Unaffordable housing and local employment growth: Evidence from California municipalities

Ritashree Chakrabarti
IHS Global Insight, USA

Junfu Zhang
Clark University, USA

Abstract

It is widely believed that unaffordable housing could drive businesses away and thus impede job growth. However, there is little evidence to support this view. This paper presents a simple model to clarify how housing affordability is linked to employment growth and why unaffordable housing could negatively affect employment growth. The paper then investigates this effect empirically using data on California municipalities. For various reasons, a simple correlation between unaffordable housing and employment growth cannot be interpreted as causal. Several empirical strategies are employed to identify the causal effect of unaffordable housing on employment growth. The estimation results provide consistent evidence that unaffordable housing indeed slows local employment growth.

Introduction

Housing prices and their growth rates vary substantially across regions in the United States. For example, the median sales price of single-family homes in the San Francisco area was $805,400 in 2007, compared with a median price of $130,000 in the Cleveland area. According to the S&P/Case-Shiller Home Price Index, home values appreciated by 354% in San Francisco from January 1987 to January 2007, whereas they rose only 122% in Cleveland.

In regions where housing prices are relatively higher or grow faster, there are always concerns that unaffordable housing could adversely affect local economic growth.[1]

Corresponding author:
Junfu Zhang, Department of Economics, Clark University, 950 Main Street, Worcester, MA 01610, USA.
Email: juzhang@clarku.edu

Among local policy makers, it appears to be a general belief that high housing prices increase the costs of living and doing business, make a region less attractive to workers and businesses, and therefore hurt the regional economy by slowing down employment growth. However, there is little empirical evidence that supports such a belief.

Fast-growing housing prices in the early 2000s inspired many studies on local housing markets in the United States, mostly focusing on the supply side.[2] Using data from California, Quigley and Raphael (2005) show that in cities with more growth-control measures, housing prices are higher and grow faster, apparently due to a slow response of new housing construction to positive labour demand shocks. Glaeser et al. (2006) similarly find that in metropolitan areas with more stringent land-use regulations, positive labour demand shocks lead to slower population growth and faster housing price appreciation. This line of research has helped us better understand why housing price varies so much across regions, but it does not directly address the question of how high housing prices would affect local employment growth.

Some earlier research has examined the interactions between local housing and labour markets. Bover et al. (1989) show that interregional differences in house-price-to-earnings ratios are correlated with regional unemployment in the United Kingdom. Using data from the southeast region of the United Kingdom, Johnes and Hyclak (1994) estimate a system of equations to assess the role of housing price in short-run regional adjustments of unemployment. Using data from four US metropolitan areas, Johnes and Hyclak (1999) estimate an error correction model to show that housing price has a significant effect on the size of local labour force.

More recently, Saks (2008) and Zabel (2012) have made significant contributions to this literature. Saks (2008) investigates how housing supply regulations affect housing and labour market dynamics in metropolitan areas across the United States. She argues that land-use and other government regulations can lower the elasticity of housing supply, which in turn can change the geographic distribution of housing prices and alter the pattern of labour migration. As a result, employment growth will be lower in places where the housing supply is more constrained. Saks presents some empirical evidence that supports this hypothesis. Along the same line of research, Zabel (2012) estimates a richer model that incorporates in- and out-migration as well as spillover effects among nearby cities. He finds that positive demand shocks tend to produce more in- and out-migration and thus more churning of workers in high-housing-cost areas, but the resulting employment and wage increases are similar between high- and low-housing-cost areas.

The present paper differs from the existing literature in two respects. First, it directly addresses the question raised by local policymakers. Despite the concerns voiced in the popular media, little research focuses narrowly on the effect of unaffordable housing on local employment growth. The existing studies are motivated by broader questions such as how labour and housing markets are interconnected and how a local economy responds to labour market shocks. In contrast, this study seeks to answer the very specific question: If housing is less affordable in a city, is employment growth going to be slower? Second, this paper directly confronts the identification problem. In empirical research, exogenous sources of variations are crucial for identifying the effect of one endogenous variable on another and such variations are difficult to find. Earlier studies in this literature make little effort to justify the choice of exogenous variables in model estimation. Saks (2008) and Zabel (2012) pay more attention to this problem. Following Bartik (1991), they

measure labour demand shocks using a weighted average of national industry employment growth where the weights are given by the industrial composition of the local economy. Saks (2008) goes further by interacting this Bartik measure with an index of local housing supply regulation to make the exogeneity assumption more plausible. The Bartik measure, although rather commonly used, is an ad hoc formulation. In contrast, this study will propose a theory that not only explains why unaffordable housing may affect local employment growth, but also provides a candidate instrumental variable for estimating this effect. Therefore, the identification strategy in this paper ties more closely to the underlying theoretical framework.

This study has two objectives. First, to develop a simple model to clarify why housing affordability varies among cities, and under what conditions unaffordable housing negatively affects local employment growth. The model reveals two insights: (1) different levels of amenities in different cities drive the variation in housing affordability; and (2) cities with unaffordable housing experience slower employment growth, because land rents are so high in those cities that they have already reached the very inelastic portion of their land supply curves.

Second, to test whether unaffordable housing indeed negatively affects employment growth. Data on California municipalities are used to empirically measure the effects of unaffordable housing on employment growth. Given the potential simultaneity and omitted-variables problems in OLS regressions, climate amenity variables are used to instrument for housing affordability, a solution suggested by the theoretical model.

Theoretical framework

This section presents a simple model to clarify thinking and motivate empirical research.

