher hours favor the city, at least for hours close to $16 - t_s$.

For the same reason, a higher skilled wage reduces the left-hand side of inequality 1, and thus makes it more likely that for any given number of hours worked, the suburban location is favored. However, once the right-hand side is close to zero, the city will be the location of choice for the skilled regardless of the wage.

Thus, our model adds a return-to-the city epilogue to the well-known urban narrative of a pre-automobile city in which the skilled and unskilled live side-by-side followed by an automobile enabled phase of suburbanization in which the skilled move to the suburb. Our model highlights the role of moderate work hours in underpinning that suburbanization phase. Longer work hours bring about a third phase of return to the city by the skilled.

We now turn to the data.

2 Data

Our primary data set is drawn from the decennial censuses of 1980, 1990 and 2000, and the American Community Survey's (ACS) pooled 5-year sample for 2010 (years 2008-2012). Our housing-price measure is based on owner-occupied 2-to-3 bedroom single family residences.[11] For home prices, we use owner-provided estimates of value. On the assumption that owners of more recently transacted units would be more knowledgeable about the going market price, we use households ten years or less in their current residence.[12]

We focus on the prime working ages, 25-55.[13] For full time, we consider two cut-offs, 40

[10] The optimal housing size, a_i is given by the implied housing stock from the purchasing power of skilled households, divided by their numbers.

[11] Two and three bedroom homes are modal, for new construction see `https://www.census.gov/construction/chars/pdf/soldbedrooms.pdf`

[12] Including all households yields similar results.

[13] These ages, after college completion but before retirement concerns, are also a key home-buying demographic [National Association of Realtors, 2014].

or 50 hours per week. Likewise, we consider two skill cut-offs, four-year college or graduate degree.

We aggregate the micro data to the census tract and use the tract centroid for the distance calculation.[14]

Further details on data set and variable construction are in the Data Appendix.

2.1 Sample cities

We limit our sample to cities that were in the top-20 (population wise) either in 1970 or 2010. There were 27 such cities which for simplicity we will refer to as "top-20." The list was topped by New York City, Los Angeles and Chicago in both years. The number 20 spot was taken by Phoenix and Memphis in 1970 and 2010 respectively.

Top-20 was chosen because it includes all large cities and a good number of medium sized cities. While the same could be said for say the top-30, we settled for top-20 for reasons of convenience, as expanding the list beyond top 20 brings in a number of similarly sized cities.[15] The restriction to large and mid-size cities is deliberate, however. In addition to their quantitative importance, the agglomeration economies and congestion costs at work behind gentrification are not those of small towns.

Whether administratively part of the city or not, we include tracts within 35 miles of the CBD. (Appendix Table A1 lists the cities and their respective CBDs). This restriction is arbitrary but arguably delineates a commutable area, the outer reaches of which would be about a one-hour commute to the CBD. We have about 65 thousand tract-year observations, or an average of 600 tracts per city and year.[16]

2.2 Bartik-type skilled labor demand shifter

We use the public use version of the decennial censuses (IPUMS) to calculate a city-specific demand shifter for skilled employment from the national growth rates of employment of college

[14] "Census tracts generally have a population size between 1,200 and 8,000 people, with an optimum size of 4,000 people." `http://www.census.gov/geo/reference/gtc/gtc_ct.html`

[15] The reasons are two-fold. First, the US population is mobile, resulting in substantial population changes. Second, while the population gap between cities at the top were sizable enough to maintain the ranking (the largest city, New York, is about twice the size of the runner up, Los Angeles, a difference of four million people), the gap narrows as one moves down the list (as predicted by Zipf's law). For instance, the population difference between number 19 (El Paso, population 649,121) and number 20 (Memphis, population 646,889) on the 2010 list was less than 3,000 – small enough to invite substantial rank mobility. Seattle, Denver, Minneapolis are some of the mid-size cities that were just outside the list, excluded not because we think they are different than some included, similarly sized cities. In fact, one reason for a top-20 cut off rather than say a top-30 is that the cities around that seam were quite similar. In other words, one way to view chosen cities is that we include all the largest cities and a sample of the mid-sized ones.

[16] For disclosure reasons, these numbers are rounded.

workers in industry h, excluding the city in question, $-j$, weighted by each industry's 1970 employment share in city j.

The base year 1970 was chosen because of the role of the IT-revolution in raising the returns to skill and the shift away from manufacturing to high human capital services with significant economies of agglomeration.

Specifically, focusing on ages 25-55, we construct our Bartik demand shifter for skilled labor demand for city j and year t as:

$$Z_{jt} = \frac{1}{N_{j,1970}} \sum_{h}^{41} n_{h,j,1970} \times (\ln n_{h,-j,t} - \ln n_{h,-j,1970}), \tag{2}$$

where

$N_{j,t}$ is the number of workers in city j and year $t = 1980, 1990, 2000, 2010$;

$n_{h,j,t}$ is the number of college educated workers in industry h, city j, year t; and

$n_{h,-j,t}$ is the number of college workers in industry h and year t, excluding city j.

In words, Z_{jt} is the predicted skilled employment growth in city j calculated as the weighted sum of college employment growth nationally, excluding city j, in 41 industries where each industry is weighted according to the number of college workers employed in the industry in the city in 1970.

Z_{jt} is book-ended by San Antonio and Washington DC. In 1980, San Antonio's predicted growth of college educated workers was 9 percent; the number for DC was 23 percent. In 2010, those numbers were 18 and 46 percent respectively (Table A1).

The Bartik shifter is city specific.[17] To allow the demand shifter to propagate differentially throughout the city we interact it with a measure of the tract distance from the CBD. We expect the demand shifter to operate more strongly in tracts close to the CBD (where skilled jobs are concentrated, see e.g., Figure 2).

2.3 Descriptives

Descriptive statistics (means) of the variables are presented in Table 1 by year and four distance intervals: 0-3, 3-10, 10-20, and 20-35 (miles from the CBD). We also present smoothed polynomials based on 1-mile distance intervals in a series of graphs.

We employ the following notation: BA-, less than four-year college; BA, four-year college but no advanced degree; MA+, advanced degree; and BA+ refers to BA and MA+ combined. Further, we let $FT(h, e)$ denote the fraction of adults 25-55 with education $e, e = BA+, MA+$

[17] Except Dallas-Fort Worth, coded as one city in the IPUMS.

and work weeks exceeding $h, h = 40, 50$ hours.

Throughout, housing prices are expressed in 1980 constant dollars.

Home prices Between 1980 and 2010, the median price for 2-to-3 bedroom one-family residences in the top-20 cities rose by 30 percent from \$92.5 to \$120.5 thousand and the increase was greater the more centrally located the tract. In the core (0-3 miles), prices more than doubled. In tracts 3-10 miles out, prices rose by 60 percent, whereas price increases were a mere 10 and 6 percent in tracts 10-20 and 20-35 miles out, respectively – changes large enough to flip the price-distance profile. In 1980, prices in the periphery were 50 percent higher than in the center. By 2010, prices in the center were higher than in the periphery by some 40 percent (Table 1 and Figure 1).

Location of jobs, distance from the CBD Our hypothesis is premised on increasing demand for skilled workers combined with skilled jobs clustering in the city center. Figure 2 shows the fraction of jobs held by a given education group (BA-, BA, BA+, MA+) by distance from the CBD (among all jobs held by adults 25-55 within 35 miles from the CBD in a given year). We see an overall reduction in unskilled jobs and the erosion is particularly pronounced in the city center. In 1980, both skilled and unskilled jobs were concentrated in the city core. By 2010, unskilled jobs had declined markedly inside the 10-mile ring; whereas skilled jobs maintained their concentration in the CBD, and this was particularly pronounced for jobs held by workers with graduate degrees.

Full-time skilled workers Over the study period, the center became more full-time and skilled. In 1980, $FT(40, BA+)$ was higher outside than inside the 10-mile perimeter. By 2010, $FT(40, BA+)$ had risen markedly in the city core (0-3 miles), making up a third of its prime working age population. The shift towards the city center was even more pronounced among those working long (50+) hours (Table 1).[18]

Breakdown by gender reveals that full-time skilled women consistently favored the city core, and more so the longer the hours and the higher the education attained (Figure 3). Over time, however, this group grew substantially.

Full-time skilled men were more willing to live away from the CBD and this was particularly pronounced at the lower hour and skill cutoffs (Figure 4). Over time, however, men, in particular $FT(50, MA+)$ men, approached women's proclivity for the central city (possibly because increasingly they were matched to women with similar education and hours, cf. Costa

[18] Meanwhile, unskilled full-time workers moved away from the center, Figure A7.

and Kahn [2000]).[19]

Since the genders exhibited similar location patterns, we will abstract from gender in the regression analysis.

Age composition Our focus on ages 25-55 may leave out important age groups, e.g., empty nesters.

Table 1 reports the population shares by age group and distance category and we see that the shares of the old and the young have declined overall, whereas the prime working age group increased by 13 percent (from 40 to 45 percent). Furthermore, the increase was concentrated in the 0-3 mile core, where their share went from 39 to 50 percent over the study period. These patterns are further shown in Figure 5. In 1980, the 25-55 age group's representation was j-shaped, bottoming out around 4 miles from the CBD. In 2010, the j had fallen on its back – the prime working ages dominated the city center, decreased with distance for the first 10 miles, to flatline beyond that point.

While young adults, 19 to 24 years old, have long favored the central city, their presence declined. Thus, this demographic does not seem to be a prime driver of gentrification, perhaps unsurprising considering their limited purchasing power.

As for ages 56 and up, this age group has grown overall but the growth was concentrated outside the 10 mile perimeter. Within 10 miles of the CBD, there was a decrease. This was true of both the 56-65 and the 66+ age groups, albeit more pronounced for the latter (Table 1).

Considering the role of public-school quality in the middle-class urban flight narrative, school-age children is an age group of particular interest. In 1980, 6-18 year olds made up about 9 percent of the population in the city core, by 2010 that number was down by a third (Table 1). In other words, the childlessness that characterized early gentrifiers (e.g., yuppies and gays [Black et al., 2002]) remains.

Other demographics Over the period, mean incomes rose by 25 percent and the gains were concentrated in tracts close to the CBD.[20] In the 0-3 mile core, incomes rose by almost 80 percent. Overall, the percent non-Hispanic white declined (12 percentage points), whereas the percent black held constant. In the central city, however, the pattern was reversed: the percent black declined while the percent non-Hispanic white was steady. As for marital status, marriage declined overall; in percentage terms the decline was steeper within 10 miles of the city center (Table 1).

[19] This pattern also held nationally (Figures A1, A2, A3, and A4).

[20] Total personal income, constant dollars.

3 Analysis

This section has four parts. First, we show the direct effect of the skilled-labor demand shock on housing prices. The relationship is positive and greater the closer to the CBD, and holds for a number of cuts of the data suggesting that it is not exclusive to a particular period or cities.

Second, we show that the skilled demand shock raises the share of the population that was full-time and skilled, as would be expected, and the effect was greater the closer to the CBD. Furthermore, and reassuringly, the effect did not extend to the full-time but unskilled, and was substantially weaker (and statistically insignificant) for just the skilled (no hours requirement).

Third, we present OLS and IV results for the relationship between housing prices and skilled labor supply, where we employ the skilled-labor demand shifter as our instrument.

Fourth, we discuss higher incomes as an alternative explanation. While the full-time and skilled group overlaps with "the rich," the two groups are not identical allowing for some daylight between them.

3.1 Housing prices and skilled-labor demand shifter

For the reduced form effect of the skilled-labor demand shock on housing prices we estimate a regression of the following form:

$$PRICE_{ijdt} = \beta_0 \times Z_{jt} + \boldsymbol{F'_{ijdt}}\boldsymbol{\beta_1} \times Z_{jt} + \alpha_j + \alpha_d + \alpha_t + \epsilon_{ijdt}, \tag{3}$$

where $PRICE_{ijdt}$ is the housing price in tract i at distance d from the CBD of city j in year t, Z_{jt} is the exogenous labor demand shock in city j in year t (see equation 2) and $\boldsymbol{F_{ijdt}}$ is an $m \times 1$ vector with functions of tract distance from the CBD, $dist_{ijdt}$, and $\boldsymbol{\beta_1}$ is the conformable vector of parameters. We allow for distance differential impacts of Z_{jt} through $\boldsymbol{F'_{ijdt}}\boldsymbol{\beta_1} \times Z_{jt}$. The coefficient vector of interest is $\boldsymbol{\beta_1}$ and we hypothesize that the effect of Z_{jt} is greater close to the CBD. To capture this distance differential effect, we consider three specifications of $\boldsymbol{F_{ijdt}}$:

$$\boldsymbol{F_{ijdt}} = \begin{cases} D3_d = (d1_d, d2_d, d3_d)' & (m = 3) \quad \text{or;} \\ d1_d, & (m = 1) \quad \text{or;} \\ (dist_{ijdt}, dist^2_{ijdt})' & (m = 2), \end{cases}$$

where $dk_d, k = 1, 2, 3$ indicates the 0-3 mile core, the 3-10 mile ring, and the 10-20 mile ring respectively (the 20-35 mile ring is the reference category). These rings were chosen based on preliminary inspection of the data. In the third specification, the distance differential effect is modelled as a quadratic function of distance.

City, distance and year fixed effects are captured by α_j, α_d and α_t. Distance fixed effects control for factors that vary systematically with distance, viz. housing density and land rents. We discretize the distance fixed effects by using the three distance dummies in $D3_d$. We also consider two-by-two interactions, city-year (α_{jt}), city-distance (α_{jd}), and distance-year (α_{dt}).

City-year fixed effects parcel out time effects common to a city, for instance, city-wide changes to law-and-order or lending policies.

Distance-year fixed effect allow for the overall price-distance relationship to change over time. Parsimony guided our choice of $D3_d$, the chosen intervals captured shared price dynamics.[21]

Ideally, we would like to control for time-invariant tract characteristics but tracts were not constant between the censuses. While there are cross-walk files, the process introduces measurement error. Therefore, in the main analysis, we favor the repeated cross-sectional data where *in lieu* of tract fixed effects we include city-specific distance fixed effects (successive 1-mile radius indicators for the first 20 miles). As a robustness exercise, we cross walk tracts and the results were robust to the inclusion of tract fixed effects.

The inclusion of the full set of fixed effects means that much of the city-year-distance level variation is absorbed – leaving little, but more exogenous, residual variation for recovering our main parameter vector of interest $\boldsymbol{\beta_1}$.

Throughout, we cluster standard errors at the city level and weigh tracts by population size.[22]

Table 2, column 1, shows the results from regressing $PRICE_{ijdt}$ on the demand shifter Z_{jt}, city, distance, and year fixed effects. We see that a higher Z_{jt} is associated with higher housing prices: 1 unit of Z_{jt} is associated with an additional \$522k. The average change in Z was 0.128 and thus the implied price change is about \$70k (out of an observed difference of \$88k, Table 1).

The average effect of the demand shifter Z_{jt} masks substantial heterogeneity (column 2). The effect in the 0-3 mile core is more than twice as large as the effect in the 3-10 mile ring (588.9/234.5=2.51), and more than six times the effect in the 10-20 mile ring. As for the 20-35 mile ring, the effect is positive but not statistically significant (the coefficient on Z_{jt}).

Columns 3-5 sequentially introduce city-year, city-distance and distance-year fixed effects. These fixed effects partial out confounders which correlate with the demand shock and vary at the corresponding level. Column 3 replaces the city- and year fixed effects with the more flexible city-year fixed effects (the main effect of Z_{jt} can no longer be recovered since it varies

[21] Results were similar for finer distance-year fixed effects, therefore the more parsimonious specification was chosen.

[22] Ages 35-55. Regression results not using weights were similar.

at the city-year level). Column 4 replaces distance fixed effects by city-distance fixed effects. The results remain similar.

Column 5 adds distance-year fixed effects, which strengthens the city-core ($d1$) effect. However, the coefficient on the demand shock in the next ring (3-10 miles) becomes statistically insignificant (but not statistically different from the effect in the 10-20 mile ring ($d3$)). Therefore, in column 6, we drop the $d2$ and $d3$ interactions but keep the full set of fixed effects – the core effect remains.

Columns 7-9 let a square polynomial capture the distance-differential effect of Z_{jt} and sequentially introduces city-year, city-distance and distance-year fixed effects. As expected, the demand-shock effect decreases with distance, at a decreasing rate. The introduction of more demanding fixed effects reduces the coefficient estimates, but they remain statistically significant throughout.

As a robustness check, we also estimate the above specifications using logged housing prices (Appendix Table A2).[23] We see similar results, except that once distance-year fixed effects are included, result only survive in the specification using a square polynomial to capture the distance-differential effect (column 9).

Overall, results remain qualitatively similar in the face of different fixed effects – reassuring evidence that the demand shock is largely unrelated to other city-year, city-distance and distance-year varying characteristics. Results are also robust to alternative distance-differential effects in $\boldsymbol{F_{ijdt}}$. Henceforth, we focus on a parsimonious-spline modelling of the distance-differential effects, that is, $\boldsymbol{F_{ijdt}}$ is either $(d1_d, d2_d, d3_d)$ or $d1_d$.

Before turning to our hypothesized mechanism – longer hours raising the salience of centrality – we investigate possible confounders.

Time-invariant tract characteristics To account for time-invariant tract effects, we use cross-walk files to construct a panel data set.[24] The procedure results in loss of fidelity, however. With this caveat in mind, column 1 of Table 3, Panels A and B, shows the results from adding tract fixed effects.[25] Results are quite similar to those found in the cross-sectional data set (cf. column 4 and 5, Table 2), suggesting that city-distance fixed effects serve as passable substitutes.

Lower crime? High crime rates arguably contributed to urban flight [Cullen and Levitt, 1999]. It stands to reason that the drastic decline in crime since the early 1990s would have

[23] Arguably, for owner-occupied homes the primary driver of willingness to pay is the user value rather than the return on investment, one reason to favor prices in levels.

[24] From US2010, `http://www.s4.brown.edu/us2010/Researcher/Bridging.htm`

[25] City-distance fixed effects are absorbed by the tract fixed effects.

the reverse effect.[26] But crime is also a potential left-hand side variable. A better educated and more employed demographic presumably reduces crime rates (e.g., they may be more law abiding and/or have higher demand for law and order).

That notwithstanding, European cities did not experience inner-city crime anywhere near US levels but have seen similar, core-centered urban renewal, e.g., Carpenter and Lees [1995], Boterman et al. [2010]. Further, Ellen and O'Regan [2010] extended Cullen and Levitt's analysis through the 1990s, thus studying periods of both rising and declining crime. They found only a limited role for crime, consistent with the pattern seen in Figure 1: the price-distance relationship pivoted already in the 1980s, years before the decline in violent crime.

Limiting our sample to 1980 and 1990, we find that effects for the years of rising crime are similar to those for the whole period (Table 3, column 2). Results remain statistically significant through the inclusion of city-year and city-distance fixed effects (Panel A). However, once once distance-year fixed effects are added (Panel B) the results for the later period are stronger (column 3). In fact, the results for years 1980 and 1990 lose significance (column 2).[27]

We also split the sample into a high and a low decline group.[28] Qualitatively, results hold in both groups (Table 3, columns 4 and 5).

Some of the cities with the steepest housing price increases saw the largest reductions in crime, New York City being a case in point. Excluding New York City, the effects remain (Table 3, column 6).[29]

In sum, while lower crime has made the central city more livable, the sources of gentrification may lie elsewhere.

Population growth Overall, our sample cities grew over the study period, but 10 of the 27 cities lost population. On average, shrinking cities lost 29% of their population (Appendix Table A8).

Grouping of cities by population growth is interesting for a number of reasons. First, growth in the number of rich households may be one reason prices have risen in the central city. Cities that lost population presumably had less crowding of the rich. Second, real estate is durable

[26]The origins of the rise and fall of crime are debated. A non-exhaustive list includes the removal of school prayers, abortion legalization [Donohue and Levitt, 2001, Foote and Goetz, 2008], more aggressive and targeted policing, greater incarceration rates, electronic surveillance, ATMs and credit cards, the crack epidemic, and environmental factors such as lead exposure [Reyes, 2007].

[27]Since the analysis breaks the sample by only two consecutive years (which are likely similar) adding the very demanding specification with both city-year and distance-year fixed effects absorbs most of the variation of the instrument.

[28]The grouping was done with a view to split the sample in terms of number tracts (Table A7). The decline was measured over the period 1985-2012, the years for which crime statistics are readily available. `http://www.ucrdatatool.gov/Search/Crime/Local/TrendsInOneVarLarge.cfm`

[29]Other reasons to exclude New York City is its population size. Its prominence in finance, business, and culture may also make its real estate market singular.

and it is possible that results are driven by the bottom falling out of the real estate market in cities with shrinking populations. While still a story of the center holding, it would be a more narrow one than that of gentrification. Third, road congestion reduces the time and convenience advantage of automobile travel and thus part of the suburban appeal [Glaeser et al., 2008]).[30] Congestion presumably eased or at least increased less in cities that shrunk. Fourth, subpar economic performance and predominant location in the rust belt suggest limited presence of foreign buyers in cities that shrunk. Fifth, zoning and other restrictions on construction presumably has a limited role in cities that shrunk.

Table 3, shows the results from splitting the sample by population growth (column 7, negative, column 8, positive). We see that the results are quite similar in the two subsamples, corroborating similar price-distance patterns (Appendix Figures A5 and A6). The inclusion of distance-year fixed effects (Panel B) affects the magnitude of the coefficient estimates but not the qualitative results. Population loss, however, resulted in lower overall price pressure. By contrast, in cities that grew, the demand shock propagated *throughout* the city (while being substantially stronger in the 0-3 mile core).

3.2 Full-time skilled work and the demand shifter

We now examine the first order effect of the Bartik shock on labor supply. Table 4, panel A shows the results for $FT(40, BA+)$: full-time workers with a college degree or more. The specifications are those of Table 2 (except its column 2, omitted for brevity).

We see that the demand shock has a positive, but insignificant effect on $FT(40, BA+)$ (column 1), but there is important heterogeneity by distance: most of the effect is concentrated in the 0-3 mile core (column 2) and this result strengthens as fixed effects are added (columns 3 and 4). We also see that inclusion of the full set of fixed effects does not leave enough variation to identify effects outside the 0-3 mile core (column 4). Therefore, we drop the ring interaction terms to focus on the 0-3 mile core (column 5). The coefficient estimate implies an increase in $FT(40, BA+)$ of 16.5 (129.4×0.128) percentage points, or 3/4 of the observed price increase in the 0-3 mile core.

The results for full-time workers with a graduate degree or more, $FT(40, MA+)$, are in panel B. The implied change in $FT(40, MA+)$ in the 0-3 mile core is 7.4 (57.52×0.128) percentage points, or 80% of the observed change (column 5).

Results using $FT(50, BA+)$ and $FT(50, MA+)$ yielded similar results (Appendix Table A3).

So far, we have looked at the intersection of skills and hours. We believe the intersection

[30] The daily commuting time has rise by about eight minutes since 1980 [McKenzie and Rapino, 2011].

of skills and hours is important for two reasons. First, skilled workers have high wages and can thus outbid the competition and resulting in upward pressure on prices in their location of choice. Second, long hours downgrade the importance of spacious residential quarters in favor of proximity to work.

However, gentrification could be driven by full-time unskilled workers or skilled workers regardless of hours. Table 5, column 1 substitutes $FT(40, BA-)$ for the dependent variable. The estimated effect of the skilled demand shock is negative (and borderline significant) in the city core while a positive effect can be found in the third ring (10-20 miles out). Thus it appears that although the skilled-demand shock could have raised demand for unskilled labor as well (e.g., due to complementarities or spill-overs), $FT(40, BA-)$ did not increase in the CBD, consistent with their being priced out. As for the skilled, unconditional on hours worked, we find a weaker centralizing effect of the demand shifter (column 2). Working fewer hours, their willingness to live away from the CBD may be higher.

The remainder of Table 5 presents results for other tract characteristics that plausibly changed with the Bartik demand shifter: percent black, white (non-Hispanic), married, and average personal income. We shall return to the role of income in Section 3.4, but for now note that since these variables correlate with $FT(40, BA+)$, we expect to find a relationship. Indeed, a higher demand shifter is linked to the city center being whiter, more married, and higher income. The reason we favor $FT(40, BA+)$ over these other demographics is its direct and causal link to growth in skill-intensive industries – race, income, or marital status we view as incidental to this primary relationship.

3.3 Housing prices and full-time skilled work

We now turn to our hypothesis: more hours worked by the skilled is behind the ascendancy of centrality. Omitted variables that vary across space and time in a way that correlates with both housing prices and tract demographics would bias results in a simple OLS regression. To guard against such confounding, we use the Bartik-type demand shifter for skilled labor. Under the assumptions that the demand shifter affects housing prices only through the posited channel and is orthogonal to any unobservable tract-by-year characteristics in the housing price equation, we can instrument the full-time skilled measure by the Bartik demand shifter in equation 4.

This identification strategy handles changes in other demographic characteristics that follow from an increase in the full-time skilled population, as long as these characteristics do not enter directly into the housing-price equation. "Permissible" indirect effects include changes to the retail landscape, level of law and order, quality of public schools, etc. that stem from the social, economic or political clout of a skilled, steadily employed population.

More worrying would be if demand growth for skilled labor affected factors that moved housing prices. For instance, a higher Bartik may result in higher tax revenues, allowing for better policing, infrastructure, or civic initiatives that improved quality of life, generally, but more so closer to the city center.

Another possible threat to our identification strategy is that a higher Bartik may signal greater demand for office space in the center – higher home prices in this scenario could be the result of dwindling supply of residential real estate. Demand-induced new construction would work against such an effect. Also, the anecdotal evidence contained in the term "residential conversion" is born out by permitting statistics [Thomas, 2009, 2010, Ramsey, 2012] and population statistics paint a similar picture. Between 2000 and 2010, the central city population grew, reversing decades of population loss [Baum-Snow and Hartley, 2016]. Still, the IV results should be interpreted with these caveats in mind.

Our regression is of the following form:

$$PRICE_{ijdt} = \alpha_1 FT(h,e)_{ijdt} + \alpha_2 dist_{ijdt}\, FT(h,e)_{ijdt} + \alpha_3 dist_{ijdt} + \alpha_{jt} + \alpha_{jd} + \alpha_{dt} + \epsilon_{ijdt}, \quad (4)$$

where the interaction term captures differential effects of FT on housing prices by tract distance from the CBD, and α_{jt}, α_{jd} and α_{dt} are the previously defined city-year, city-distance, and distance-year fixed effects.

Table 6, column 1, shows the OLS results from estimating equation 4 with FT alone. Panel A shows the results for $FT(40, BA+)$ and the estimated coefficient implies an average increase in housing prices by \$21k (1.829×12.57). Column 2 adds distance-year fixed effects and the result remains virtually unchanged. Columns 3 and 4 allow for a distance-differential effect of FT and we see that while of the expected sign and statistically significant (column 3), this distance differential effect wanes once distance-year fixed effects are included (column 4), bringing the estimated effect of full-time skilled workers close to that obtained without the interaction term (column 2). Panel B shows the results for $FT(40, MA+)$ and we see similar results (results using $FT(50, BA+)$ and $FT(50, MA+)$ were also similar, see Appendix Table A4).

Since the OLS results could be biased, we now turn to the IV regression. However, the full set of fixed effects leaves little residual variation for the demand shock exploited in our IV strategy. This is particularly problematic when instrumenting more than one endogenous variable.

Therefore, we present two types of IV specifications. Ones where we instrument FT only and ones where we allow the effect of FT to vary with distance. In the latter case, we need more than one instrument. To that end, we interact the Bartik shifter Z with the three distance dummies. However, the strength of the instruments vary inversely with distance and is further weakened once the full set of fixed effects are included. Therefore, we also present a specification where we include only city-year and city-distance fixed effects.

Table 7 presents the results, Panel A for $FT(40, BA+)$ and Panel B for $FT(40, MA+)$. In column 1, FT is instrumented by $(d1_d, d2_d, d3_d) \times Z_{jt}$. We see a statistically significant positive effect, similar in magnitude to the OLS results. However, statistical significance is lost once distance-year fixed effects are included (column 2). In this specification, the FIRST STAGE effect of the demand shock on FT is confined to the first distance ring (cf. Table 4, column 4); the test for weak instruments also points in this direction (the test statistics indicate that the 2SLS bias could be as large as 30%). This is perhaps unsurprising given that the fixed effects absorb much of the variation in the instruments.

To maintain a strong first stage while including the full set of fixed effects, we restrict the instrument list to $d1_d \times Z_{jt}$ (cf. Table 4, column 5); the resulting IV estimate is twice the size of the OLS estimate, and statistically significant (Table 7, column 3).

Columns 4 and 5 allow the effect of FT to vary with distance to the CBD. The inclusion of two potentially endogenous variables necessitates more than one instrument and we revert to using $(d1_d, d2_d, d3_d) \times Z_{jt}$. In line with expectations, we find evidence of the effect of FT being strongest closer to the CBD, and the negative coefficient on the interaction term remains statistically significant also when the full set of fixed effects is included (unlike the OLS results). Unsurprisingly, however, the inclusion of the full set of instruments weakens the instruments (column 5).

In sum, Table 7 shows that whenever the full set of fixed effects are included and weak instruments are a concern (columns 2 and 5), the IV results line up in magnitude with the OLS results. However, when instruments are strong (columns 3 and 4), the IV results are larger in magnitude.

We take these findings as evidence that the causal effect of FT on housing prices exceeds the OLS estimates. New construction is a candidate reason why OLS would underestimate the causal effect of FT. While the bulk of residential construction is greenfield development, since the early 1990s there has been a marked shift towards brownfield development [Thomas, 2009, 2010, Ramsey, 2012].[31]

Figures 6 and 7 show the results from our two favored specifications. Figure 6 shows the predicted price changes from the Table 7, column 3 specification, favored because of its strong first stage. We see that the model does well close to the CBD but over predicts price changes away from the CBD, suggestive of the presence of a distance differential effect possibly from

[31] We present the IV results for the alternative specifications and samples of Table 3 in the Appendix.

distance-varying housing supply elasticity.

Figure 7 shows the implied price changes from the Table 7, column 4 specification, and we see that the predicted changes track the observed ones closely. The stronger first stage is our reason for favoring this specification, a strength that comes at the price of its only including city-year and city-distance fixed effects.

3.4 Why is this not just about more rich people?

Since 1980, the US population has grown by almost 100 million, and incomes have risen especially in the right tail of the distribution [Kopczuk et al., 2010].[32] Thus, the last decades' rich are both richer and more numerous than in the post-WWII period, which could account for rebound of the central city [Gyourko et al., 2013].

This alternative story emphasizes higher income rather than hours and turns on the central city having an advantage in providing high-end amenities. We now discuss this possibility, starting with the city as provider of consumption amenities and then turning to the hours vs. income distinction.

Consumption amenities Historically significant or difficult-to-repurpose real estate, or facilities with particular space demands have a measure of exogeneity; once in place, there is a strong presumption in favor of their continued existence. Examples include historical buildings, museums, parks, or golf courses. Gentrification could be explained by a shift in tastes from activities where the suburbs have an advantage (e.g., golf) to ones where the central city has an advantage (e.g., museums) [Couture and Handbury, 2015]. However, such an explanation leaves the taste shift unaccounted for.

The city center may also be a natural home for consumption amenities prized by high-income individuals, for instance, restaurants and bars [Glaeser et al., 2001]. Greater demand for such amenities could be one reason for gentrification. But is it truly the case that the central city has an advantage in providing consumption amenities? Restaurants of all stripes can be found outside of city centers.[33] While it may be the case that restaurant offerings in the city-center have improved, that could be the outcome rather than the driver of gentrification [Albouy and Lue, 2015, Hwang and Lin, forthcoming].

The "spreading" through space – one location's good amenities boosting the attractiveness of nearby areas – suggests a handle on endogeneity [Guerrieri et al., 2013]. However, the question of ground zero remains.

[32] See Davis and Heathcote [2007] for the rising share of land rent in housing values.
[33] The Michelin restaurant guide was conjured up as a marketing ploy by the eponymous tire maker.

Reverse commuting has been taken as evidence that urban revival is driven by centrally located consumption amenities [Glaeser et al., 2001, Couture and Handbury, 2015]. Reverse commuting is also consistent with our proposed explanation. Once the center attracts a more affluent demographic, we would expect local amenities to follow suit and to generate their own gravitational field.

Reverse commuters in our hypothesis increasingly have to compete with a high income demographic who in addition to access to local amenities can enjoy living close to work. Therefore, we expect reverse commuting to be modest in size, a conjecture confirmed in our data. Defining reverse commuting as living inside the 3 mile radius core but working outside of it, we found reverse commuters to be a small (about 3 percent) and dwindling share of the prime working age in the 27 cities analyzed and this held whether conditioning on education level and full-time work.[34]

Hours vs. Income The reason we have focused on the full-time and skilled rather than just the rich is the role of leisure scarcity in the location decision. However, the substantial overlap between the two groups makes separation difficult. Ideally, we would like to control for income in the housing price equation, but for that we would need an additional instrument.

Wanting that, we take households of a given income but different $FT(40, BA+)$ status and compare location patterns. Since for the topic at hand, high income households are of particular importance, we focus on the top of the income distribution.

To avoid having the income distribution being mechanically affected by the share of households headed by singles (which has increased), we consider the individual income distribution.[35] Singles are simply assigned their total personal income. For couples, we assume income pooling and each spouse is assigned half of the couple's combined total personal income.[36]

The income distribution is computed separately for each city and year, the underlying assumption being that competition for housing is local to the city. We restrict the sample to household heads and spouses as applicable.[37] For disclosure reasons, we present descriptives by 20-quantiles (ventiles), but results were similar using percentile bins (not reported).

We classify householders as "high" or "low" depending on their $FT(40, BA+)$ status. Singles

[34] The numbers for no-college, just college, and graduate degree was 3.4, 3.1 and 4.2 percent respectively in 1980, and 1.8, 2.6 and 3.6 in 2010. Conditioning on working full time (40 hours per week or more), the numbers were 3.5, 3 and 4 percent in 1980, and 1.8, 2.8 and 3.8 in 2010.

[35] The distinction between households and individuals becomes less important with positive assortative matching [Juhn and Murphy, 1997, Heim, 2007, DiCecio et al., 2008, McGrattan and Rogerson, 2008, Schwartz, 2010], lower degree of intra-household specialization, and more singles.

[36] In the preliminary analysis, we considered alternative ways to compute income for individuals in couples, such as the sum of incomes divided by factors other than 2. Our findings were not sensitive to these alternative calculations.

[37] This restriction is only made for in this section where we analyze location by income.

are high if they have four-year college and work more than 40 hours per week, and low otherwise. For couples, to be in the high group, we require both spouses to have the above attributes.

Figure 8 shows the 1980 and 2010 distance distributions of individuals in the 20-quantiles starting at percentile 50, 75, 85, and 95 respectively. We see that higher income indeed increased the probability of locating in the city center. While there is a pronounced 'suburban hump' for lower incomes, it is much attenuated at top incomes. Interestingly, the tendency of 'the rich' to located close to the city center has not increased despite considerable income growth at the top of the distribution.

To study how the income distribution varies by distance to the CBD, we divide the population into four income groups (0-50, 50-75, 75-95, and 95-100th percentile), and look at their representation in each 1-mile distance ring from the CBD (Figure 9). Focusing on the very first mile, we see that top-5 percent's representation actually shrunk somewhat between 1980 and 2010; this first mile became richer however, thanks to healthy growth of the 50-95 income percentile segment.

For our purposes, however, the most interesting change is that within each income group, high types gained on low types, and the growth was stronger the higher the income group and the closer to the CBD.

To show these changes more clearly, we difference the two years and scale by group size (5 in the case of the 95-100 group, 20 in the case of the 75-95 group, etc.). Figure 10, left panel, shows the changes and we see a clear income divide. The change is most striking for the near suburbs (5-15 miles) where above-median income groups reduced their presence. This change is mirrored near the CBD, with increased representation of the 50-95 percentiles. The right panel breaks out high and low types in each income group and we see that the pattern is driven by the overall growth of high-income high types and their favoring of the center.[38]

So far we have ignored the differences between couples and single. The last decades have seen a growing corps of professional singles, a group who may have particular affinity for city life. Figure 11 shows the same series as Figure 8, but separates couples and singles. We see that neither group increased their propensity to locate in the city center. As for the high/low distinction, Figure 12 shows that for both years, high-type couples are more likely to locate in or near the city center and the gap becomes more pronounced a higher incomes. Figure 13 shows the analogous series for singles and a similar pattern is evident. High-type singles are much more prone to locate centrally.

As a final piece of descriptives, Figure 14 shows the composition of heads and spouses, by

[38]Figure 10 is also suggestive of why price developments in exclusive suburbs have been much more muted. For instance, while real estate prices in Manhattan rose by 33% in the last 10 years, they shrunk in suburbs such as White Plains and Chappaqua, affluent towns in the NYC metro area, `http://www.neighborhoodscout.com/`.

type (high/low) and marital status (single/couple) by income percentile (20-quantiles, centered on the marker point). We see substantial growth of high types, particularly towards the top of the distribution. In 1980, the majority of individuals were in low-type households and this was true throughout the income distribution. In 2010, past the 80th percentile, high types constitute the majority.

Thus, we find that full-time skilled households increased over our study period and the increase was greater towards the top of the income distribution. Further, while single households increased, high-type couple households dominated growth at the top. For both singles and couples, and throughout the income distribution[39] high-type households were more likely to locate in or close to the CBD relative to their low-type counterpart. This was true both in 1980 and 2010. Looking at the income and type distribution by distance to the CBD, two things stand out: (i) the growth of higher income groups in the CBD was broad based, extending beyond the very rich; and (ii) the gain was driven by full-time skilled workers.

4 Discussion

Suburbanization and central-city decline dominated the US urban landscape for most of the 20th Century. As the century drew to a close, however, change was clearly underway. Today, gentrification has outgrown its erstwhile niche status and headlines a broad-based revival of the central city.

The underlying factor, we have proposed, can be found in high-income households spending more time at work and less at home, as exemplified by the decline of the traditional breadwinner-housewife household in favor of dual earner couples and singles at the top of the income distribution.

That this development favors central city over suburban location follows from standard models of land and time use – suburban living offers space against a commute. Less non-market time reduces the enjoyment of the former and heightens the burden of the latter. To the best of our knowledge, however, this is the first paper to emphasize dwindling non-market time of high-income households as the driver of gentrification.

The empirical analysis has pooled men and women, despite the importance of women's entry into the labor force for reducing high-income households' non-market time. In the preliminary analysis we looked at men and women separately but did not find gender differential effects of note, possibly because the expansion of hours has been qualitatively similar for skilled men and women. A counter factual worth considering, however, is whether gentrification would have

[39]Except at the very bottom, not reported.

grown out of its niche status if married women had stayed home.