The starting point is a simplified version of the well-known Roback (1982) model, which uses variations in urban amenities to explain differentials in land rents and income across cities. Some explicit assumptions are then made about land supply in different cities, which within the Roback framework imply a relationship between housing affordability and local employment growth.

Consider an economy that consists of many cities, each endowed with some amenity level a. One could think of the amenity as the total number of sunny days in a year, or the average daily temperature in winter.

City residents are all workers. A representative worker consumes a composite good x and land s (lot size, as part of housing), and enjoys amenity a. He has the following utility maximisation problem:

$$\text{Max } U(x, s, a)$$
$$\text{s.t. } x + rs = w + m \tag{1}$$

where r is land rent; w is wage income; and m is non-labour income. The price of the composite good is determined on the international market. It is used as the numeraire and normalised to 1. To keep notation clean, city indexes are suppressed here. However, it should be noted that a, r, and w all vary across cities.

Equation (1) defines the representative worker's indirect utility function $V(w, r, a)$. Assume that workers can freely move from one city to another at no cost. In equilibrium, every worker attains the same level of utility u:

$$V(w, r, a) = u \tag{2}$$

V increases with a and w and decreases with r. That is, $V_a > 0$, $V_w > 0$, and $V_r < 0$.

There are also firms located in these cities. All firms have access to the same technology, which uses land and labour to produce the composite good. The production function is written as $f(n, d)$, where n is

the number of workers and d is the quantity of land used in production. Following Roback (1982) and subsequent literature, the analysis here ignores any capital used in production. Alternatively, one could admit the use of capital, but assume a fixed capital-to-labour ratio (such as one computer for each worker). In that case, as long as the price of capital is not determined locally, the production function can still be written this way without an explicit capital input.

Assume that function f exhibits constant returns to scale. Then the production technology can be represented using the unit cost function $C(w, r)$, which gives the minimum cost of producing one unit of good x. Firms are free to enter or exit the market, and can move costlessly from one city to another. This implies that in equilibrium firms everywhere have the same unit cost, which equals output price:

$$C(w, r) = 1 \qquad (3)$$

Note that $C_w > 0$ and $C_r > 0$.

Workers consume land only as part of housing. Housing here refers to a physical structure attached to a piece of land. For simplicity, assume that every worker lives in the same kind of physical structure, which is produced and assembled on the international market. Therefore, the physical structure is just a part of the composite good x. Let b be the amount of the composite good that constitutes the physical structure of housing, and assume that in equilibrium, $b \ll x$ for any worker. A worker's spending on housing is therefore $b + rs$. Following common practice, housing affordability is measured using the ratio of housing price to labour income:

$$h = (b + rs)/w \qquad (4)$$

A higher h implies that housing is less affordable.

Next, examine how h varies from one city to another. Differentiating equations (2) and (3) with respect to a yields:

$$dw/da = V_a C_r / (V_r C_w - V_w C_r) < 0 \qquad (5)$$

$$dr/da = V_a C_w / (V_w C_r - V_r C_w) > 0 \qquad (6)$$

These equations imply that in equilibrium, a city with a higher level of amenity has a lower wage rate and a higher land rent. It is an intuitive result. Because workers enjoy amenity, they are willing to accept a lower wage and pay a higher rent in a city with higher amenity. At the same time, firms' production is not affected by amenity. In a city with higher amenity, a firm can still break even: although it has to pay a higher rent, it offers a lower wage to workers, so its unit cost remains the same.

Further, assume that a worker's demand for land is inelastic, so an increase in land rent will never lower the worker's expenditure on land: $d(rs)/dr \geq 0$. Together with equations (5) and (6), this assumption implies $dh/da > 0$, meaning that housing is less affordable in a city with a higher level of amenity.

This modelling framework just laid out allows us to examine how unaffordable housing affects local employment growth. Starting with an equilibrium, consider a change in the total number of workers in the national economy. One may imagine that a cohort of college graduates just entered the labour force, or that a group of immigrants just arrived. We ask the following question: Under what conditions will a city with less-affordable housing experience slower employment growth? Because the composite good is sold on the international market, the influx of workers to a city does not affect the price of x. However, in principle, a change in the number of workers (N) in a city would affect r and w, and thus a worker's utility in this city.

Differentiating equations (2) and (3) with respect to N and substituting for dw/dN yield:

$$(V_r - V_w C_r/C_w)(dr/dN)\Delta N = \Delta u \qquad (7)$$

Equation (7) shows that a change in N affects r, which in turn causes a change in a worker's utility through two channels. First, a change in r directly affects a worker's utility (by V_r). Second, the change in r also causes firms to adjust the wage rate so their unit cost remains the same, which in turn affects utility (by $-V_w C_r/C_w$).

To illustrate the idea in the simplest way, assume that $(V_r - V_w C_r/C_w)$ is constant across cities. Note that under both the original equilibrium and the new equilibrium – before and after an influx of workers into the economy – indirect utility has to be the same everywhere. Therefore, Δu will be the same in all cities. However, dr/dN may vary from one city to another, which implies that ΔN will be different in different cities. In particular, a city with a higher dr/dN will have a lower ΔN.

In each city, equilibrium land rent is determined by land supply and demand in the city. Land supply refers to the quantity of land available for industrial or residential uses as a function of land rent. Land demand refers to the quantity of land demanded by workers and firms as a function of land rent, which is ultimately determined by the number of workers who reside in the city and the amenity level in the city.