Gentrification is not unique to the US, but decades of suburbanization lends it a distinct character, e.g., Saint-Paul [2015]. US cities' lacking in historically significant real estate has been pointed to as one explanation for the different starting points [Hohenberg and Lees, 1995, Brueckner et al., 1999]. Our paper suggests another: the greater purchasing power of the US middle class. Circa 1850, US overtook Western Europe in terms of GDP per capita, and by 1950 the latter was a mere half of the former.[40] In other words, the limited purchasing power of the European middle class post WWII may have contributed to their favoring existing city-core real estate over new construction, leaving the latter for low-cost or public housing.

In future, the current trend of long hours among "the rich" may stall or reverse. If so, another wave of urban flight may be in the cards as the affluent find time to golf and garden. In the meantime, however, a combination of high density housing and mass transit can help keep the city not just productive, but also affordable and livable.

[40] The Maddison-Project `http://www.ggdc.net/maddison/maddison-project/home.htm`

References

Mark Aguiar and Erik Hurst. *The Increase in Leisure Inequality 1965-2005*. The AEI Press, 2009.

David Albouy and Bert Lue. Driving to opportunity: Local rents, wages, commuting, and sub-metropolitan quality of life. *Journal of Urban Economics*, 89:74 – 92, 2015.

David H. Autor, Lawrence F. Katz, and Melissa S. Kearney. Trends in U.S. wage inequality: Revising the revisionists. *The Review of Economics and Statistics*, 90(2):300–323, May 2008.

Timothy J. Bartik. *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research, Kalamazoo, MI, 1991.

C. F. Baum. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *Stata Journal*, 7(4):465–506, 2007.

Nathaniel Baum-Snow and Daniel Hartley. Causes and consequences of central neighborhood change, 1970-2010. May 2016.

Paul Beaudry, Mark Doms, and Ethan Lewis. Should the Personal Computer Be Considered a Technological Revolution? Evidence from US Metropolitan Areas. *Journal of Political Economy*, 118(5):988–1036, October 2010.

Dan Black, Gary Gates, Seth Sanders, and Lowell Taylor. Why do gay men live in San Francisco? *Journal of Urban Economics*, 51(1):54 – 76, 2002.

Dan A. Black, Natalia A. Kolesnikova, and Lowell J. Taylor. Why do so few women work in New York (and so many in Minneapolis)? labor supply of married women across U.S. cities. *Journal of Urban Economics*, 79(January):59–71, 2014.

Willem R. Boterman, Lia Karsten, and Sako Musterd. Gentrifiers settling down? patterns and trends of residential location of middle-class families in Amsterdam. *Housing Studies*, 25(5): 693 – 714, 2010.

Jan K. Brueckner and Stuart S. Rosenthal. Gentrification and neighborhood housing cycles: will America's future downtowns be rich? *Review of Economics and Statistics*, 91(4):725–743, November 2009.

Jan K. Brueckner, Jacques-Francois Thisse, and Yves Zenou. Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review*, 43(1):91 – 107, 1999.

Juliet Carpenter and Loretta Lees. Gentrification in New York, London and Paris: an international comparison. *International Journal of Urban and Regional Research*, 19(2):286, 1995.

Dora Costa and Matthew E. Kahn. Power couples: Changes in the locational choice of the college educated, 1940-1990. *Quarterly Journal of Economics*, 115(4):1287–1315, November 2000.

Dora L. Costa. The wage and the length of the work day: From the 1890s to 1991. *Journal of Labor Economics*, 18(1):156–181, 2000.

Victor Couture and Jessie Handbury. Urban revival in America, 2000 to 2010. November 2015.

Julie B. Cullen and Steven D. Levitt. Crime, urban flight, and the consequences for cities. *Review of Economics and Statistics*, 81(2):159–169, May 1999.

Morris A. Davis and Jonathan Heathcote. The price and quantity of residential land in the United States. *Journal of Monetary Economics*, 54(8):2595 – 2620, 2007.

Rebecca Diamond. The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 2016.

Riccardo DiCecio, Kristie M. Engemann, Michael T. Owyang, and Christopher H. Wheeler. Changing trends in the labor force: A survey. *Federal Reserve Bank of St. Louis Review*, 90 (1), January/February 2008.

John J. III Donohue and Steven D. Levitt. The impact of legalized abortion on crime. *Quarterly Journal of Economics*, 116(2):379–420, May 2001.

Gretchen Livingston DVera Cohn and Wendy Wang. After decades of decline, a rise in stay-at-home mothers. Social & demographic trends, Pew Research Center, April 2014.

Lena Edlund, Cecilia Machado, and Maria Sviatschi. Bright minds, big rent: Gentrification and the rising returns to skill. Working Paper 21729, National Bureau of Economic Research, November 2015.

Ingrid Gould Ellen and Katherine O'Regan. Crime and urban flight revisited: The effect of the 1990s drop in crime on cities. *Journal of Urban Economics*, 68(3):247 – 259, 2010.

Giovanni Favara and Jean Imbs. Credit supply and the price of housing. *American Economic Review*, 105(3):958–92, 2015.

Christopher Foote and Christopher Goetz. The impact of legalized abortion on crime: Comment. *Quarterly Journal of Economics*, 123(1), February 2008.

Leigh Gallagher. *The End of the Suburbs: Where the American Dream Is Moving*. Portfolio/Penguin, 2014.

Edward L. Glaeser and Joseph Gyourko. Urban decline and durable housing. *Journal of Political Economy*, 113(2):pp. 345–000, 2005.

Edward L. Glaeser, Jed Kolko, and Albert Saiz. Consumer city. *Journal of Economic Geography*, 1(1):27–50, 2001.

Edward L. Glaeser, Matthew E. Kahn, and Jordan Rappaport. Why do the poor live in cities? the role of public transportation. *Journal of Urban Economics*, 63(1):1 – 24, 2008.

Edward L. Glaeser, Joshua D. Gottlieb, and Joseph Gyourko. Can cheap credit explain the housing boom. Technical report, Harvard University and Wharton, 2010.

Claudia Goldin. A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119, 2014.

Veronica Guerrieri, Daniel Hartley, and Erik Hurst. Endogenous gentrifcation and housing price dynamics. *Journal of Public Economics*, 100:45–60, 2013.

Joseph Gyourko, Christopher Mayer, and Todd Sinai. Superstar cities. *American Economic Journal: Economic Policy*, 5(4):167–99, 2013.

Bradley T. Heim. The incredible shrinking elasticities: Married female labor supply, 1978-2002. *The Journal of Human Resources*, 42(4):pp. 881–918, 2007.

Paul M. Hohenberg and Lynn Hollen Lees. *The making of urban Europe, 1000-1994*. Harvard University Press, Cambridge, MA, 1995.

Jackelyn Hwang and Jeffrey Lin. What have we learned about the causes of recent gentrification? *Cityscape*, forthcoming.

William R. Johnson. House prices and female labor force participation. University of Virginia, January 2012.

Chinhui Juhn and Kevin M. Murphy. Wage inequality and family labor supply. *Journal of Labor Economics*, 15(1):72–97, January 1997.

Chinhui Juhn, Kevin M. Murphy, and Brooks Pierce. Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3):410–442, June 1993.

Lawrence F. Katz and Kevin M. Murphy. Changes in the relative wages, 1963-87: supply and demand factors. *Quarterly Journal of Economics*, 107(1):35–78, February 1992.

Wojciech Kopczuk, Emmanuel Saez, and Jae Song. Earnings inequality and mobility in the united states: Evidence from social security data since 1937. *Quarterly Journal of Economics*, 125(1):91 − 128, 2010.

Peter J. Kuhn and Fernando A. Lozano. The expanding workweek? understanding trends in long work hours among U.S. men, 1979-2004. Discussion Paper 1924, IZA, 2006.

Stephen F LeRoy and Jon Sonstelie. Paradise lost and regained: Transportation innovation, income, and residential location. *Journal of Urban Economics*, 13(1):67–83, 1983.

Janice Fanning Madden. Consequences of the growth of the two-earner family urban land use and the growth in two-earner households. *American Economic Review*, 70(2):191–197, 1980.

Ellen McGrattan and Richard Rogerson. Changes in the distribution of family hours worked since 1950. Research Department Staff Report 397, Federal Reserve Bank of Minneapolis, 2008.

Brian McKenzie and Melanie Rapino. Commuting in the United States, 2009. American Community Survey Reports ACS-15, Census Bureau, Washington, DC, 2011.

Peter Mieszkowski and Edwin S. Mills. The causes of metropolitan suburbanization. *Journal of Economic Perspectives*, 7(3):135–147, Sum 1993.

Enrico Moretti. Real wage inequality. *American Economic Journal-Applied Economics*, 5(1): 65–103, Jan 2013.

National Association of Realtors. Home buyer and seller generational trends. Technical report, 2014. http://www.realtor.org/sites/default/files/reports/2014/2014-home-buyer-and-seller-generational-trends-report-full.pdf.

Walter Oi. Residential location and labor supply. *Journal of Political Economy*, 84(4):S221–S238, 1976.

Francois Ortalo-Magne and Andrea Prat. On the political economy of urban growth: Homeownership versus affordability. *American Economic Journal: Microeconomics*, 6(1):154–81, 2014.

Kevin Ramsey. Residential construction trends in Americas metropolitan regions. Technical report, U.S. Environmental Protection Agency, 2012.

Jordan Rappaport. Monocentric city redux. Research Working Paper 14-09, The Federal Reserve Bank of Kansas City, 2014.

Jordan Rappaport. Productivity, congested commuting, and metro size. Research Working Paper 16-03, The Federal Reserve Bank of Kansas City, 2016.

Jessica Wolpaw Reyes. Environmental policy as social policy? The impact of childhood lead exposure on crime. *B E Journal of Economic Analysis & Policy*, 7(1), 2007.

Gilles Saint-Paul. Bobos in paradise: Urban politics and the new economy. Discussion Paper 10879, Centre for Economic Policy Research, 2015.

Christine R. Schwartz. Earnings inequality and the changing association between spouses earnings. *American Journal of Sociology*, 115(5):pp. 1524–1557, 2010.

James H. Stock and Motohiro Yogo. Testing for weak instruments in linear iv regression. In Donald W. K. Andrews and James H. Stock, editors, *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge University Press, 2005.

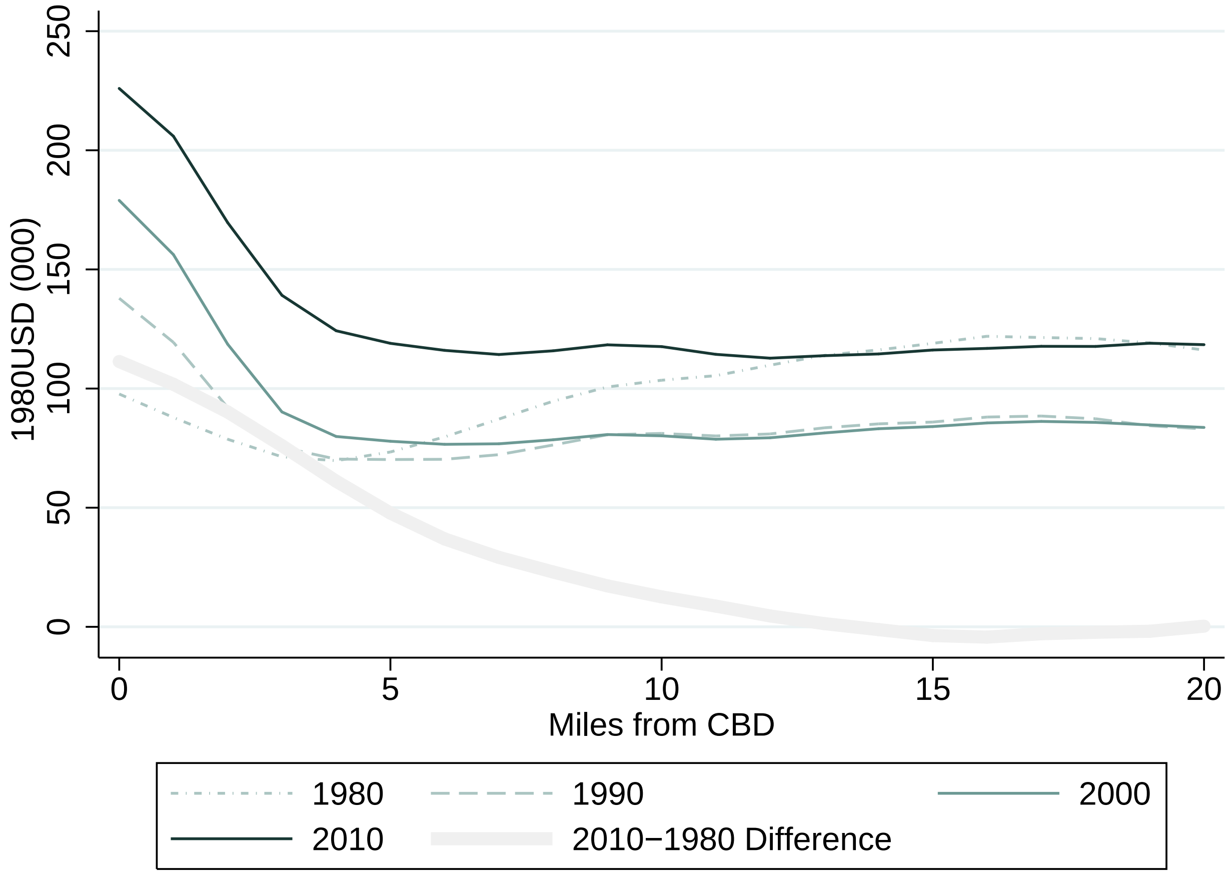
John Thomas. Residential construction trends in Americas metropolitan regions. Technical report, U.S. Environmental Protection Agency, 2009.

John Thomas. Residential construction trends in Americas metropolitan regions. Technical report, U.S. Environmental Protection Agency, 2010.

Michelle J. White. A model of residential location choice and commuting by men and women workers. *Journal of Regional Science*, 17(1):41–51, 1977.

Graphs

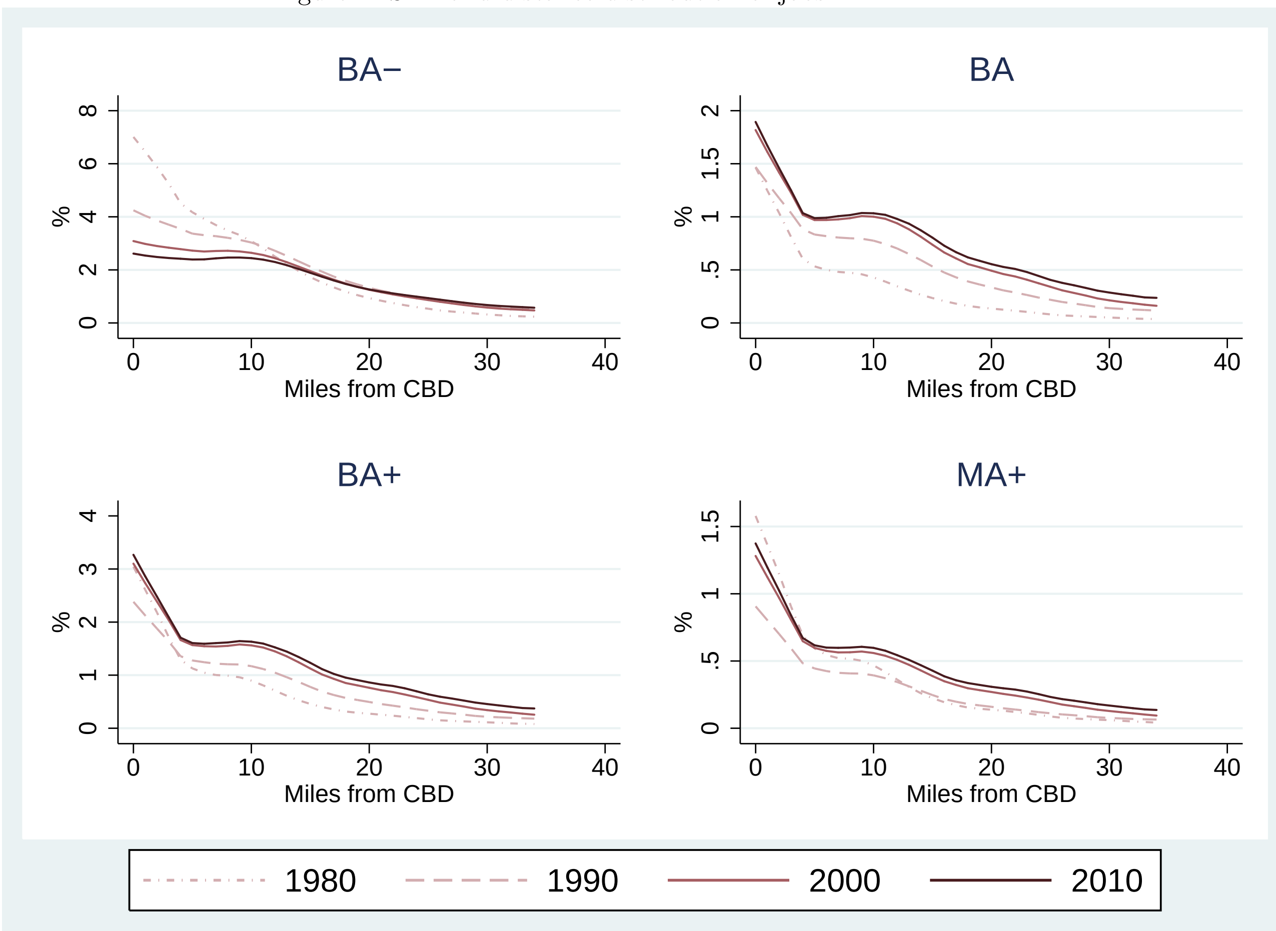
Figure 1: Home prices by distance from the CBD



Source: 1980, 1990, 2000 census and 5-year sample for 2010 American Community Survey, restricted use data. The sample corresponds to tracts in our top-20 US cities (27 in all, see table A1), within 35 miles of the CBD.

Notes: The figure shows the median home price in tract by the tract's distance from the CBD, for each census year. Home prices (in 1980 dollars) are for owner occupied, 2-3 bedrooms, and one-family units. 20 miles includes 20-35 miles. See the Appendix for further variable and sample construction details.