Assume that the land supply function in each city has the following property: land is perfectly elastically supplied initially. As long as city residents and firms are willing to pay the opportunity cost of land in the agricultural sector, they can use more land and expand the city. However, this process cannot go on forever. After the city boundary reaches a certain limit, reflecting local land-use regulations or geographical constraints, land (for urban uses) can be supplied only at an ever-higher cost (Saiz, 2010; Saks, 2008).

More specifically, it is assumed that city i has an (inverse) land supply function, as follows:

$$r = \begin{cases} r_a & \text{if } q \le \bar{q}_i \\ r_a + (q - \bar{q}_i)^\rho & \text{if } q > \bar{q}_i \end{cases}$$

where r is land rent in the city, and q is the quantity of land available for residential and industrial uses in the city; r_a is the cost of land in the agricultural sector, which, for simplicity, is assumed to be the same everywhere; and $\bar{q}_i$ is the maximum amount of land that can be supplied to the city at the opportunity cost in agriculture. Note that $\bar{q}_i$ varies from one city to another due to local regulations and geographic conditions. Further assume that $\rho > 1$, so that dr/dq increases with q when q is higher than $\bar{q}_i$.

It is straightforward to show that in the system of cities described above, employment growth is slower in cities with less affordable housing. Consider a simple heuristic example depicted in Figure 1: an economy with only two cities. Both cities have the same land supply curve, that is, $S_1 = S_2$. City 1 has a higher amenity level than city 2, that is, $a_1 > a_2$. Suppose the initial equilibrium is attained when N_1 workers live in city 1 and N_2 workers live in city 2, and thus their land demand curves are labelled as $D(N_1)$ and $D(N_2)$. Note that equilibrium
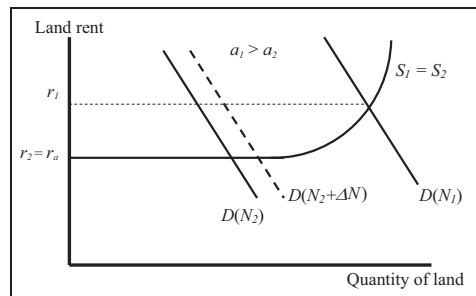

Figure 1. Responses to an influx of workers in an economy with two cities.

land rent in city 1 has to be higher ($r_1 > r_2$), because city 1 has a higher level of amenity. Also, because city 1 has a higher level of amenity, equilibrium wage rate is lower and therefore housing is less affordable in city 1.

Imagine a small number of workers, ΔN, are now added to the economy. In equilibrium, these new workers will be absorbed by city 2, because the in-migration does not influence equilibrium land rent in city 2. More generally, when there are many cities, one would expect that the increase in the number of workers will be smaller in a city where the equilibrium land rent is already on the inelastic portion of the land supply curve.

In other words, in a city with unaffordable housing, employment growth is smaller because equilibrium land rent is very high and land supply is more inelastic. An inelastic land supply implies that even a small increase in demand for land as a result of employment growth pushes the city's land rent much higher and drives workers to other cities with lower rents. Therefore the city can accommodate only moderate employment growth.

In summary, the model presented above has two implications:

Housing is less affordable in cities with higher amenities. This is because higher amenities lead to lower wages and higher housing prices, which together imply less affordable housing. Higher housing prices in high-amenity cities are mainly driven by higher land rents.

Cities with less-affordable housing experience slower employment growth. Less affordable housing reflects higher equilibrium land rent, and higher land rent indicates inelastic land supply that can be sustained only by more stringent regulatory and/or geographic constraints on the supply of urban land. Thus unaffordable housing is essentially an indicator of binding constraints on land supply. In cities with unaffordable housing, these constraints restrict local employment growth.

Empirical strategy

This section discusses the problems associated with estimating the effect of unaffordable housing on employment growth. The estimation equation is as follows:

$$y_{i,t} = \alpha + \beta h_{i,t-1} + \lambda X_{i,t-1} + \tau_t + \varepsilon_{i,t} \quad (8)$$

where the dependent variable $y_{i,t}$ is employment growth in city i and period t; $h_{i,t-1}$ is the key independent variable measuring housing affordability in city i and period $t-1$; $X_{i,t-1}$ represents a vector of control variables; τ_t is a year fixed effect; and $\varepsilon_{i,t}$ is the error term of the regression.

Note that the theoretical model follows the standard practice of Roback (1982) to treat all city residents as workers. Of course, in reality city population and total employment are two different concepts. The empirical specification in equation (8) uses employment growth rate as the left-hand-side variable. However, one might suspect that for various reasons cities of different population sizes may experience different rates of employment growth. Therefore, we will include city population size in the equation as one of the control variables.

A simple OLS regression of equation (8) will likely produce a biased estimate of β. In the theoretical model, equilibrium is assumed to occur instantaneously after any shocks hit the system of cities – the adjustment process is ignored. In empirical work one must treat this assumption with caution, because the adjustment to a new equilibrium takes time, and data collected in out-of-equilibrium situations may bias the coefficients in a simple OLS regression.

There are two types of potential biases. First, there may be some simultaneity bias. The goal of empirical analysis is to investigate how unaffordable housing affects employment growth. However, a simple OLS regression may also pick up a reverse causal effect. For example, rapid job growth

in a city, resulting from exogenous shocks, can raise the land rent and thus housing price in the city in the short run (if the city has already reached the inelastic portion of its land supply curve). Over time, workers and businesses will migrate to other cities to take advantage of the lower rents in those places, pushing land rent and housing price back toward their original equilibrium levels. If data are collected during this adjustment period, a simple OLS regression may show a positive relationship between unaffordable housing and employment growth, even if unaffordable housing leads to slower employment growth in equilibrium.

The second concern in estimating equation (8) is the problem of omitted variables. A simple OLS regression might fail to take into account some relevant but unobserved factors. For example, suppose some cities have just introduced a new regulation to restrict land use. This will push land rent, and thus housing price, higher in the short run. Again, over time, workers and businesses will migrate out, so the prices will move back toward their original equilibrium levels. During this adjustment period, both housing affordability and employment will change, although neither one is causing the other to change. In general, many new regulations and land-supply shocks could influence both housing price and employment growth. If not properly controlled, they will contaminate the estimated effect of unaffordable housing on employment growth that we intend to measure.

A few empirical strategies are employed to tackle the problems with simple OLS regressions. First, empirical analysis is conducted at the city level within a single state, California. This helps avoid the potential bias caused by unobserved heterogeneities at the state or higher levels that we expect to confound empirical studies based on nationwide data.

Second, predetermined affordability is used to predict employment growth in all empirical specifications. The idea is that if growth is not anticipated, it will not affect predetermined affordability measures. Thus the use of independent variables measured at the end of the last period should help mitigate the simultaneity bias.

Third, some specifications will include city fixed effects in the main equation, using within-city variations over time to identify the effect of unaffordable housing on local employment growth. This approach also helps mitigate the potential bias from unobserved heterogeneities across cities.

The fourth strategy is to use the instrumental variables (IV) approach to correct for both the simultaneity and omitted-variables biases in simple OLS regressions. To identify the effect of housing affordability on employment growth in a city, one needs a variable (or a set of variables) that affects local housing affordability but does not directly influence local employment growth. The theoretical model predicts that housing affordability is a function of the amenity level in a city. Thus local amenity measures are natural candidates for instrumental variables used to isolate the effect of housing affordability on employment growth.

Specifically, we will use weather variables (average July maximum temperature and average January minimum temperature) interacted with state-level energy cost (electricity price) to instrument for housing affordability. The primary reason for using the interaction terms instead of weather variables alone is that energy prices may affect climate (dis)amenities. For example, a hot summer may not be that unbearable if air conditioning is cheap. Therefore, an interaction between extreme temperature and electricity price gives a more accurate measure of the amenity that really matters.

These amenity variables qualify as valid instruments if: (1) they are strongly

correlated with housing affordability; and (2) they can be excluded from the main equation. Condition (1) is implied by the theoretical model, and, as will become evident below, is born out in the data. Condition (2) is a strong assumption we made to attain model identification. One might worry about the exclusion condition because nice weather tends to attract residents to a city, which is assumed in our model here and documented in earlier studies (e.g. Graves, 1980; Poelhekke, 2006; Rappaport, 2007). That is, equilibrium city population is generally related to climate amenities in the city. However, such amenities do not necessarily affect employment growth rate directly, especially if one believes that climate-induced population migration across cities had reached equilibrium before the study period. Nonetheless, to be cautious, control variables are added in the regressions, including log city population, the proportion of adult population with a bachelor's degree, and time dummies. After controlling for a city's equilibrium population level and its equilibrium employment of skilled labour, it seems reasonable to assume that employment growth rate is not directly affected by weather.

With instrumental variables, the effect of housing affordability on local employment growth is estimated using the two-stage least squares (2SLS) method, treating housing affordability as an endogenous variable in equation (8). The first stage equation is given by:

$$h_{i,t-1} = \gamma + \theta A_{i,t-1} + \delta X_{i,t-1} + \mu_{t-1} + \eta_{i,t-1} \tag{9}$$

where, as before, the subscripts i and t index cities and years; $X_{i,t-1}$ is a vector of city characteristics as controls; μ_{t-1} represents a year fixed effect; and $A_{i,t-1}$ is the set of instruments. The predicted housing affordability ($\hat{h}$) from this first-stage regression is

used to estimate the employment growth equation in the second stage.

Ideally, one would estimate fixed-effects models using instrumental variables, which presumably will produce the most convincing results. Unfortunately, the climate amenity variables used as instruments do not vary a lot within a small area, especially over a short period of time. Consequently, these instruments are not useful in fixed-effects models. Therefore, the empirical analysis presented below tries the IV approach and the fixed-effects approach separately.

Data and variables

Our empirical analysis is conducted at the level of California municipalities. The main advantage of focusing on a single state is that one does not need to worry about heterogeneity in state policies, because all the jurisdictions within a state are subject to the same state-level regulations. However, a state may be too small to have wide regional variations. For this reason, we have chosen a large state where both housing prices and climate amenities vary drastically across regions. Regression analysis at the municipality level instead of the county or metropolitan-area level ensures a reasonably large number of observations. Our analysis focuses on employment growth over two-year periods from 1993 to 2004. The choice of these time intervals is largely dictated by data availability.

Dependent variable

Employment growth. For California cities, two-year employment growth is calculated using data from the state's Employment Development Department. This database contains average yearly employment at the city level, collected by the department's Labor Market Information Division in cooperation with the US Department of Labor and the Bureau of Labor Statistics.