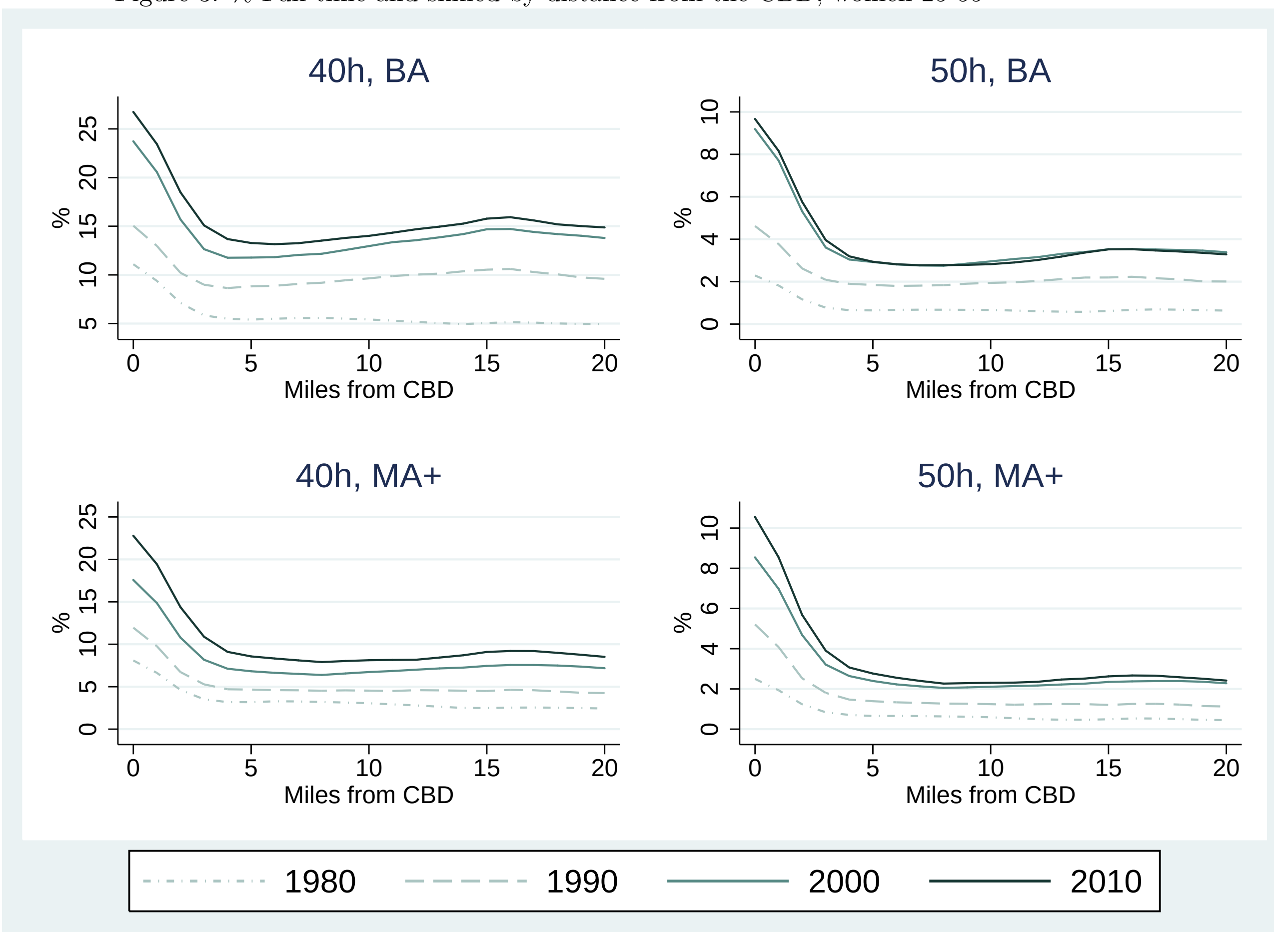
Figure 2: Skill and distance distribution of jobs



Source: 1980, 1990, 2000 census and 5-year sample for 2010 American Community Survey, restricted use data. The sample corresponds to tracts in our top-20 US cities (27 in all, see table A1), within 35 miles of the CBD.

Notes: The figure shows the percent of jobs held by individuals (ages 25-55) with education $e = BA-, BA, BA+, MA+$ by job location (tract) distance from the CBD, and census year. BA+ corresponds to a four-year college degree or more; BA- is its complement. BA corresponds to a four-year college degree only; MA+ corresponds to an advanced degree. Each line integrates to the fraction of jobs help by an education group in a given year. For example, in 2010 about 1.8×35=60% of all jobs within 35 miles were held by individuals with less than a BA degree. See the Appendix for further variable and sample construction details.

Figure 3: % Full time and skilled by distance from the CBD, women 25-55



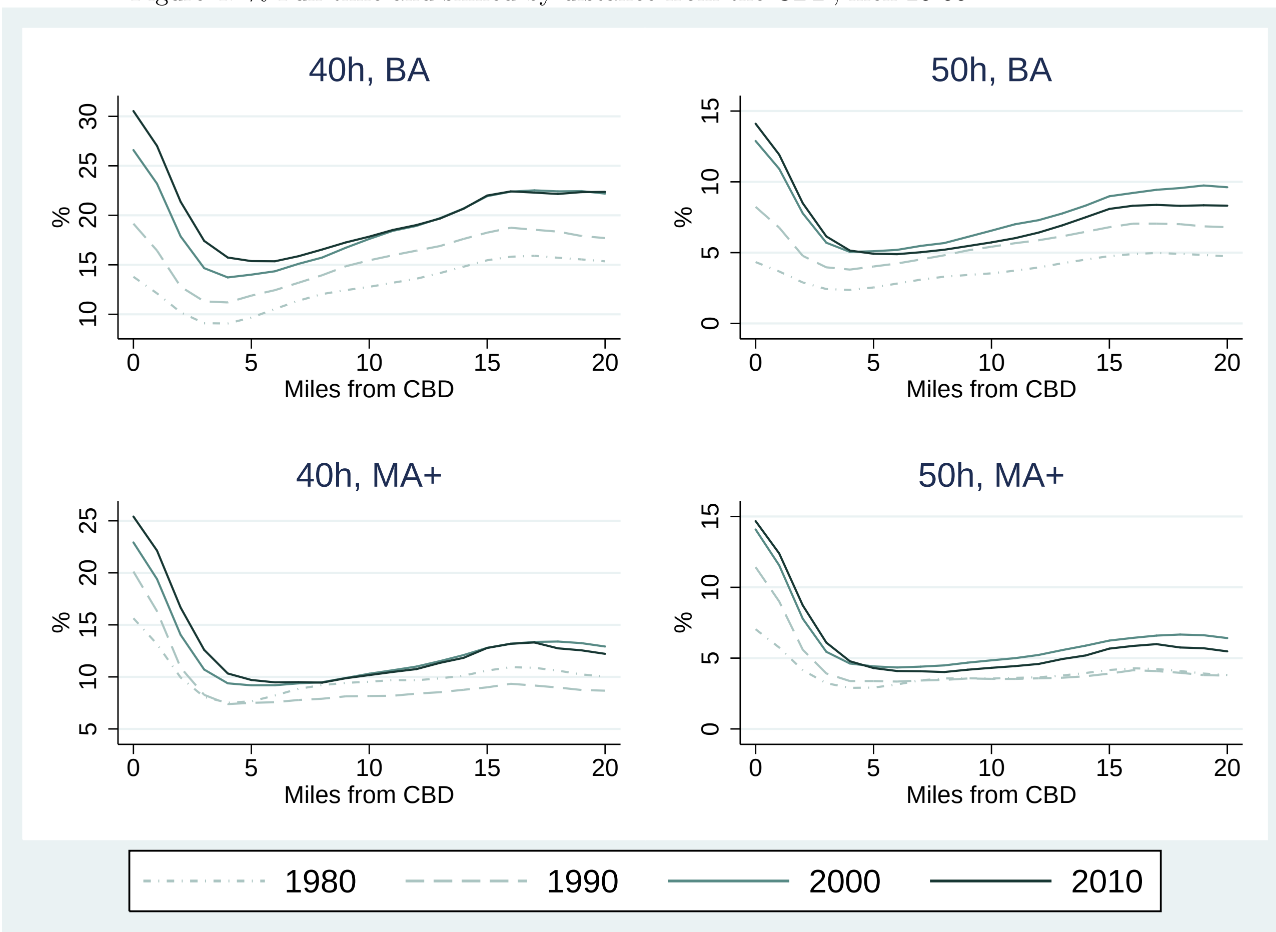
Source: 1980, 1990, 2000 census and 5-year sample for 2010 American Community Survey, restricted use data. The sample corresponds to tracts in our top-20 US cities (27 in all, see table A1), within 35 miles of the CBD.

Notes: The figure shows the percent of women (25-55) that work more than $h = 40, 50$ hours/week and have education $e = BA, MA+$ by tract distance from the CBD and census year. Values at distance 20 gives the average for the distance interval 20-35 miles.

BA corresponds to a four-year college degree and MA+ corresponds to an advanced degree or more.

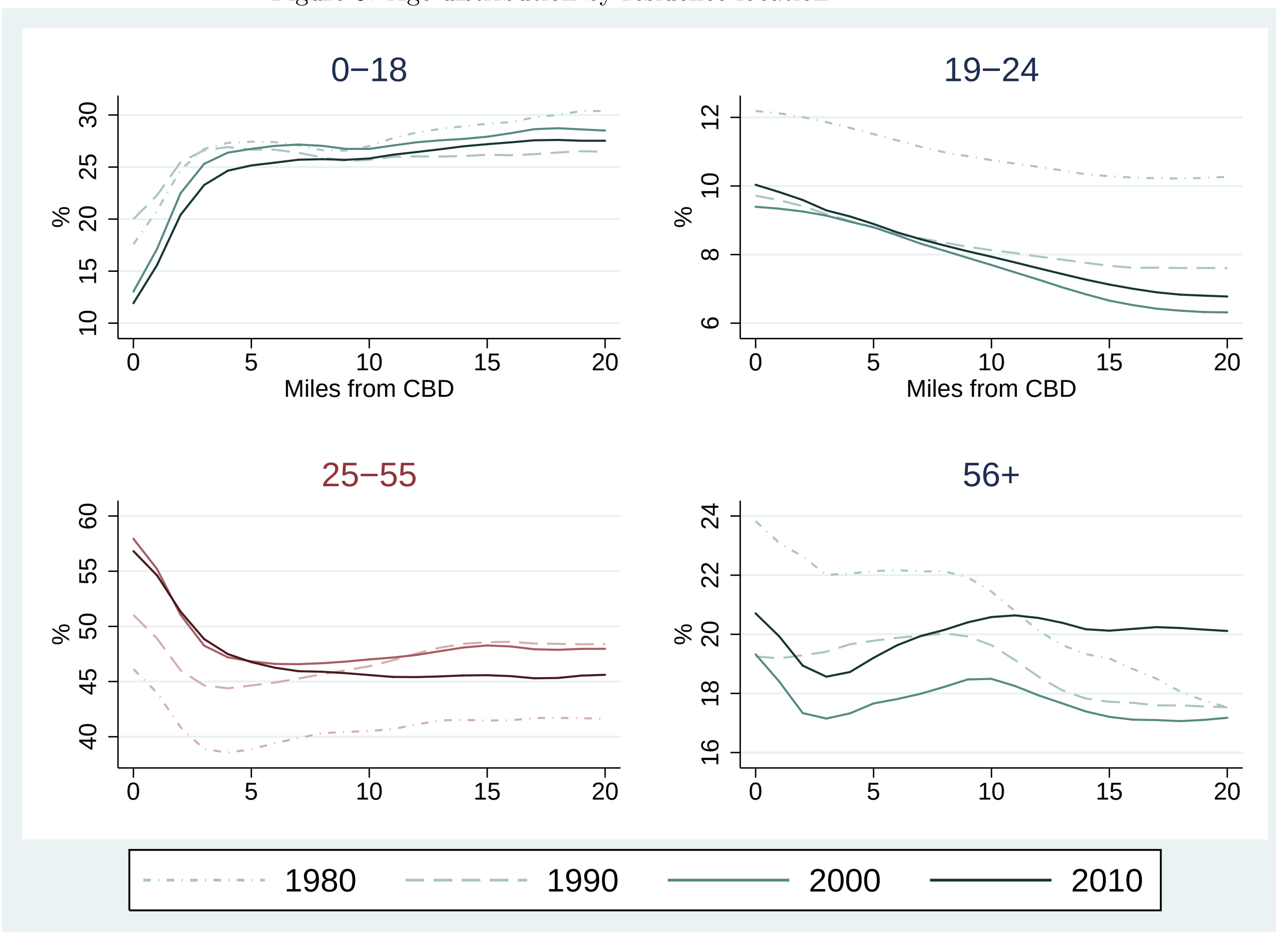
For example, in 2010, around 26% of women (25-55) living in the CBD (miles from CBD=0) had a BA degree and worked 40 hours/week or more. See the Appendix for further variable and sample construction details.

Figure 4: % Full time and skilled by distance from the CBD, men 25-55



Source: 1980, 1990, 2000 census and 5-year sample for 2010 American Community Survey, restricted use data. The sample corresponds to tracts in our top-20 US cities (27 in all, see table A1), within 35 miles of the CBD.
Notes: The figure shows the percent of men (25-55) that work more than $h = 40, 50$ hours/week and have education $e = BA, MA+$ by tract distance from the CBD and census year. Values at distance 20 gives the average for the distance interval 20-35 miles.
BA corresponds to a four-year college degree and MA+ corresponds to an advanced degree or more.
For example, in 2010 around 30% of men (25-55) living in the CBD (miles from CBD=0) had a BA degree and worked 40 hours/week or more.
See the Appendix for further variable and sample construction details.

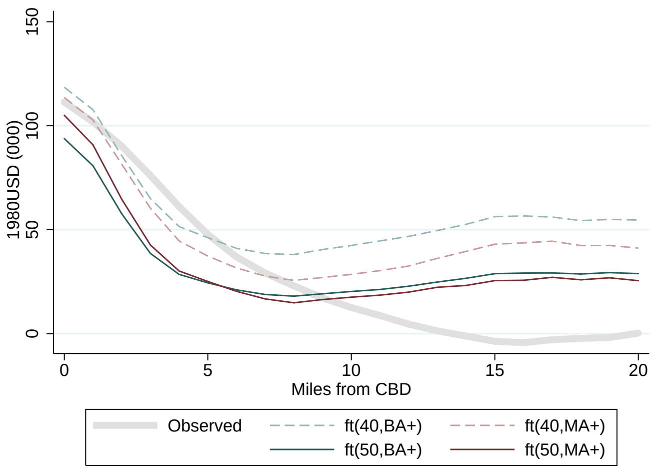
Figure 5: Age distribution by residence location



Source: 1980, 1990, 2000 census and 5-year sample for 2010 American Community Survey, restricted use data. The sample corresponds to tracts in our top-20 US cities (27 in all, see table A1), within 35 miles of the CBD.

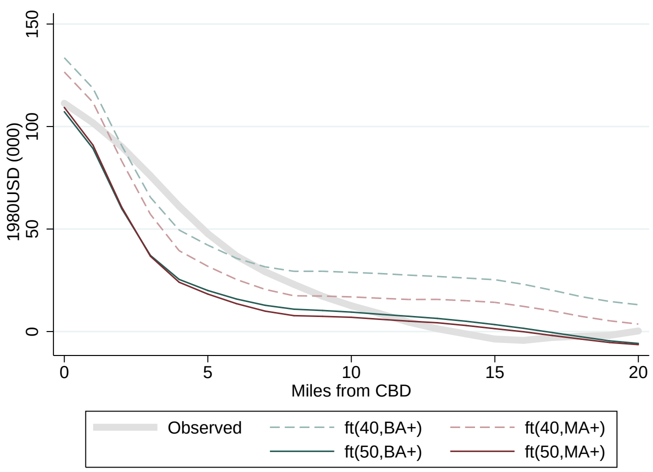
Notes: The figure shows the (tract) population age distribution by tract distance to the CBD. Values at distance 20 gives the average for the distance interval 20-35 miles.
For example, in 2010 about 12% of individuals living in the CBD (miles from CBD=0) belonged to age group 0-18. See the Appendix for further variable and sample construction details.

Figure 6: Predicted v. Actual Price increase, '000 1980$



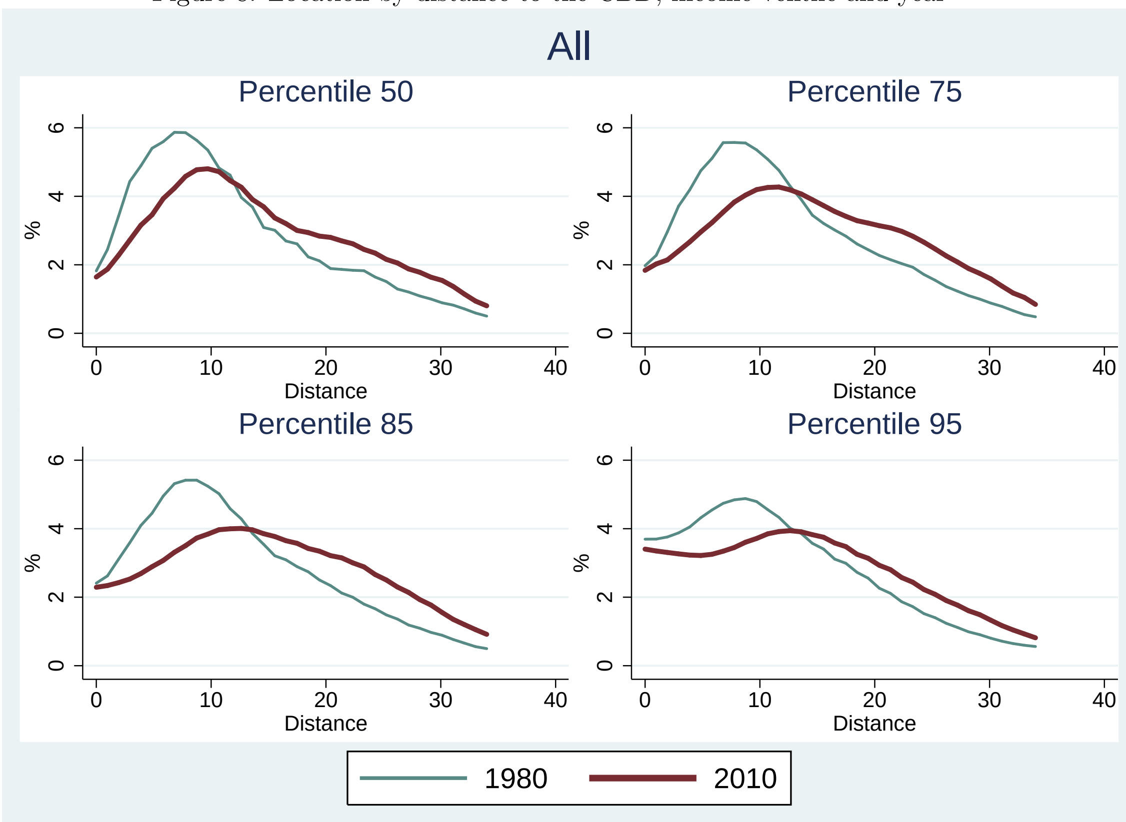
Notes: The figure shows the predicted price change using the regression results from Table 7, column 3 and observed changes in the explanatory variable $FT(h, e)$ by distance to the CDB. Values at distance 20 gives the average for the distance interval 20-35 miles.