Independent and control variables

Housing affordability. This variable is calculated by dividing city-level median housing price by county-level median household income. Data on median housing prices are downloaded from the Business and Economic Statistics Section of RAND California (see RAND California, no date).[3]

County-level median household income data come from the Small Area Income and Poverty Estimates (SAIPE) programme of the US Census Bureau. The programme provides estimates of key income and poverty statistics for small geographic areas in non-census years. Prior to 1998, the bureau produced county-level income data in odd-numbered years only, so data are not available for 1994 and 1996. The missing data for these two years are imputed by taking the average of the preceding and the following year.

Fraction of adult population with a bachelor's degree. Data on educational attainment come from the 1990 and 2000 census, respectively. These data give the share of the adult population 25 years and older with a bachelor's degree in each city for the two census years. The data for other years from 1992 to 2004 are imputed, assuming a linear trend.

City population. Population data for California cities are downloaded from the website of the state Department of Finance. They estimate the total population of each city each year based on the 1990 and 2000 census data.[4]

Instrumental variables

July and January temperature. These weather variables are obtained from the National Climatic Data Center's monthly surface data files (see NOAA National Climatic Data Center, no date). Data on the July maximum temperature and January minimum temperature are extracted for all available, active California weather stations. Both temperature variables are calculated by averaging the daily maximum or minimum temperature over the month's 31 days. City names (contained in weather station names) allow us to match the weather variables with other city-level data.

California state level electricity price. Data on annual electricity prices for California are obtained from the Energy Information Administration (see US Energy Information Administration, 2012).

Empirical results

Descriptive statistics

There are 478 incorporated cities and towns in California, but the NCDC city list for California is much shorter, because many smaller cities have no weather stations. After matching all the variables from different sources, the study sample covers 115 cities in the state for the period 1993–2004, including all the large cities. The housing price variable and the weather variables may be missing for certain years for some cities, and therefore this sample is not a balanced panel. Table 1 presents descriptive statistics for the dependent, independent/control, and instrumental variables.

The average California city in the sample has 140,295 residents, with an employment level of 64,014. Average employment growth is 2.0% over one year and 3.9% over two years. The key independent variable – the housing affordability ratio – averages 4.5. It varies substantially, ranging from a minimum of 0.78 to a maximum of 16.42. Table 2, which lists the housing affordability ratio for a selected group of cities in California, illustrates this variation. For example, the mean ratio in the inland city of Fresno is 2.97, while it is as high as 11.46 in the coastal city of Santa Monica.

Table 1. Descriptive statistics for California cities.

Variables	Variable description	Summary statistics			
		Mean	Std. dev.	Min	Max
Dependent variables					
total_emp	Total city employment in California	64,014	175,159	570	1,733,300
lmid_empg1	City-level employment growth over a one-year period in California	0.02	0.055	−0.684	0.488
lmid_empg2	City-level employment growth over a two-year period in California	0.039	0.078	−0.662	0.498
Independent/control variables					
affordability_ ratio	Median home price in the city, divided by median household income in the county	4.49	2.11	0.78	16.42
percent_bachelor	Percentage of population 25 years and older in the city with a bachelor's degree	15.31	7.58	3.06	37.99
tot_pop	City-level annual population	140,295	389,261	1374	3,822,955
Instrumental variables					
jul_max	July mean maximum temperature	86.07	10.78	62.2	111.8
jan_min	January mean minimum temperature	40.99	5.73	9.9	52.4
elect_price	Annual electricity price in California	11.41	0.60	10.6	12.64

Note: Employment growth is defined as the difference in log employment.

The weather variables, shown in Table 1, also reveal a great deal of variation across cities. The July maximum temperature ranges from a minimum of 62.2 degrees to a maximum of 111.8 degrees, and the January minimum temperature ranges from a minimum of 9.9 degrees to a maximum of 52.4 degrees.

Figure 2 shows a scatterplot of two-year employment growth over affordability ratio. The employment growth rate indeed appears to be negatively correlated with the affordability ratio, which is clearly shown by the fitted straight line. That is, municipalities with a higher affordability ratio (and thus less affordable housing) tend to have slower employment growth. This is only an unconditional correlation, but it is indicative that

unaffordable housing may have a negative effect on employment growth.

OLS and IV estimates

Table 3 presents OLS and IV regression results for California cities. To focus on the independent variable of interest, the coefficients of control variables are not included in the table. Standard errors are clustered by county, allowing for both spatial and serial correlations among all observations within a county.

The OLS results are in the upper panel of Table 3. In the left column is the specification without city fixed effects. In this regression, housing affordability has a negative coefficient, but its magnitude is small

Table 2. Summary statistics on the housing affordability ratio for selected California cities, 1993–2004.

City name	Housing-price-to-income ratio			
	Mean	Std. dev.	Minimum	Maximum
Anaheim	3.75	0.53	3.28	4.97
Bakersfield	2.88	0.31	2.42	3.40
Burbank	5.86	1.04	4.99	8.49
Fresno	2.97	0.37	2.52	3.48
Laguna Beach	9.32	2.10	7.08	13.16
Lodi	3.74	0.44	3.32	4.69
Long Beach	4.62	0.61	4.00	6.00
Los Angeles	4.81	0.76	4.05	6.49
Los Banos	4.57	0.87	3.84	5.92
Los Gatos	8.01	1.30	6.54	10.09
Madera	2.93	0.29	2.69	3.50
Merced	3.45	0.47	3.05	4.12
Modesto	3.22	0.57	2.75	4.38
Monterey	8.26	1.22	6.90	9.97
Newark	5.27	1.14	4.11	7.33
Oceanside	4.21	0.62	3.71	5.72
Sacramento	2.84	0.54	2.38	3.94
San Bernardino	2.20	0.23	1.94	2.64
San Diego	4.58	0.65	4.16	6.26
San Francisco	7.78	1.50	6.45	10.19
San Jose	4.60	0.86	3.82	6.23
Santa Barbara	8.71	1.68	7.15	11.73
Santa Cruz	6.69	1.59	5.14	9.17
Santa Monica	11.46	1.46	9.87	14.46
Stockton	3.19	0.61	2.57	4.45
Turlock	3.36	0.53	2.97	4.44
Victorville	2.59	0.29	2.26	3.03
Visalia	3.58	0.23	3.26	3.97
Vista	4.29	0.61	3.71	5.69

Note: The statistics for each city are calculated using housing affordability ratios in the city over different years.