Figure 7: Predicted v. Actual Price increase, '000 1980$



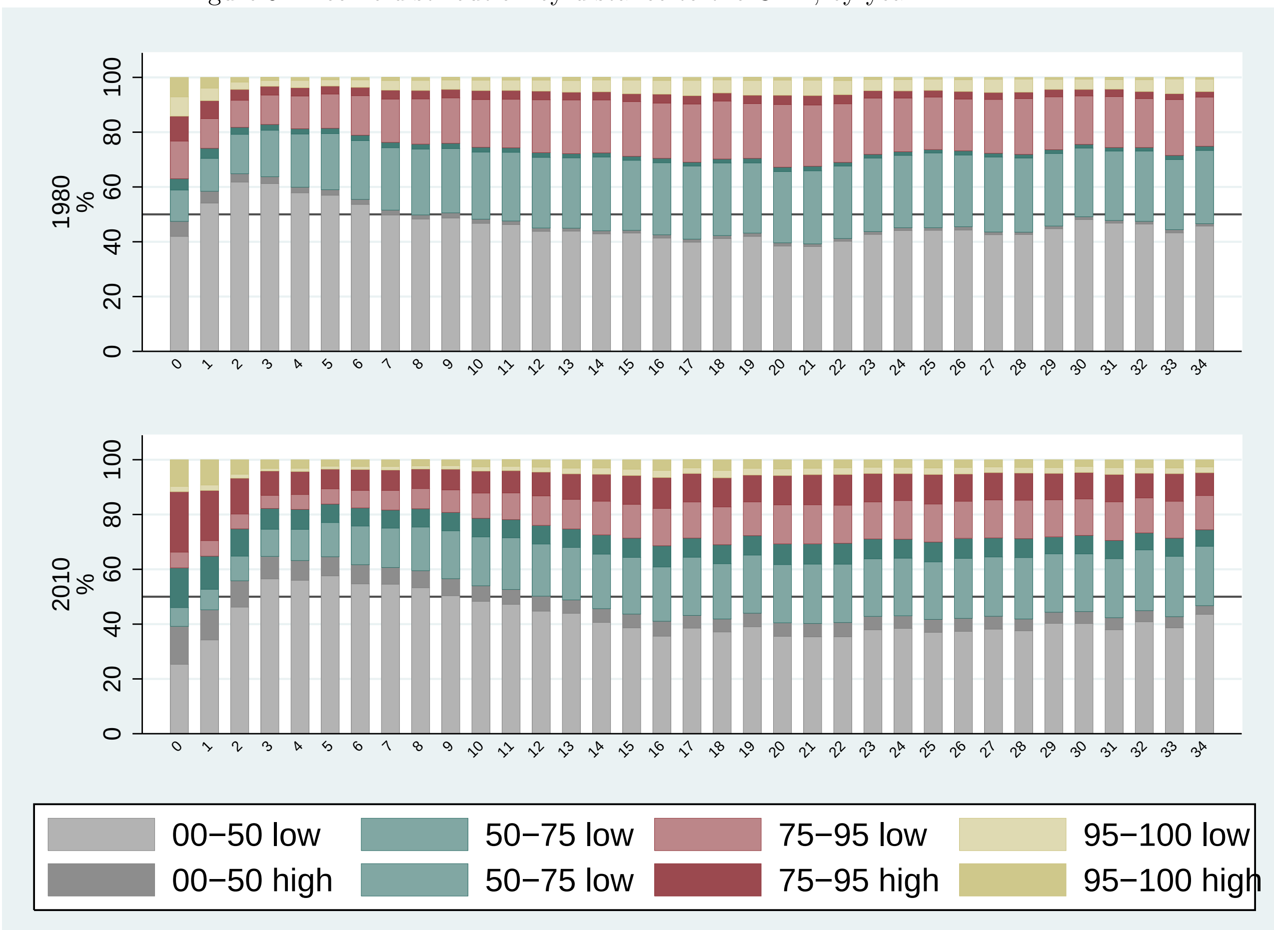
Notes: The figure shows the predicted price change using the regression results from Table 7, column 4 and observed changes in the explanatory variable $FT(h, e)$ by distance to the CDB. Values at distance 20 gives the average for the distance interval 20-35 miles.

Figure 8: Location by distance to the CBD, income ventile and year



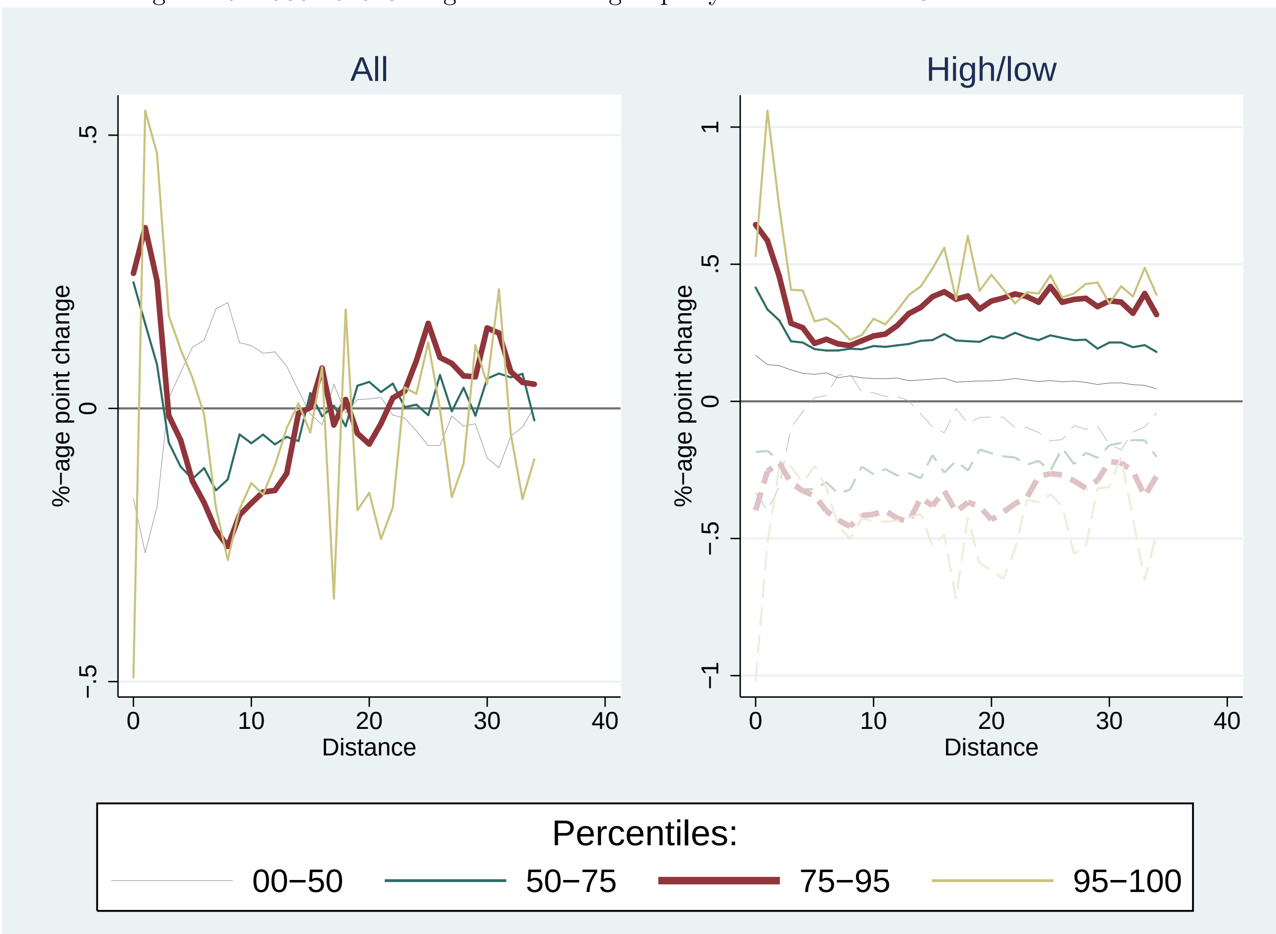
Source: 1980 decennial census and 5-year sample for 2010 American Community Survey, restricted use data. Notes: The sample is restricted to heads and spouses, 25-55 years old.
The income distribution is the individual income distribution where income for singles is total personal income. For couples, each person is assigned half of the spouses' combined total personal incomes. The income distribution is computed separately for each city and year.
Example: In 1980, 2 percent of individuals in the 75-80 ventile lived within 1 mile of the CBD, whereas almost 4 percent of the top 5 percentiles did so. The numbers for 2010 were similar with a slight decrease in the fraction of individuals in the 95-100 percentile who lived within 1 mile of the CBD.

Figure 9: Income distribution by distance to the CBD, by year



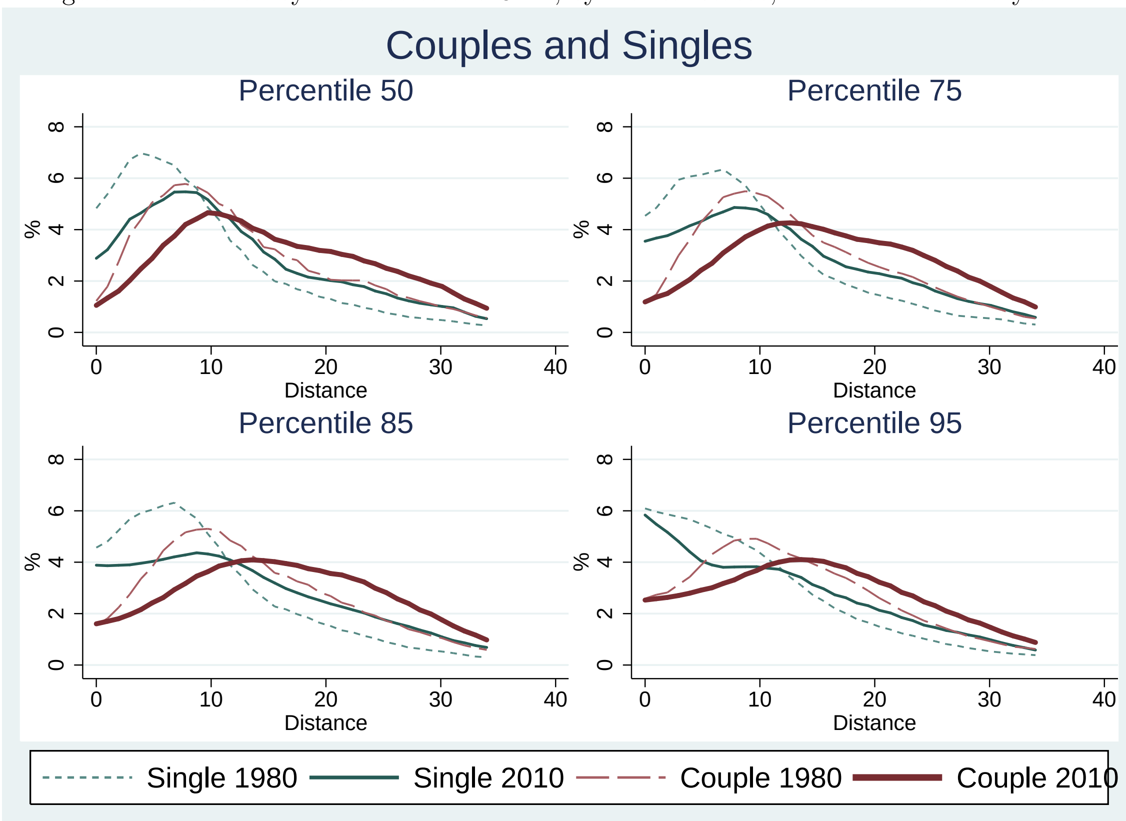
Source: 1980 decennial census and 2010 5-year sample American Community Survey, restricted use data. Notes: The sample is restricted to heads and spouses, 25-55 years old.
The income distribution is the individual income distribution where income for singles is total personal income. For couples, each person is assigned half of the spouses' combined total personal incomes. The income distribution is computed separately for each city and year.
High – works 40 hours a week or more and have a four-year college degree, if part of a couple, these requirements need to be fulfilled by both spouses.
Low – not high.
Example: The bar labeled 0 on the x-axis shows that within the first mile of the CBD center point, in 1980, 14.3 percent had incomes in the 95-100th percentile, about half of whom were of the high type. In 2010, the 95-100th percentile's fraction had fallen to 11.9 percent, and the group was overwhelmingly of the high type.

Figure 10: 1980-2010 Changes in income groups by distance to the CBD



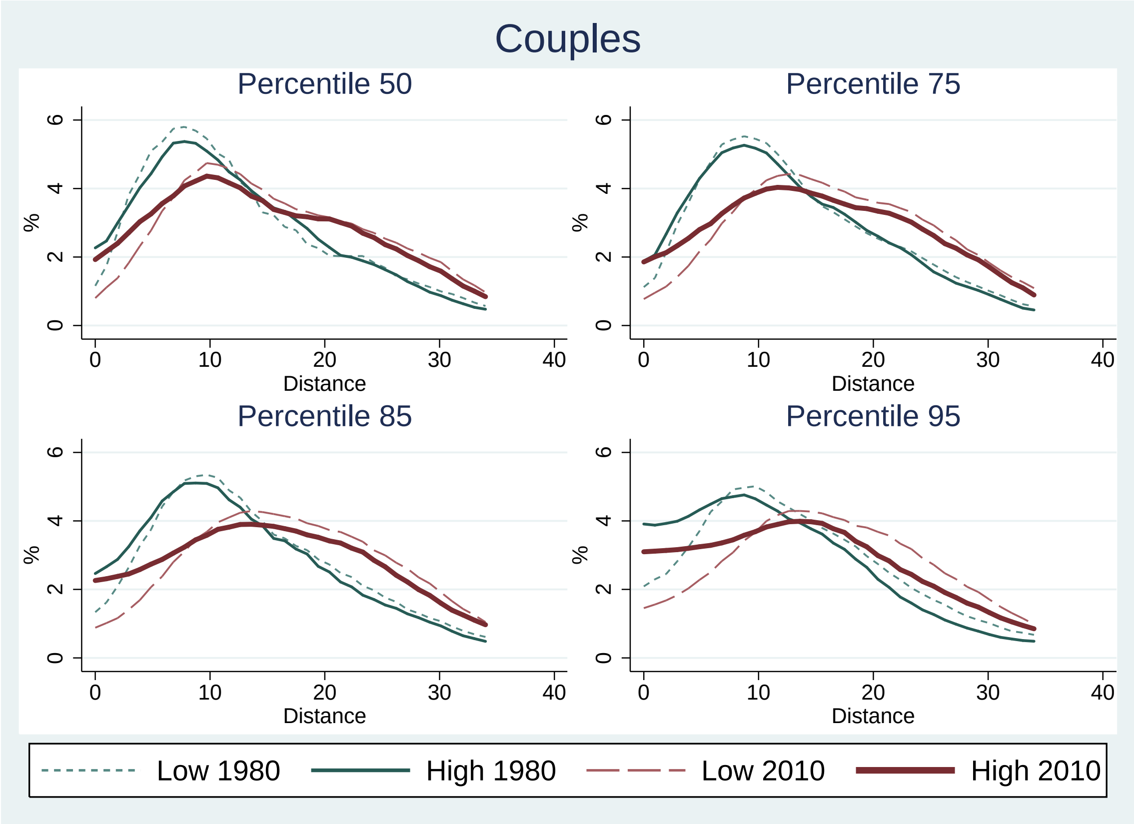
Source: 1980 decennial census and 2010 5-year sample American Community Survey, restricted use data.
Notes: The sample is restricted to heads and spouses, 25-55 years old.
The income distribution is the individual income distribution where income for singles is total personal income. For couples, each person is assigned half of the spouses' combined total personal incomes. The income distribution is computed separately for each city and year.
This graph illustrates the change in income groups' representation, the difference between the panels in Figure 9, scaled by the group size.
Left panel:
Income groups, without regard for whether high or low.
Right panel:
High – Solid line. Works 40 hours a week or more and have a four-year college degree, if part of a couple, these requirements need to be fulfilled by both spouses.
Low – Dashed line. Not high.
Example: The left panel shows the change from 14.3 to 11.9 percent in the 0-1 mile (from the CBD) for the 95-100th percentile group seen in Figure 9 as -0.5 (≈-2.4/5). The right panel shows that this decline in the top income group was driven by low-types.

Figure 11: Location by distance to the CBD, by marital status, income ventile and year



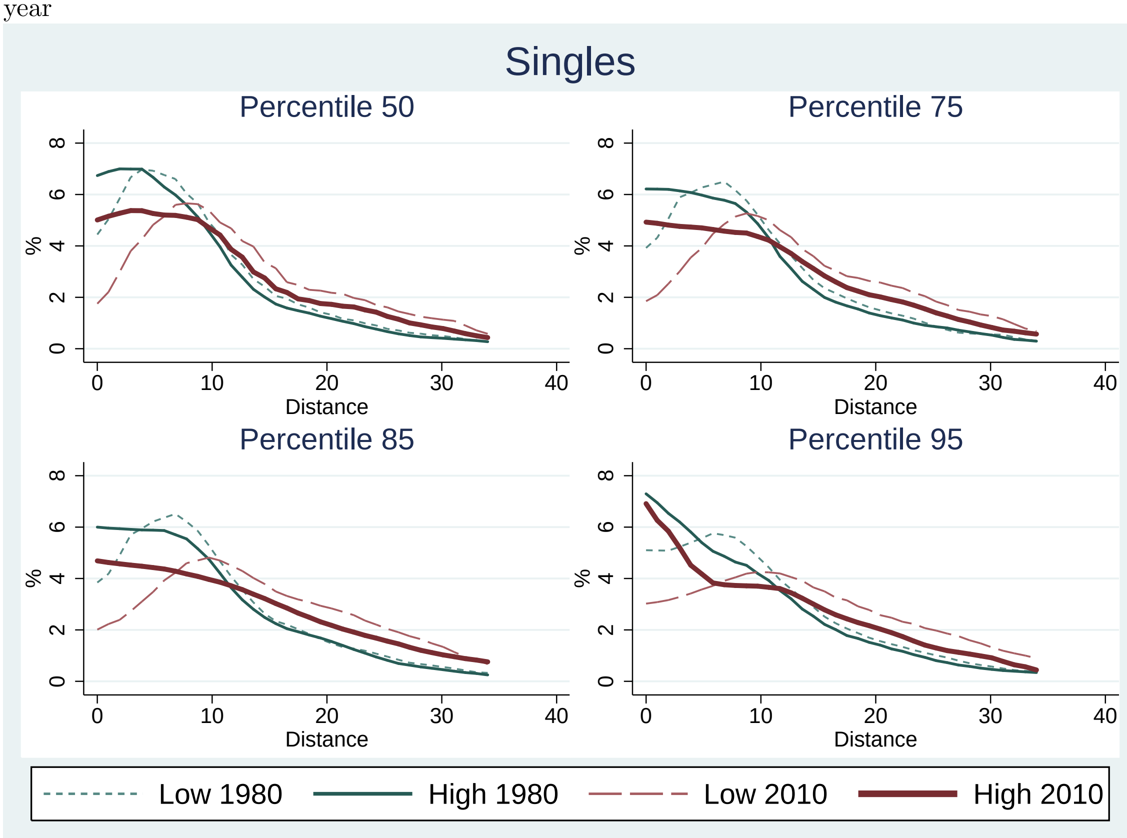
Source: 1980 decennial census and 2010 5-year sample American Community Survey, restricted use data. Notes: The sample is restricted to heads and spouses, 25-55 years old.
The income distribution is the individual income distribution where income for singles is total personal income. For couples, each person is assigned half of the spouses' combined total personal incomes. The income distribution is computed separately for each city and year. The income percentile indicates the ventile's starting percentile, i.e., percentile 50 indicates incomes in the 50-55th percentile range.
Example: The top right panel shows that among those with incomes in the 75-80th percentile, in both years, about two percent lived within one mile of the CBD center point.