(−0.3%), and it is not statistically significant. That is, a simple comparison across cities does not reveal slower employment growth in less-affordable cities. The right column in the upper panel shows the estimate from the specification with city fixed effects. Here the coefficient is still negative but much larger (−2.5%), and it is statistically significant at the 1% level. Therefore, when focusing on within-city changes over time, we do find that slower employment growth tends to follow years with less-affordable housing.

The two regressions in the upper panel of Table 3 indicate that the OLS estimate

without controlling city fixed effects is biased toward zero. To understand this bias, let us consider two cities: San Jose and Fresno. Suppose indeed that within both cities slower employment growth follows periods with less affordable housing. Thus the fixed-effects OLS regression will reveal the negative effect of unaffordable housing on employment growth. Additionally, suppose that in San Jose, innovation creates new jobs and thus increases employment as well as housing prices. As a result, San Jose has both less affordable housing and faster employment growth than Fresno. That is, although within-city variations imply a negative effect
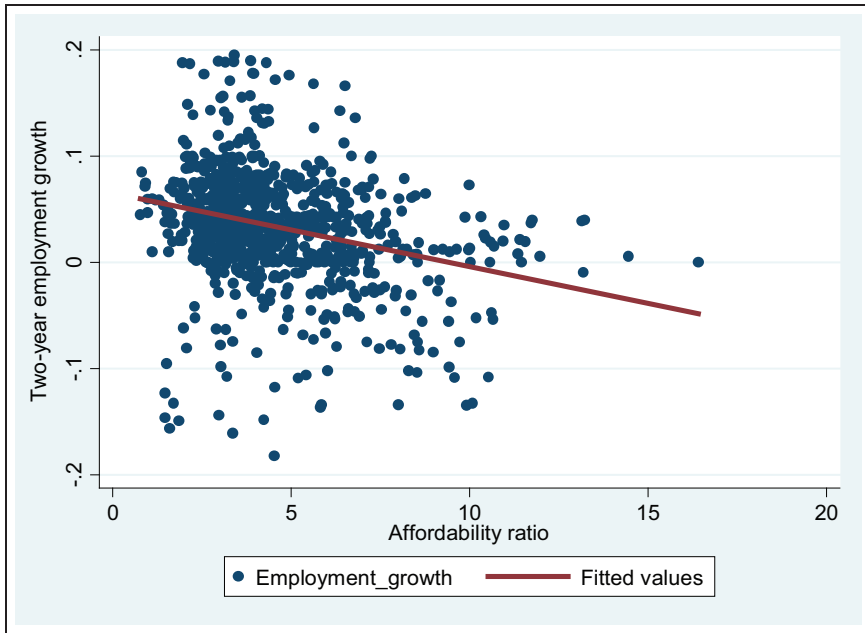
Figure 2. Housing affordability and two-year employment growth in California cities.
Note: A higher housing-price-to-income 'affordability ratio' means that housing is less affordable. Employment growth is regressed on affordability ratio, and the estimated coefficient is then used to predict the 'fitted value' at each affordability ratio.

of unaffordable housing on employment growth, between-city variations imply a positive effect. An OLS regression without controlling for city dummies mixes up these two effects and thus biases the estimate toward zero. An IV regression excluding city dummies also uses both within- and between-city variations. However, since the variations are exogenous and come from climate amenities, the IV regression helps correct the omitted variables biases generated from the between-city variations.

The lower panel of Table 3 shows the IV estimates, without controlling for city fixed effects. Different variables are used to instrument for the housing affordability ratio in the first-stage regression, including: (1) July maximum temperature interacted with electricity price; (2) January minimum temperature interacted with electricity

price; and (3) both interaction variables. Correspondingly, three sets of IV estimates are reported.

All the IV estimates show a statistically significant negative relationship between the housing affordability ratio and city-level employment growth. In each case, the coefficient is less negative (–1.2, –1.9 or –1.5 versus –2.5%) than the OLS estimate with city fixed effects. To understand this discrepancy, it is important to recognise that the IV regression estimates the coefficient mainly based on exogenous cross-sectional variation in housing affordability because weather variables (even interacted with electricity price) contain little over-time variation within a city. In contrast, the fixed-effect OLS estimate entirely relies on within-city variation over time. Given that the within-city variation is not necessarily exogenous, one might find

Table 3. Effects of unaffordable housing on city employment growth in California, 1993–2004.
(Dependent variable: employment growth in a California city over two years).