Figure 12: Location by distance to the CBD, couples by high/low status, income ventile and year



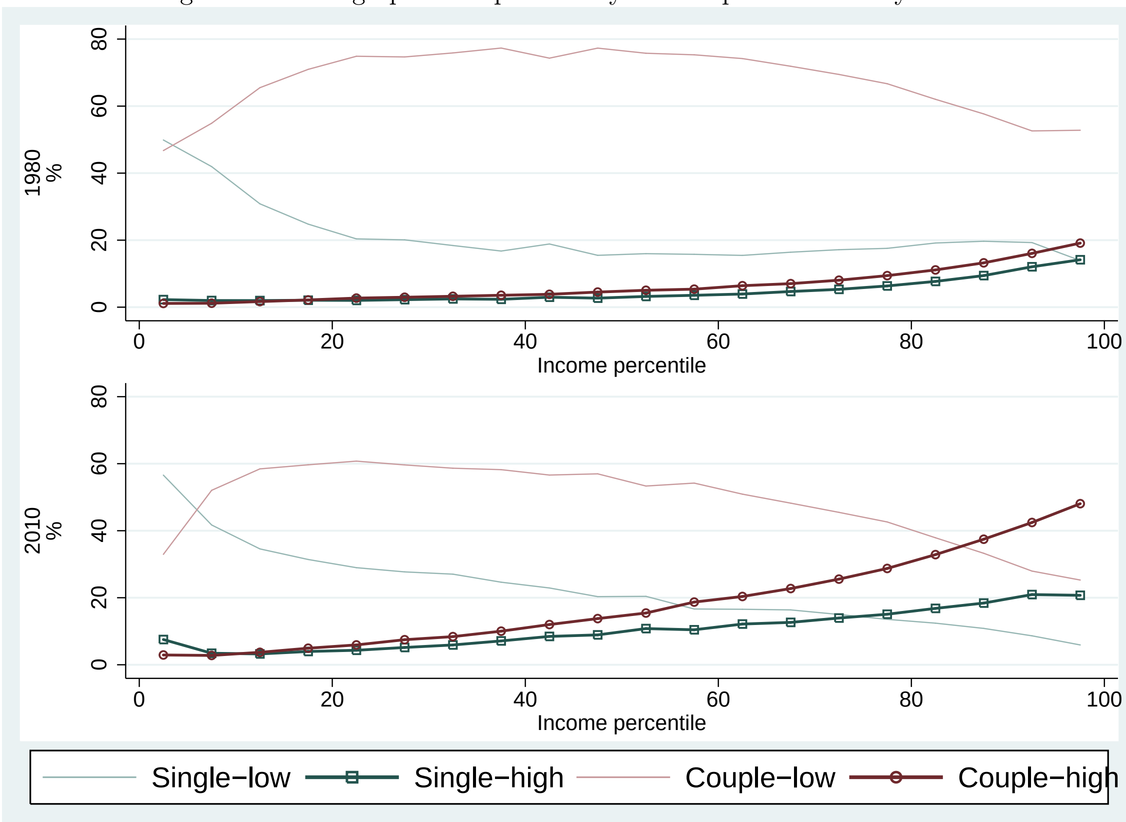
Source: 1980 decennial census and 2010 5-year sample American Community Survey, restricted use data. Notes: The sample is restricted to heads and spouses, 25-55 years old.
The income distribution is the individual income distribution where for couples, each person is assigned half of the spouses' combined total personal incomes. The income distribution is computed separately for each city and year. The income percentile indicates the ventile's starting percentile, i.e., percentile 50 indicates incomes in the 50-55th percentile range.
Couple-high – both spouses work 40 hours a week or more and have four-year college degrees.
Couple-low – couple, not high.
Example: The top right panel shows that among couples whose income fell within the 75-80th percentile, two percent of high types and about one percent of low types lived within one mile of the CBD center point in 1980 and 2010.

Figure 13: Location by distance to the CBD, singles by high/low status, income ventile and year



Source: 1980 decennial census and 2010 5-year sample American Community Survey, restricted use data. Notes: The sample is restricted to heads and spouses, 25-55 years old.
The income distribution is the individual income distribution where income is total personal incomes. The income distribution is computed separately for each city and year. The income percentile indicates the ventile's starting percentile, i.e., percentile 50 indicates incomes in the 50-55th percentile range.
Single-high – works 40 hours a week or more and has four-year college degree.
Single-low – single, not high.
Example: The top right panel shows that among couples whose income fell within the 75-80th percentile, two percent of high types and about one percent of low types lived within one mile of the CBD center point in 1980 and 2010.

Figure 14: Demographic composition by income percentile and year



Source: 1940-2010 Decennial censuses, Integrated Public Use Microdata Series (IPUMS).
Notes: The sample is restricted to heads and spouses, 25-55 years old.
The income distribution is the individual income distribution where income for singles is total personal income. For couples, each person is assigned half of the spouses' combined total personal incomes. The income distribution is computed separately for each city and year.
Single-high – works 40 hours a week or more and has four-year college degree.
Single-low – single, not high.
Couple-high – both spouses work 40 hours a week or more and have four-year college degrees.
Couple-low – couple, not high.
The demographic composition is computed for each 20-quantiles (ventiles), with markers at the ventile midpoint.
Example: In 1980, 20 percent of individuals in the 95-100 ventile were high-type couples, in 2010, that figure had risen to almost 50 percent.

Tables

Unless otherwise specified, the analysis is based on the decennial censuses and the American Community Survey, restricted use data. Throughout, sample sizes have been rounded to nearest 1000 for disclosure reasons. The repeated cross section has 65 thousand tract-years (53 thousand excluding New York); and the panel data set has 48 thousand tract-years.

Table 1: Summary Statistics

		(1)	(2)	(3)	(4)	(5)
		Distance to the CBD (miles)				
Variable	Year	$d1$=[0,3]	$d2$=(3,10]	$d3$=(10,20]	$d4$=(20,35)	[0,35)
Home price	1980	66.64	79.16	108.23	114.06	92.5
('000 1980$)	2010	154.77	115.07	119.22	120.5	120.54
	Δ	88.13	35.91	10.99	6.44	28.04
FT(40,BA+)	1980	13.78	12.92	15.19	15.05	14.05
	2010	35.13	23.15	27.22	28.29	26.62
	Δ	21.35	10.23	12.03	13.24	12.57
FT(40,MA+)	1980	6.37	5.42	6.12	5.64	5.77
	2010	15.5	8.92	10.1	10.35	10.13
	Δ	9.13	3.5	3.98	4.71	4.36
FT(50,BA+)	1980	4.03	3.43	4.3	4.44	3.92
	2010	14.32	7.25	8.81	9.69	8.86
	Δ	10.29	3.82	4.51	5.25	4.94
FT(50,MA+)	1980	2.29	1.82	2.06	1.93	1.96
	2010	7.28	3.4	3.8	4.02	3.95
	Δ	4.99	1.58	1.74	2.09	1.99
Income	1980	10.53	12.06	14.25	14.3	12.94
('000 1980$)	2010	17.98	13.99	17.22	18.77	16.53
	Δ	7.45	1.93	2.97	4.47	3.59
White	1980	48.99	62.43	75.61	85.73	68.87
	2010	44.97	38.81	53.91	66.37	51.06
	Δ	-4.02	-23.62	-21.7	-19.36	-17.81
Black	1980	32.89	23.61	13.41	6.78	18.7
	2010	25	27.22	15.43	7.49	18.25
	Δ	-7.89	3.61	2.02	0.71	-0.45

Table 1: Summary Statistics, continued

		(1)	(2)	(3)	(4)	(5)
		Distance to the CBD (miles)				
Variable	Year	$d1$=[0,3]	$d2$=(3,10]	$d3$=(10,20]	$d4$=(20,35)	[0,35)
Married	1980	46.29	62.67	72.22	77.35	66.32
	2010	36.13	48.25	60.72	66.59	56.14
	Δ	-10.16	-14.42	-11.5	-10.76	-10.18
Ages:						
0-5	1980	8.03	7.85	7.51	8.23	7.82
	2010	6.99	8.32	8.04	8.12	8.08
	Δ	-1.04	0.47	0.53	-0.11	0.26
6-12	1980	9.04	9.56	10.09	11.02	9.9
	2010	6.58	9.02	9.63	10.12	9.33
	Δ	-2.46	-0.54	-0.46	-0.9	-0.57
13-18	1980	8.97	9.66	10.62	11.35	10.16
	2010	6.05	8.12	8.64	8.83	8.33
	Δ	-2.92	-1.54	-1.98	-2.52	-1.83
19-24	1980	12.45	11.21	10.41	10.46	10.97
	2010	10.31	8.58	7.36	6.77	7.84
	Δ	-2.14	-2.63	-3.05	-3.69	-3.13
25-55	1980	39	38.7	40.52	40.81	39.63
	2010	50.45	45.53	44.64	44.63	45.34
	Δ	11.45	6.83	4.12	3.82	5.71
56-65	1980	9.74	10.86	10.67	9.21	10.42
	2010	9.67	9.88	10.7	10.88	10.39
	Δ	-0.07	-0.98	0.03	1.67	-0.03
66-	1980	12.76	12.16	10.18	8.93	11.09
	2010	9.95	10.55	10.99	10.64	10.68
	Δ	-2.81	-1.61	0.81	1.71	-0.41

Note: percent unless otherwise indicated. Income is total personal income. Source: Decennial censuses and the American Community Survey, restricted use data.

Table 2: Effects of Bartik-type demand shock on housing values

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Dependent Variable: Home price ('000) 1980$								
Z	522.0*	425.7							
	[278.0]	[264.1]							
$Z\times$									
$d1$		588.9***	580.7***	633.1***	777.4***	561.7**			
		[168.6]	[169.1]	[123.1]	[264.1]	[259.7]			
$d2$		234.5***	249.7***	285.2***	235.6				
		[52.21]	[49.95]	[54.47]	[144.8]				
$d3$		88.25***	98.22***	153.8***	319.5*				
		[18.25]	[18.49]	[41.01]	[173.0]				
$Z\times$									
$dist$							-50.30***	-47.79***	-42.65***
							[13.92]	[7.843]	[9.668]
$dist^2$							1.053***	0.928***	0.819***
							[0.307]	[0.156]	[0.150]
R^2	0.317	0.333	0.362	0.400	0.401	0.401	0.353	0.401	0.402
Fixed effects:									
City	✓	✓							
Year	✓	✓							
City-Year			✓	✓	✓	✓	✓	✓	✓
City-Distance				✓	✓	✓		✓	✓
Distance-Year					✓	✓			✓

Regressions 1-3 include distance controls $d1, d2, d3$; regressions 7-9 include linear and square continuous distance controls $(dist, dist^2)$.
Z is the Bartik demand shifter, see Equation 2, also Table A1. $d1, d2$, and $d3$ indicate distance intervals 0-3, 3-10, 10-20 miles from the CBD, respectively.
The unit of observation is a tract and each regression has 65-thousand tract-years.
Significance levels: *** – 0.01; ** – 0.05; * – 0.1. Standard errors are clustered at the city level; tracts are weighted by population size (ages 25-55).
Source: Decennial censuses and the American Community Survey, restricted use data.

Table 3: Effects of Bartik-type demand shock on housing values – robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent Variable: Home price ('000) 1980$							
	Sample:							
		Cross Section:						
		Years:		Crime decline:			Population:	
	Panel	1980/1990	2000/2010	High	Low	not NYC	Shrunk	Grew
	Panel A. City-year and City-distance FE							
$Z\times$								
$d1$	655.5***	752.3***	1164.2***	728.6***	511.2***	493.3***	596.9***	661.3***
	[169.4]	[143.5]	[179.5]	[187.8]	[92.62]	[62.16]	[88.75]	[190.4]
$d2$	299.7***	449.8***	531.4**	285.2***	290.3**	266.7***	352.6***	241.0***
	[60.29]	[112.4]	[228.5]	[43.39]	[114.6]	[71.79]	[100.7]	[66.12]
$d3$	125.2*	298.7***	175.8**	99.77**	241.4***	150.1**	239.9**	102.8**
	[61.03]	[107.0]	[79.86]	[37.41]	[70.42]	[57.38]	[77.41]	[41.01]
R^2	0.591	0.188	0.613	0.356	0.346	0.408	0.241	0.436
	Panel B. City-year, City-distance, and Distance-year FE							
$Z\times$								
$d1$	856.2**	533.2	2464.4***	998.3	565.7***	591.8***	622.2***	1190.0**
	[375.1]	[454.2]	[635.1]	[618.4]	[108.0]	[111.3]	[148.9]	[514.0]
$d2$	271	-146.3	1489.6	359.5	244.8**	225.1*	264.2	552.3**
	[194.1]	[444.8]	[1000.5]	[231.5]	[112.0]	[125.3]	[149.0]	[223.9]
$d3$	431.6*	456.6	280.2	473.6***	398.3***	325.8*	474.6***	354.8**
	[251]	[552.0]	[180.3]	[120.5]	[117.2]	[170.1]	[146.7]	[151.3]
R^2	0.562	0.188	0.613	0.356	0.348	0.409	0.244	0.436

Column 1 includes tract fixed effects.
Z is the Bartik demand shifter, see Equation 2, also Table A1. $d1, d2$, and $d3$ indicate distance 0-3, 3-10, 10-20 miles from the CBD, respectively.
The unit of observation is a tract and sample sizes are as follows: column 1: 48 thousand; columns 2-5, 7-8: about 65/2 thousand; column 6: 53 thousand tract-years.
Significance levels: *** – 0.01; ** – 0.05; * – 0.1. Standard errors are clustered at the city level; tracts are weighted by population size (ages 25-55). Source: Decennial censuses and the American Community Survey, restricted use data.

Table 4: Effects of Bartik-type demand shock on full-time skilled workers

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable:				
	A. $FT(40, BA+)$				
Z	11.56				
	[13.91]				
$Z\times$					
$d1$		97.27***	33.10**	120.6*	129.4***
		[12.58]	[13.43]	[59.46]	[39.51]
$d2$		16.94	-31.30***	24.02	
		[14.58]	[8.369]	[44.49]	
$d3$		6.977	-16.11*	-54.85*	
		[16.93]	[8.549]	[29.14]	
R^2	0.185	0.203	0.318	0.320	0.320
	B. $FT(40, MA+)$				
Z	11.73**				
	[4.520]				
$Z\times$					
$d1$		52.17***	14.22***	52.46**	57.52***
		[6.621]	[4.024]	[21.24]	[16.07]
$d2$		12.34	-15.66***	6.307	
		[10.97]	[3.892]	[17.57]	
$d3$		8.080	-7.354	-22.21	
		[10.78]	[4.823]	[15.33]	
R^2	0.167	0.185	0.283	0.285	0.284
Fixed effects:					
city	✓				
year	✓				
city-year		✓	✓	✓	✓
city-distance			✓	✓	✓
distance-year				✓	✓

Regressions 1-2 include distance controls $d1, d2, d3$.
Z is the Bartik demand shifter, see Equation 2, also Table A1. $d1, d2$, and $d3$ indicate distance 0-3, 3-10, 10-20 miles from the CBD, respectively.
The unit of observation is a tract and each regression has 65-thousand tract-years.
Significance levels: *** – 0.01; ** – 0.05; * – 0.1. Standard errors are clustered at the city level; tracts are weighted by population size (ages 25-55).
Source: Decennial censuses and the American Community Survey, restricted use data.

Table 5: Effects of Bartik-type demand shock on other demographics

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable:					
	%					('000) 1980$
	$FT(40, BA-)$	$BA+$	Black	White	Married	Income
$Z\times$						
$d1$	-0.62	85.72	-92.67	268.2***	103.9***	63.70*
	[0.376]	[55.53]	[67.05]	[49.59]	[24.01]	[34.16]
$d2$	0.029	14.69	-52.56	141.7**	70.97**	9.523
	[0.314]	[59.31]	[52.06]	[58.50]	[27.07]	[14.24]
$d3$	0.516*	-76.31*	-17.5	63.12	51.39**	-5.304
	[0.284]	[39.51]	[51.28]	[58.72]	[18.62]	[13.84]
R^2	0.261	0.291	0.279	0.444	0.477	0.295

All regressions include city-year, city-distance and distance-year fixed effects.
Z is the Bartik demand shifter, see Equation 2, also Table A1. $d1, d2$, and $d3$ indicate distance intervals 0-3, 3-10, 10-20 miles from the CBD, respectively.
The unit of observation is a tract and each regression has 65-thousand tract-years.
Significance levels: *** – 0.01; ** – 0.05; * – 0.1. Standard errors are clustered at the city level; tracts are weighted by population size (ages 25-55). Source: Decennial censuses and the American Community Survey, restricted use data.

Table 6: The relationship between full-time skilled workers and housing prices

	(1)	(2)	(3)	(4)
	Dependent Variable: Home price ('000) 1980$			
	A. $FT(40, BA+)$			
FT	1.829***	1.831***	2.037***	1.838***
	[0.167]	[0.167]	[0.164]	[0.147]
$FT \times dist$			-0.0170**	-0.000484
			[0.00686]	[0.00501]
R^2	0.475	0.482	0.475	0.482
	B. $FT(40, MA+)$			
FT	3.502***	3.507***	3.635***	3.334***
	[0.340]	[0.342]	[0.335]	[0.309]
$FT \times dist$			-0.0115	0.0144
			[0.0163]	[0.0130]
R^2	0.475	0.482	0.475	0.482
Fixed effects:				
City-Year	✓	✓	✓	✓
City-Distance	✓	✓	✓	✓
Distance-Year		✓		✓

Regressions 3-4 include continuous distance control, *dist*.
The unit of observation is a tract and each regression has 65-thousand tract-years.
Significance levels: *** – 0.01; ** – 0.05; * – 0.1. Standard errors are clustered at the city level; tracts are weighted by population size (ages 25-55). Source: Decennial censuses and the American Community Survey, restricted use data.

Table 7: Effects of full-time skilled workers on housing values – IV results

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: Home price ('000) 1980$				
	A. $FT(40, BA+)$				
FT	2.007*	1.71	4.341***	5.026***	3.429**
	[1.201]	[1.363]	[1.215]	[1.049]	[1.508]
$FT \times dist$				-0.194***	-0.401**
				[0.0185]	[0.166]
K-P LM test (p)	0.00615	0.0353	0.00704	0.0024	0.03
C-D $Wald$ stat.	$77.35^{(b5,s10)}$	$32.45^{(b5,s10)}$	$53.9^{(b10)}$	$76.86^{(b5,s10)}$	$12.49^{(b5,s15)}$
K-P $Wald$ stat.	$20^{(b5,s10)}$	$6.007^{(b30)}$	$10.82^{(b10)}$	$18.99^{(b5,s10)}$	3.023
Overid. test (p)	0.0456	0.105	.	0.228	0.29
	B. $FT(40, MA+)$				
FT	3.212	4.595	9.765***	11.11***	5.943
	[2.605]	[3.232]	[2.759]	[2.408]	[4.049]
$FT \times dist$				-0.499***	-0.666**
				[0.0541]	[0.320]
K-P LM test (p)	0.00438	0.0233	0.0053	0.0018	0.0589
C-D $Wald$ stat.	$63.02^{(b5,s10)}$	$20.08^{(b5,s10)}$	$38.88^{(b10)}$	$62.22^{(b5,s10)}$	$9.431^{(b5,s15)}$
K-P $Wald$ stat.	$30.01^{(b5,s10)}$	4.68	$12.92^{(b10)}$	$20.62^{(b5,s10)}$	1.54
Overid. test (p)	0.0589	0.149	.	0.195	0.219
Fixed effects:					
City-Year	✓	✓	✓	✓	✓
City-Distance	✓	✓	✓	✓	✓
Distance-Year		✓	✓		✓
Instruments:					
$Z \times (d1, d2, d3)$	✓	✓		✓	✓
$Z \times d1$			✓		

Notes on separate page.

Notes to Table 7.

All specifications include city-year, city-distance and distance-year fixed effects.

Regressions 4-5 include continuous distance control, $dist$.

Z is the Bartik demand shifter, see Equation 2, also Table A1. $dist$ is distance in miles from the CBD, $d1, d2$, and $d3$ indicate distance intervals 0-3, 3-10, 10-20 miles from the CBD, respectively.

The unit of observation is a tract and each regression has 65-thousand tract-years.

Significance levels: *** – 0.01; ** – 0.05; * – 0.1. Standard errors are clustered at the city level; tracts are weighted by population size (ages 25-55). Source: Decennial censuses and the American Community Survey, restricted use data.

K-P LM test (p) corresponds to the p-value of the Kleinbergen-Paap LM test. The null hypothesis is that the structural equation is under identified (i.e., the rank condition fails).

C-D $Wald$ stat and K-P $Wald$ statistics are the Cragg-Donald and Kleibergen-Paap Wald statistics for testing weak identification. In both cases, the critical values are the Stock and Yogo [2005] critical values initially tabulated for the C-D $Wald$ stat, see Table 8. We follow Baum [2007] and also apply the Stock and Yogo critical values to the K-P $Wald$ stat (critical values for the K-P $Wald$ stat do not exist). In each case, we specify whether the test statistics rejects the null hypothesis (at the 5% level) that the bias of of the IV estimates exceeds the OLS bias by 5, 10, 20 and 30 percent (b5, b10, b20, b30), and whether test statistics rejects the null hypothesis (at the 5% level) that the size of the test is greater than 10, 15, 20 and 25 percent (s20, s15, s20, s25). C-D $Wald$ stat and K-P $Wald$ statistics reduce to the the standard non-robust and heteroscedasticity robust first-stage F-statistics, respectively. The critical values for the relative bias test cannot be computed for the case of 2 endogenous and 3 instruments and we use the more conservative critical values for the case of 2 endogenous and 4 instruments, see Stock and Yogo [2005] and Table 8.

Overid. test (p) corresponds to the p-value of the test of the overidentifying restrictions. The null hypothesis is that the instruments are valid instruments.

Source: Decennial censuses and the American Community Survey, restricted use data.

Table 8: Threshold values for Cragg-Donald and Kleibergen-Paap $Wald$ statistics

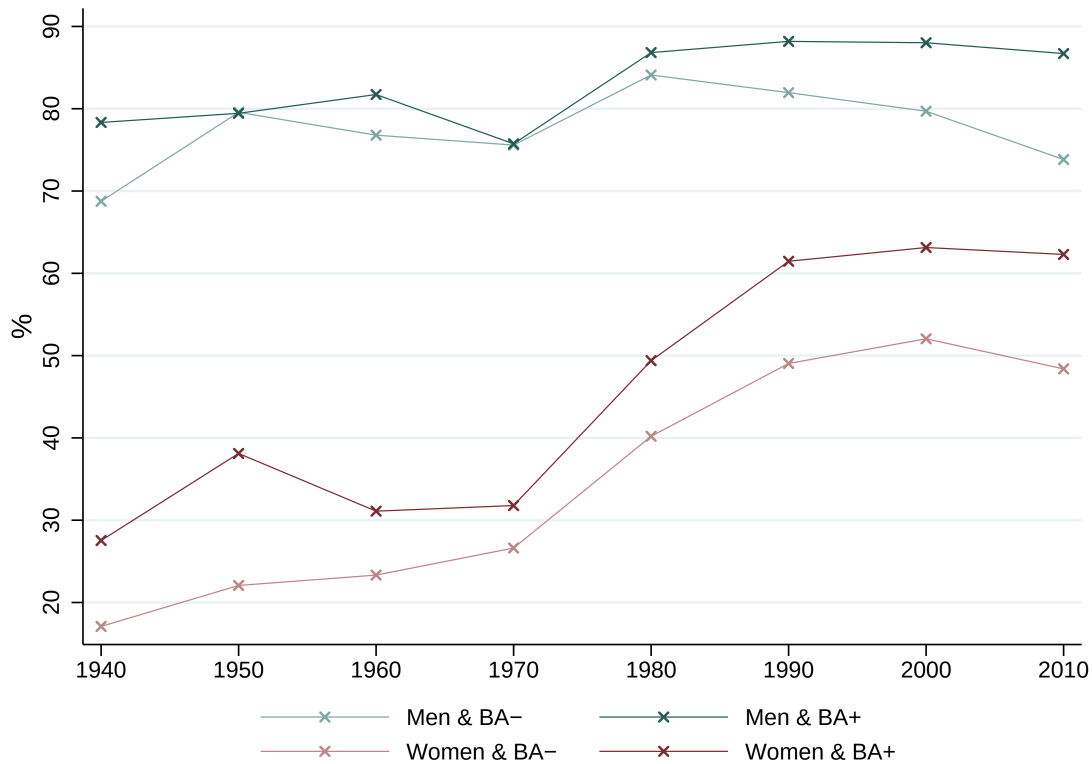
	(1)	(2)
Number of variables:		
Endogenous	1	2
Exogenous	3	4^a
Stock-Yogo 2SLS Bias		
0.05	13.91	11.04
0.10	9.08	7.56
0.20	6.46	5.57
0.30	5.39	4.73
Stock-Yogo 2SLS Size		
0.10	22.30	13.43
0.15	12.83	8.18
0.20	9.54	6.40
0.25	7.80	5.45

^{a} – used for our case of 2 endogenous and 3 exogenous (more conservative).
Source: Stock and Yogo [2005, table 5.1 and 5.2].

Appendix

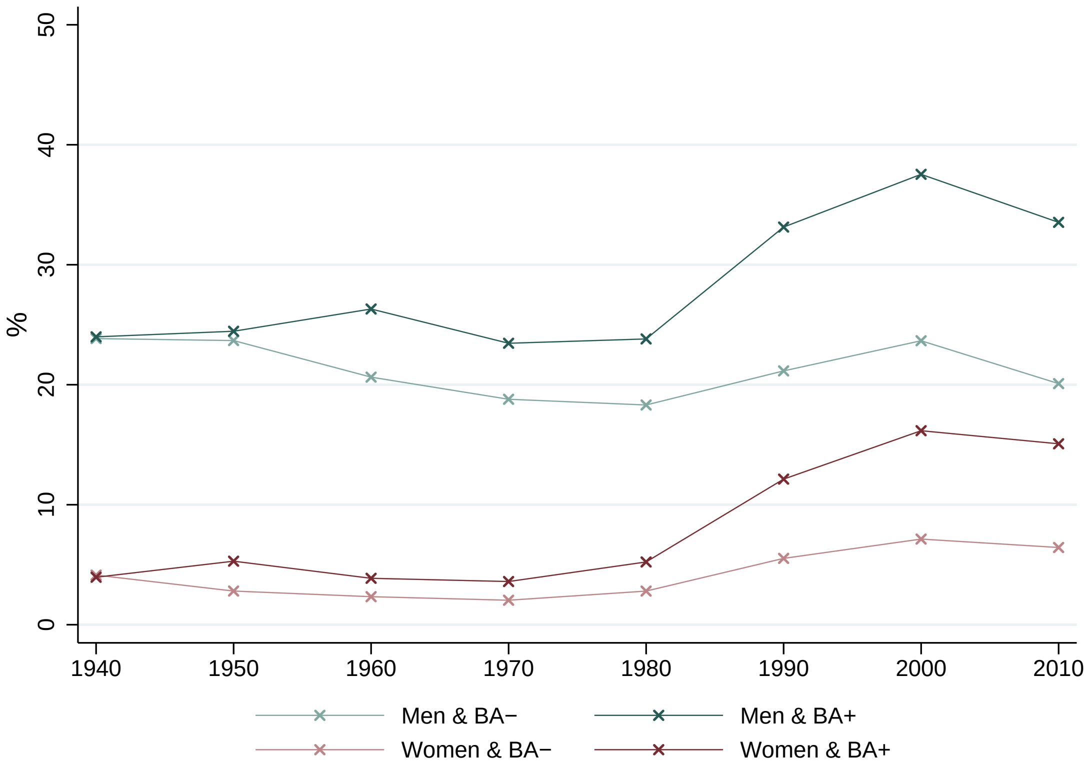
Figures

Figure A1: % Full time, 40+h/week, 1940-2010



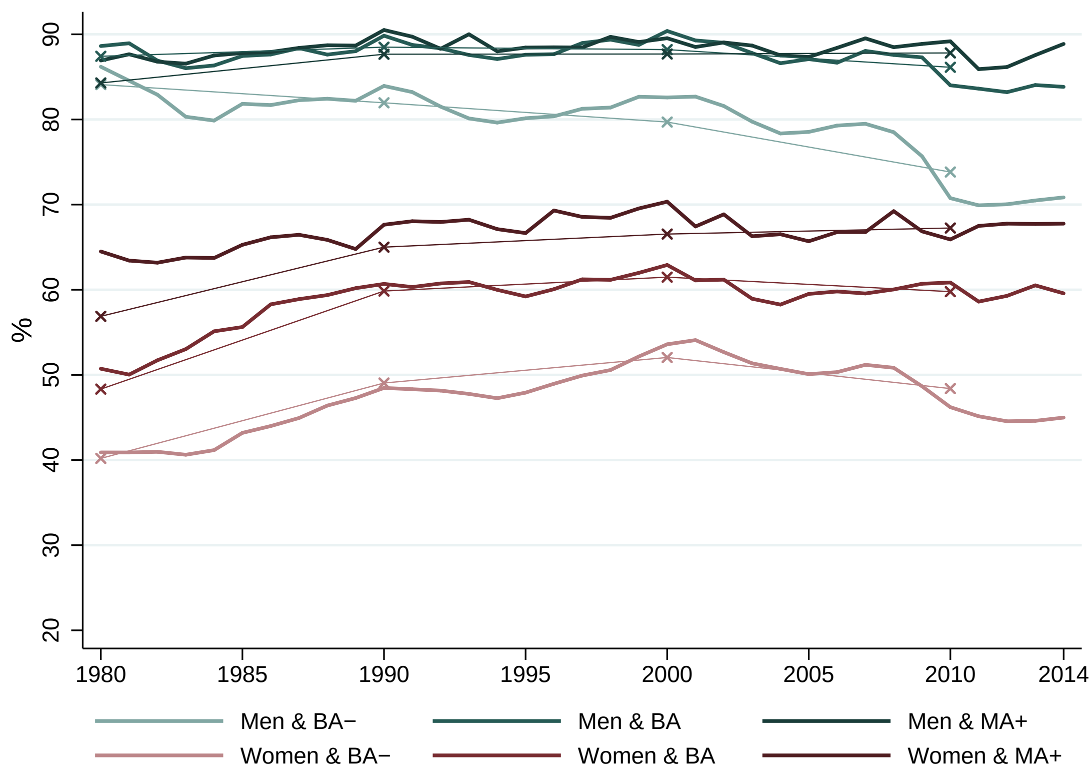
Source: 1940-2010 decennial censuses, Integrated Public Use Microdata Series (IPUMS).
Notes: The figure shows the percent of individuals (25-55) that worked more than 40 hours/week by census year, by sex and education. BA+ corresponds to a four-year college degree or more, BA- is the complement. See the Appendix for further variable and sample construction details.

Figure A2: % Full time, 50+h/week, 1940-2010



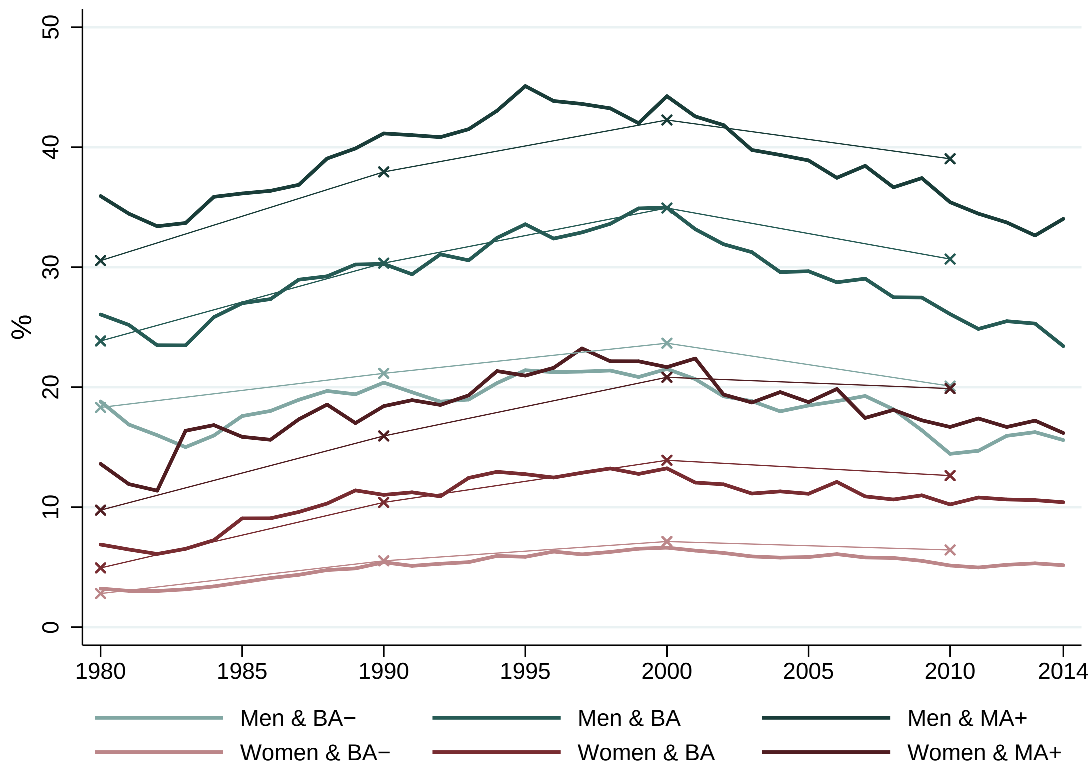
Source: 1940-2010 decennial censuses, Integrated Public Use Microdata Series (IPUMS).
Notes: The figure shows the percent of individuals (25-55) that worked more than 50 hours/week by census year, by sex and education. BA+ corresponds to a four-year college degree or more, BA- is the complement. See the Appendix for further variable and sample construction details.

Figure A3: Full time, 40+h/week, 1980-2014



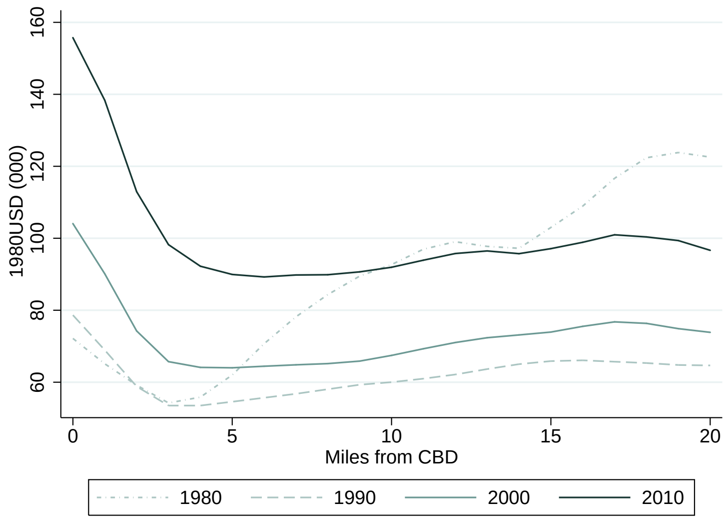
Note: Men and women ages 25-55.
Sources: Decennial data: Decennial censuses, integrated public use micro data series (IPUMS).
Annual data: The Current Population Surveys (CPS).

Figure A4: Full time, 50+h/week, 1980-2014



Note: Men and women ages 25-55.
Sources: Decennial data: Decennial censuses, integrated public use micro data series (IPUMS).
Annual data: The Current Population Surveys (CPS).

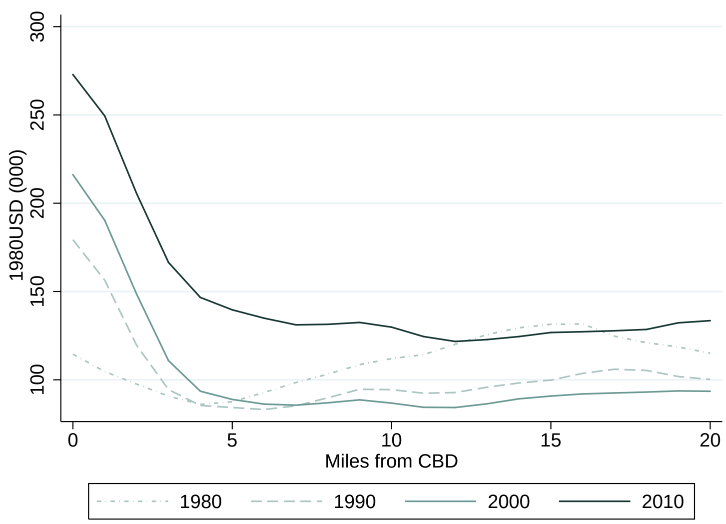
Figure A5: Home prices by distance from the CBD, cities that lost population



Notes: The figure shows the median price (1980$) for owner-occupied, 2-3 bedroom, one-family homes in our top-20 US cities, by distance from the CBD. See the Appendix for further details on variable and sample construction, especially Table A8. 20 miles includes 20-35 miles.

Source: Decennial censuses and the American Community Survey, restricted use data.

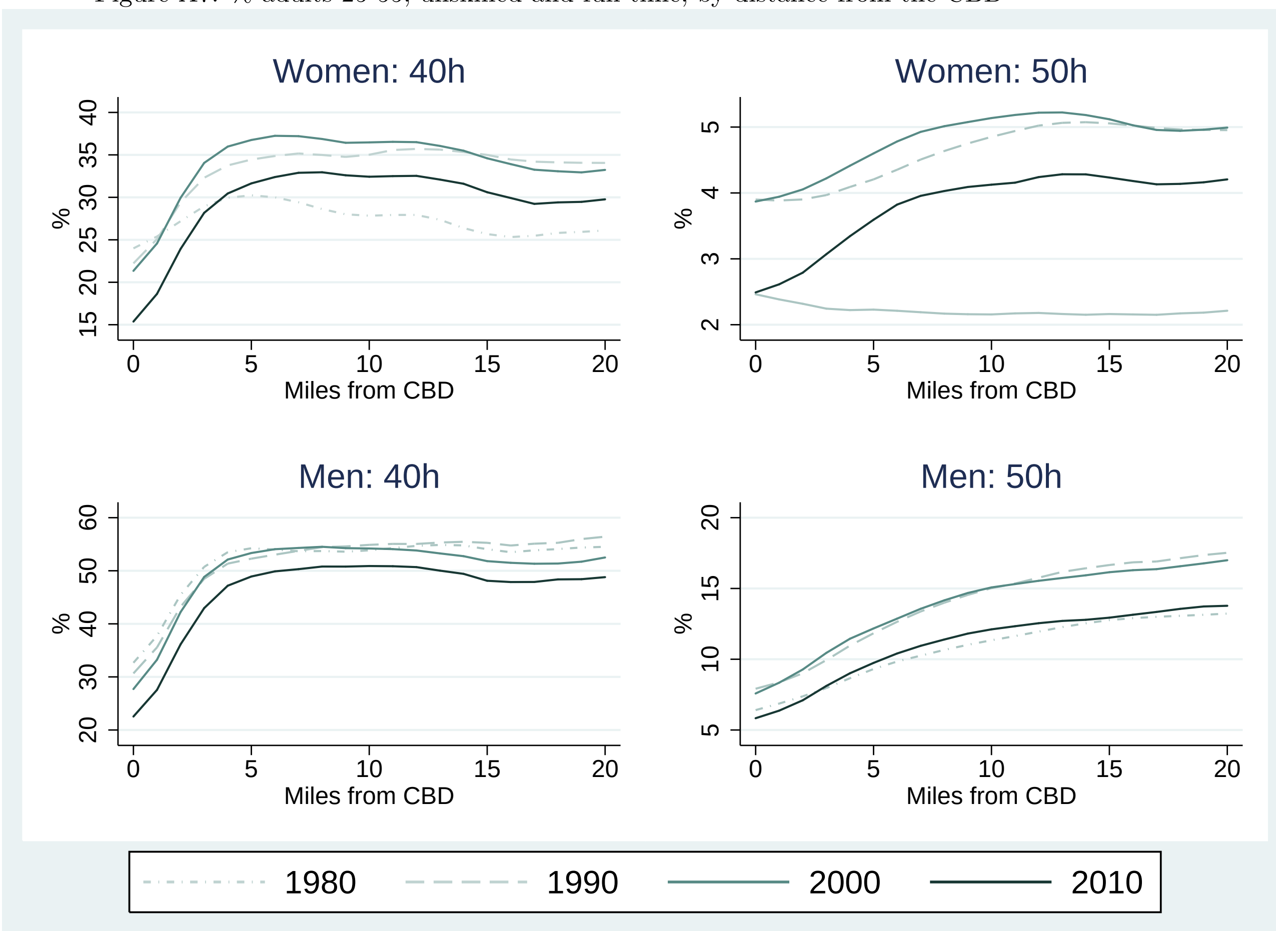
Figure A6: Home prices by distance from the CBD, cities that gained population



Notes: The figure shows the median price (1980$) for owner-occupied, 2-3 bedroom, one-family homes in our top-20 US cities, by distance from the CBD. See the Appendix for further details on variable and sample construction, especially Table A8. 20 miles includes 20-35 miles.