	OLS regressions	
	(1)	(2)
Affordability ratio	−0.003	−0.025***
	(0.002)	(0.008)
Controls	Yes	Yes
Year fixed effects	Yes	Yes
City fixed effects	No	Yes
R^2	0.10	0.31
No. of observations	903	903

	2SLS regressions		
	(1) IV: jul_max× elect_price	(2) IV: jan_min× elect_price	(3) IV: jul_max×elect_price & jan_min×elect_price
Affordability ratio	−0.012**	−0.019***	−0.015***
	(0.005)	(0.007)	(0.005)
	[0.005]	[0.006]	[0.004]
	{0.006}	{0.009}	{0.006}
Controls	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
City fixed effects	No	No	No
First-stage F-statistic	38.35	15.70	22.68
Hansen's J test p-value	–	–	0.319
Endogeneity test p-value	0.083	0.035	0.027
No. of observations	902	903	870
Standard deviation of affordability ratio in sample	2.125	2.083	2.085

*p < 0.10; **p < 0.05; ***p < 0.01.
Note: Standard errors clustered by county are in parentheses. For the IV results, alternative standard errors are also reported. Standard errors clustered by county-year are in square brackets; standard errors clustered by city are in curly brackets. Levels of significance, Hansen's J tests, and endogeneity tests are all based on standard errors clustered by county. Although not reported in this table, we included an intercept and controlled for log city population and the percentage of adult population in the city with a bachelor's degree in all specifications.

the IV estimates more convincing than the fixed-effects OLS estimate.

Consider our preferred IV specification, the third column that uses both interaction variables as instruments. The coefficient of the affordability ratio is –0.015. This implies that if a city's housing affordability ratio is higher by one unit (or about half a standard deviation), its employment growth rate over two years is expected to be lower by one and a half percentage points. This is a rather large effect, given that total employment in the average city grows by only 3.9% over two years (as shown in Table 1).

For all IV regressions, Table 3 also presents standard errors of the key coefficient based on alternative clustering methods. Standard errors clustered on county-years, allowing for spatial correlations within a county in a single year, are shown in square

brackets under the estimated coefficients. Standard errors clustered by city, allowing for serial correlations within the city over different years, are shown in curly brackets. In general, clustering on county-year leads to smaller standard errors than in the baseline regressions, and clustering on city leads to larger standard errors. However, in both cases the standard errors change only slightly from the baseline results, and the coefficient of the housing affordability ratio remains statistically significant under all specifications.

Table 3 also presents some results regarding the validity of the instrumental variables and regarding the presence of endogeneity biases in the OLS estimates. All these results are based on standard errors clustered by county.

Valid instruments must be correlated with the endogenous variable, and orthogonal to the error term. An F test for the joint significance of the instrumental variables in the first-stage regressions is performed to check the correlation between the instruments and the endogenous variable. This statistic ranges from 15.7 to 38.3 across different IV specifications, and is consistently greater than 10, suggesting that the instruments used in these regressions have good explanatory power.

In the third column, two interaction variables are used as instruments, although there is only one endogenous variable. This allows the use of an overidentifying test to check the validity of the instruments. Specifically, we conduct Hansen's J test for the null hypothesis that both instruments are proper instruments. The p-value reported in Table 3, 0.319, suggests that the null hypothesis cannot be rejected.

Although the discussion of estimation problems clearly suggested the potential endogeneity between employment growth and housing affordability, it is still instructive to empirically test for the presence of endogeneity here. Table 3 shows the p-values

of the statistics from the endogeneity tests. For all three specifications, these tests reject the null hypothesis (at the 10% level) that the OLS coefficient is unbiased. That is, it is likely that endogeneity is present when measuring the effect of unaffordable housing on employment growth and can potentially lead to seriously biased estimates in OLS regressions. Therefore, an IV estimate is preferred in this case.

The 2SLS regressions in Table 3 do not include county or city fixed effects. Presumably, such fixed-effects IV regressions would be the preferred specifications, because they would be the most conservative approach to dealing with unobserved heterogeneities among the cities. However, as noted above, year-to-year variations in both housing affordability and the weather variables would be small within any small geographic region over a short period. Therefore the correlation between these two variables would be too weak to identify the effects of unaffordable housing in fixed-effects models.

Because the affordability ratio variable and city-level weather variables are observed in every year, it is certainly possible to run the regressions with county or city fixed effects. Indeed, with county or city fixed effects, the first-stage F-statistics are small across all the specifications, confirming that the year-to-year correlation between weather variables and housing affordability within a county or a city is not strong enough to help identify the effects of unaffordable housing. The estimated coefficients under these fixed effects models blow up in some cases, and can never be precisely estimated. Since these exercises are not informative, we consider the baseline IV regressions (without controlling for county or city fixed effects) as our best-attainable results.

Table 4 reports results from some robustness checks. The first alternative specification adds another control variable: a dummy for coastal counties. The negative

Table 4. Effects of unaffordable housing on city employment growth in California, 1993–2004: alternative specifications (dependent variable = employment growth in a California city over two years).

| | A. IV results on CA city employment growth, specification with coastal dummy | | |
	IV: jul_max× elect_price	IV: jan_min× elect_price	IV: jul_max×elect_price & jan_min×elect_price
Affordability ratio	−0.011	−0.024**	−0.016**
	(0.011)	(0.011)	(0.008)
First-stage F-statistic	15.56	5.62	9.80
No. of observations	902	903	870
	B. IV results on CA city employment growth, specification with additional controls		
	IV: jul_max× elect_price	IV: jan_min× elect_price	IV: jul_max×elect_price & jan_min×elect_price
Affordability ratio	−0.012*	−0.022***	−0.016**
	(0.006)	(0.007)	(0.006)
First-stage F-statistic	30.53	28.52	12.07
No. of observations	639	634	613

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Note: Standard errors clustered by county are in parentheses. Although not reported in this table, in all regressions, we included an intercept and controlled for the year fixed effect, log city population, and the percentage of adult population in the city with a bachelor's degree in all specifications. In addition, in panel B, we also controlled for per capita government general expenditure; share of government expenditure spent on road, fire protection, and police; retail and wholesale employment as a share of total employment; manufacturing employment as a share of total employment; and violent crime rate.

effect observed in the baseline IV regressions might result from the fact that housing is generally less affordable in coastal cities (refer to Table 2). In other words, the estimated coefficients in Table 3 might be picking up only this 'coast effect'. By introducing the coastal dummy, this alternative specification identifies the effects of unaffordable housing using only the variations within the coastal or inland areas.