Source: Decennial censuses and the American Community Survey, restricted use data.

Figure A7: % adults 25-55, unskilled and full time, by distance from the CBD



Notes: 40h(50h) denotes working 40(50) hours or more per week.
Source: Decennial censuses and the American Community Survey, restricted use data.

Bartik

We group industries into 41 categories using the 1990 industry classification available in IPUMS (see IPUMS website or Edlund et al. [2015]. In the preliminary analysis we also used a more aggregate grouping into seven industries, the results were very similar, but we favor the more disaggregate grouping because of the reduction in within-group industry heterogeneity.

Table A1: Bartik, Base Year 1970

City	1980	1990	2000	2010	Central Business District (CBD)
Austin	0.169	0.259	0.320	0.350	Texas Capitol
Baltimore	0.110	0.169	0.206	0.223	W. Lexington × Park Ave.
Boston	0.144	0.226	0.277	0.298	South Station
Charlotte	0.102	0.162	0.192	0.206	Charlotte Convention Center
Chicago	0.110	0.175	0.214	0.229	LaSalle × W. Congress Parkway
Cleveland	0.102	0.162	0.196	0.212	Soldiers' and Sailors' Monument
Columbus	0.126	0.197	0.242	0.262	E. Long Street × Route 3
Dallas	0.122	0.196	0.232	0.251	Dallas Convention center
Detroit	0.090	0.145	0.171	0.186	Grand Circus Park
El Paso	0.112	0.170	0.211	0.230	El Paso Art Institute
Fort Worth	0.122	0.196	0.232	0.251	Fort Worth Convention Center
Houston	0.137	0.209	0.243	0.266	Houston Center
Indianapolis	0.093	0.147	0.177	0.191	Monument Circle
Jacksonville	0.094	0.148	0.178	0.194	Bank of America Tower
Los Angeles	0.126	0.200	0.242	0.260	Pershing Square
Memphis	0.091	0.144	0.176	0.190	Beale St. × Riverside Drive
Milwaukee	0.102	0.159	0.192	0.207	Milwaukee County Court House
New Orleans	0.112	0.169	0.196	0.216	New Orleans Morial Convention Center
New York	0.136	0.217	0.266	0.286	Rockefeller Center
Philadelphia	0.105	0.165	0.202	0.218	City Hall
Phoenix	0.116	0.183	0.221	0.237	Phoenix Convention Center
Saint Louis	0.103	0.163	0.197	0.213	Federal Reserve
San Antonio	0.089	0.138	0.166	0.180	Tower of the Americas
San Diego	0.125	0.196	0.240	0.259	Horton Plaza
San Francisco	0.172	0.273	0.332	0.362	Transamerica Pyramid
San Jose	0.163	0.253	0.312	0.330	1 Infinity Loop (Apple Inc. headquarters)
Washington D.C.	0.229	0.344	0.426	0.461	White House

Tables

Table A2: The Effects of Demand Shock on Housing Values ('000) 1980$, logged

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Dependent Variable: Home price ('000) 1980$, Logged					
Z	1.995	1.212							
	[1.585]	[1.622]							
$Z\times$									
$d1$		4.988***	5.031***	3.618***	-1.148	-2.363**			
		[1.157]	[1.172]	[0.581]	[1.438]	[1.009]			
$d2$		2.442***	2.588***	1.798***	1.386				
		[0.605]	[0.602]	[0.529]	[1.501]				
$d3$		0.626***	0.682***	0.720***	1.725*				
		[0.188]	[0.182]	[0.259]	[0.860]				
$Z\times$									
$dist$							-0.532***	-0.335***	-0.246***
							[0.122]	[0.0506]	[0.0833]
$dist^2$							0.0116***	0.00703***	0.00526***
							[0.00270]	[0.00119]	[0.00144]
R^2	0.479	0.498	0.528	0.591	0.592	0.592	0.528	0.592	0.593
Fixed effects:									
City	✓	✓							
Year	✓	✓							
City-Year			✓	✓	✓	✓	✓	✓	✓
City-Distance				✓	✓	✓		✓	✓
Distance-Year					✓	✓			✓

Notes, see Table 2.

Table A3: Effects of Bartik-type demand shock on full-time skilled workers

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: $FT(50, BA+)$				
Z	11.73				
	[8.483]				
$Z\times$					
$d1$		54.97***	35.51**	103.3**	95.50**
		[11.44]	[14.73]	[44.69]	[42.18]
$d2$		8.181	-14.00***	25.65*	
		[7.638]	[3.863]	[14.84]	
$d3$		2.720	-8.558**	-9.807	
		[7.520]	[4.119]	[13.34]	
R^2	0.145	0.167	0.294	0.296	0.296
	$FT(50, MA+)$				
Z	8.160***				
	[1.981]				
$d1$		29.22***	16.79***	43.35**	40.07**
		[4.266]	[5.363]	[17.56]	[17.98]
$d2$		6.175	-6.297***	10.76	
		[5.229]	[2.014]	[7.386]	
$d3$		3.229	-3.270	-4.061	
		[4.468]	[2.224]	[6.950]	
R^2	0.128	0.147	0.253	0.255	0.254
Fixed effects:					
city	✓				
year	✓				
city-year		✓	✓	✓	✓
city-distance			✓	✓	✓
distance-year				✓	✓

Notes, see Table 4.

Table A4: The relationship between full-time skilled workers and housing prices

	(1)	(2)	(3)	(4)
		$FT(50, BA+)$		
FT	4.003***	3.995***	4.320***	3.964***
	[0.371]	[0.371]	[0.325]	[0.322]
$FT \times dist$			-0.0260**	0.00227
			[0.0124]	[0.0118]
R^2	0.488	0.494	0.488	0.494
		$FT(50, MA+)$		
FT	6.732***	6.708***	6.740***	6.210***
	[0.651]	[0.654]	[0.625]	[0.587]
$FT \times dist$			-0.00124	0.0422
			[0.0320]	[0.0283]
R^2	0.480	0.486	0.480	0.486
Fixed effects:				
City-Year	✓	✓	✓	✓
City-Distance	✓	✓	✓	✓
Distance-Year		✓		✓

Notes, see Table 6.

Table A5: Effects of full-time skilled workers on housing values – IV results

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: Home price ('000) 1980$				
	$FT(50, BA+)$				
FT	6.017***	4.328***	5.882***	7.067***	6.887***
	[0.889]	[1.325]	[1.047]	[0.801]	[2.613]
$FT \times dist$				-0.401***	-1.483
				[0.0430]	[1.136]
K-P LM test (p)	0.0163	0.0174	0.00521	0.00677	0.177
C-D $Wald$ stat.	169.4	53	121.8	147.9	2.511
K-P $Wald$ stat.	7.615	4.755	5.17	8.768	1.982
Overid. test (p)	0.0299	0.405	.	0.328	0.91
	$FT(50, MA+)$				
FT	13.15***	10.34***	14.02***	15.28***	10.37**
	[1.973]	[3.234]	[2.421]	[1.787]	[5.090]
$FT \times dist$				-0.915***	-2.401
				[0.107]	[2.362]
K-P LM test (p)	0.0164	0.0142	0.00437	0.0025	0.199
C-D $Wald$ stat.	111.2	28.59	65.86	96.59	2.211
K-P $Wald$ stat.	10.05	4.506	5.008	11.44	0.907
Overid. test (p)	0.0406	0.378	.	0.336	0.463
Fixed effects:					
City-Year	✓	✓	✓	✓	✓
City-Distance	✓	✓	✓	✓	✓
Distance-Year		✓	✓		✓
Instruments:					
$Z \times (d1, d2, d3)$	✓	✓		✓	✓
$Z \times d1$			✓		

Notes, see Table 7.

Table A6: Effects of full-time skilled workers on housing values – IV results, robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent Variable: Home price ('000) 1980$							
	Sample:							
		Cross Section:						
		Years:		Crime decline:			Population:	
	Panel	1980/1990	2000/2010	High	Low	not NYC	Shrunk	Grew
	$FT(40, BA+)$							
FT	4.132***	5.267	6.339**	3.767***	3.477*	3.402***	4.266***	4.473***
	[1.407]	[3.768]	[2.599]	[0.793]	[1.943]	[0.671]	[1.379]	[1.647]
K-P LM test (p)	0.0037	0.0444	0.146	0.128	0.0337	0.0496	0.172	0.017
C-D $Wald$ stat.	436.7	6.407	6.727	13.13	40.27	36.78	11.37	43.02
K-P LM test (p)	10.23	12.68	9.509	22.64	22.84	13.03	23.63	31.48
	$FT(40, MA+)$							
FT	9.750***	8.132	8.083***	7.075***	8.704*	7.570***	8.175***	11.39***
	[2.687]	[5.489]	[2.142]	[2.258]	[4.796]	[1.612]	[2.787]	[4.037]
K-P LM test (p)	0.0028	0.0692	0.0943	0.092	0.0318	0.0349	0.183	0.0112
C-D $Wald$ stat.	240.4	10.27	14.87	13.88	22.91	27.74	10.03	26.71
K-P LM test (p)	8.121	6.568	27.51	11.08	16.78	18.36	27.4	15.52
	$FT(50, BA+)$							
FT	5.306***	5.203*	15.67**	7.124***	4.487***	5.326***	7.154***	5.756***
	[1.182]	[2.977]	[7.663]	[1.345]	[1.572]	[1.197]	[1.622]	[1.145]
K-P LM test (p)	0.00436	0.0212	0.132	0.141	0.0193	0.0542	0.114	0.00663
C-D $Wald$ stat.	720.1	35.13	4.338	16.2	94.37	65.25	17.15	106.3
K-P LM test (p)	5.354	8.844	6.511	22.75	6.095	8.368	13.83	9.061
	$FT(50, MA+)$							
FT	12.69***	8.600*	19.72***	15.77***	10.92***	12.72***	15.65***	14.55***
	[2.379]	[4.846]	[5.624]	[4.675]	[3.563]	[2.897]	[3.431]	[2.749]
K-P LM test (p)	0.00548	0.0278	0.0862	0.0993	0.0183	0.0399	0.102	0.00676
C-D $Wald$ stat.	347.8	37.95	8.411	10.14	48.88	35.46	9.747	56.16
K-P LM test (p)	4.376	7.545	20.72	11.22	5.489	8.8	12.24	7.126

The instrument is $Z \times d1$, corresponding to column 3, Table 7.
For cities in columns 4 and 5, see Appendix Table A7.
For cities in columns 7 and 8, see Appendix Table A8.
All specifications include city-year, city-distance and distance-year fixed effects.
Z is the Bartik demand shifter, see Equation 2, also Table A1. $d1, d2$, and $d3$ indicate distance 0-3, 3-10, 10-20 miles from the CBD, respectively.
The unit of observation is a tract and sample sizes are as follows: column 1: 48 thousand; columns 2-5, 7-8: about 65/2 thousand; column 6: 53 thousand tract-years.
Significance levels: *** – 0.01; ** – 0.05; * – 0.1. Standard errors are clustered at the city level; tracts are weighted by population size (ages 25-55). Source: Decennial censuses and the American Community Survey, restricted use data.
See Table 7 for further notes.

Table A7: Change in Violent Crime 1986-2012

High Crime Reduction Cities	
New Orleans	-.64
New York City	-.60
Boston	-.55
Los Angeles County	-.54
Detroit	-.42
Baltimore City	-.41
Dallas	-.41
San Francisco	-.39
Low Crime Reduction Cities	
St. Louis	-.34
Jacksonville	-.28
City Of Fort Worth	-.28
Washington Metropolitan	-.26
San Diego County Sheriff	-.18
El Paso County Sheriff	-.18
Cleveland	-.17
San Jose	.02
Philadelphia	.10
Columbus*	.18
Memphis	.18
Phoenix	.25
San Antonio	.28
Houston	.31
Austin	.44
Milwaukee	.91
Indianapolis	1.15
Chicago	n.a.
Charlotte	n.a

* The entry for Columbus is for the period 1986-2011.

Violent crimes include murder, rape, robbery, aggravated assault. For a detailed description, see `http://www.ucrdatatool.gov/offenses.cfm`.

Source: Uniform Crime Report `http://www.ucrdatatool.gov/Search/Crime/Local/TrendsInOneVarLarge.cfm`

Table A8: Population Change

City	Population 1970	2010	Change	% Change
		Shrunk		
Detroit	1,511,482	713,777	-797,705	-53
St. Louis	622,236	319,294	-302,942	-49
Cleveland	750,903	396,815	-354,088	-47
New Orleans	593,471	343,829	-249,642	-42
Baltimore	905,759	620,961	-284,798	-31
Philadelphia	1,948,609	1,526,006	-422,603	-22
DC	756,510	601,723	-154,787	-20
Chicago	3,366,957	2,695,598	-671,359	-20
Milwaukee	717,099	594,833	-122,266	-17
Boston	641,071	617,594	-23,477	-4
Sum	11,814,097	8,430,430	-3,383,667	-29
		Grew		
New York	7,894,862	8,175,133	280,271	4
Memphis	623,530	646,889	23,359	4
Indianapolis	744,624	820,445	75,821	10
SF	715,674	805,235	89,561	13
LA	2,816,061	3,792,621	976,560	35
Columbus	539,677	787,033	247,356	46
Jacksonville	528,865	821,784	292,919	55
Dallas-FW	1,237,877	1,939,012	701,135	57
Houston	1,232,802	2,100,263	867,461	70
San Diego	696,769	1,307,402	610,633	88
El Paso	322,261	649,121	326,860	101
San Antonio	654,153	1,327,407	673,254	103
San Jose	445,779	945,942	500,163	112
Phoenix	581,562	1,445,632	864,070	149
Charlotte	241,178	731,424	490,246	203
Austin	251,808	790,390	538,582	214
Sum	19,527,482	27,085,733	7,558,251	39

Source: US Census.

Variable construction

PRICE **2-3 Bedroom single-family home** The decennial censuses and the ACS ask the owner of owner-occupied single-family homes the estimated value of the value of their home. While self-reported assessments of values, we will refer to them as housing prices.

In order to obtain numbers that are close to market prices, we restrict the sample to those who moved in within the last ten years, on the assumption that homes that owners of more recently bought and sold units would be more knowledgeable about price developments than those with longer tenure.

To obtain a price that refers to comparable units while preserving sample size, we focus on two or three bedroom single-family homes. Housing prices are given in intervals and the bracket values vary across the years. Therefore, we focus on the median bracket value. For tracts where the median bracket is the top code, we assign the dollar value that would be the mean if the top bracket had had the same range as the penultimate bracket. For instance, if the penultimate bracket ranged from 800 thousand to 1 million, and 1 million and above were the top bracket, we would give houses in the top bracket a 1.1 million dollar valuation. We chose this rule because it is conservative and if anything result in an underestimate of the price increases close to the CBD.

We focus on the median home price in the tract. Top coding is our reason for focusing on the median rather than a more selective percentile.

In our main specification, we impute a value that is the same distance from the top-code threshold as the immediately preceding midpoint value. For instance, if the top bracket is 800-1000, and values above 1000 are top coded, we assign the value 1100 to the top code. This method is conservative if the housing price distribution, like the wealth and income distributions, is right skewed.

We also tried alternative top codings, including simply imputing the threshold value. Our qualitative results were not sensitive to these variations.

For the study period 1980-2010, our measure shows a 30 percent increase in constant dollars, a rise largely inline with the Case-Shiller national index's rise of 26 percent. The Case-Shiller national index went from 43.44 to 145.0, and the CPI from 100 to 264, yielding a constant dollar housing price increase of 26 percent ($=\frac{145}{43.44} \times \frac{100}{264} - 1$).

Another benchmark is offered by the Census Bureau's constant price index for new-family homes sold rose from \$72 to \$272 thousand, current dollars. In constant dollars, the increase was 30 percent ($=\frac{261}{76} \times \frac{100}{264} - 1$). `https://www.census.gov/const/www/constpriceindex_excel.html`

We use the Consumer Price Index (CPI) to deflate housing prices. A theoretical concern is that housing accounts for some 40 percent of the CPI and therefore a price index that excluded housing would be preferable, for instance the food price index. Conveniently, the food price index moved very closely with the CPI and on those grounds we use the CPI despite its theoretical shortcomings.

Center point, Central Business District (CBD) For each city we identified the center with the help of Google maps. All cities had a clear central area identifiable from the convergence of roads, the presence of a main railway station, clusters of hotels with the national chain name prefixed by "down town" and a concentration of signature institutions and monuments. We used the thus identified area to designate a city center, with one exception. For San Jose, we placed downtown in Silicon Valley.

The only city with more than one clear center was New York City, where both midtown and downtown can claim that title. We picked the midtown center but locating the center downtown resulted in similar results. Because of this ambiguity, in a robustness test we exclude the New York metropolitan area and results are robust to this exclusion (Table A6, column 3).

Within each center we picked a center point, a salient building or monument, and obtained its latitude and longitude from iTouchMap.com. For instance, for Washington DC, the CBD is given by the White House and for San Jose, Apple Inc. Headquarters. While clearly there are alternative points, but most contenders would be within a mile or two of the points picked, listed in Table A1.

Distance to the CBD (*dist*) From tract shape files, we have generated latitude and longitude for the (population weighted) centroid of the tract, allowing us to calculate the distance between a tract and the CBD. We restrict our sample to tracts that are within 35 miles of the CBD.

In the preliminary analysis, we found our variables of interest to exhibit a pronounced j-shape with respect to distance from the CBD, the minimum located somewhere in the 3-7 mile range. Therefore we grouped distance from the CBD as follows:

d1 0-3 miles

d2 3-10 miles

d3 10-20 miles

d4 20-35 miles, reference category.

FT **Full Time** $FT(h,e)$ denotes the fraction of adults 25-55 who worked more that $h, h = 40, 50$ hours per week and had education $e, e = BA+, MA+$.

Z **Metropolitan Area Demand Shifter for Skilled Workers** See description in Section 2.2

IV.1 Panel data set construction

We use US2010 cross-walk files to create a tract-level panel data set, where for each tract in 2010 we construct its equivalent in previous years using cross-walk files from `http://www.s4.brown.edu/us2010/Researcher/Bridging.htm`. These files provide a mapping of tracts in a census year and 2010, as well as weights. For example, if there were a new tract B in 2010 that was the result of combining blocks from year 2000 tracts B1 and B2, then we would create a year 2000 tract B as a weighted average of B1 and B2. Thus, the ability to include tract fixed effects comes at the cost of data quality, and the problem worsens with the number of years to 2010 (more time allows for more changes to tracts). In order to explore the role of race, we are interested in parsing the sample by initial fraction blacks. Therefore, we restrict the sample to tracts for which we have an observation in 1980 and also remove tracts for which we only have one observation. We are left with a sample of about 48 thousand tracts.