Panel A of Table 4 shows the results from the regressions with the coastal dummy. As with the baseline regressions, standard errors are clustered by county. The IV coefficients are all negative. They are estimated with less precision but two of the three are still statistically significant. This suggests that although the variations between coastal and inland cities play a role in identifying the effect of unaffordable housing on employment growth, they alone are not driving the results. Our preferred specification, which uses both interaction variables as instruments, gives an estimated coefficient of 0.016, almost identical to the estimate without controlling for the coastal dummy.

Panel B of Table 4 reports results from specifications using more city-level controls. In all the specifications presented above, log city population and the percentage of adult population with a bachelor's degree in the city are included as controls. Here we check whether our baseline results are sensitive to including more city level controls. High frequency data on city characteristics are not

easily available. We hand-collect information from the 1994, 2000, and 2007 editions of the County and City Data Book published by the US Census Bureau, and impute the data for the years in between by linear interpolation. The extra city level control variables are related to local public finance, economic structure, and living environment, including: (1) per capita government general expenditure; (2) share of government expenditure spent on road, fire protection, and police; (3) retail and wholesale employment as a share of total employment; (4) manufacturing employment as a share of total employment; and (5) violent crime rate. The choice of these variables is dictated entirely by data availability. Since these variables are missing for some cities, we end up with a smaller sample size. As it turns out, adding these extra control variables has no significant effect on our qualitative results. In fact, the estimated coefficients are very similar to those in the baseline regressions.[5]

Overall, results from these alternative specifications also consistently point to a negative effect of unaffordable housing on local employment growth in California. Our preferred specification, which uses both amenity variables as instruments, suggests that a one-unit increase in the housing-price-to-income ratio reduces city-level employment growth by 1.6 percentage points over two years, almost identical to the baseline estimate.[6]

Conclusion

A simple theoretical model is proposed to explain why housing affordability could affect regional employment growth. In the model, the variation of housing affordability is driven by heterogeneities in location-specific amenities. Cities with less-affordable housing tend to experience slower employment growth, because equilibrium land rents are

so high that the supply of land must have reached some constraints. These land supply constraints are the ultimate restrictions on local employment growth. Note that many land-supply constraints are actually man-made and created by land-use regulations. This model implies that when high housing prices are sustained by tight land-use regulations, they lead to slower employment growth. This is consistent with the findings of the existing literature that focuses on land-use regulations, such as Glaeser and Gyourko (2003), Quigley and Raphael (2005), and Saks (2008).

Potential endogeneity biases create a major problem in empirical studies of the effect of unaffordable housing. While housing prices may affect local employment growth, job growth could also make the local economy more prosperous, leading to an increase in housing prices. In addition, omitted variables also cause complications in empirical analysis. Many unobserved factors could affect housing affordability and employment growth simultaneously, resulting in spurious correlations between the two variables. All of these can bias the estimates obtained through simple OLS regressions.

This study has employed various empirical strategies to correct for the potential simultaneity and omitted-variables biases. Most importantly, identification relies on the instrumental variables approach. As suggested by the theoretical model, the empirical study uses measures of climate amenities, such as July maximum temperature and January minimum temperature (interacted with electricity price), to instrument for the housing affordability measure.

The IV regression results show significant negative effects of unaffordable housing on local employment growth. For California cities, other things being equal, a one-unit lower housing affordability ratio is associated with a two-year employment growth

rate that is about one and a half percentage points higher.

One limitation of this study is that the instrumental variables do not contain enough over-time variations to help identify the within-city effect of unaffordable housing. The OLS estimate controlling for city dummies suggests that the within-city effect is also negative. However, this estimate is not based on exogenous variations and may be subject to some endogeneity bias. Therefore one must interpret this result on within-city effect with caution and further research is warranted along this line.

Acknowledgements

Funding

Notes

1. See Box 1 in Chakrabarti and Zhang (2010) for quotes from newspaper articles reflecting widespread anxiety over high housing prices in California and other US regions.
2. The bulk of this literature focuses on how land-use regulations restrict land supply and in turn lead to higher housing prices. See, for example, Glaeser and Gyourko (2003); Glaeser and Ward (2009); Glaeser et al. (2005a, 2005b); and Ihlanfeldt (2007). Hwang and Quigley (2006) examine how a broad range of economic conditions and regulations affect the outcomes in local housing markets.
3. RAND originally acquired these data from the California Association of Realtors. The price reflects both sales of new homes and resales. RAND has price data from 1991 to 2002, all measured in nominal dollars.
4. These city-level population estimates, and the methodology used for the estimates, are available at http://www.dof.ca.gov/research/demographic/reports/view.php.
5. We have also tried including both the coastal dummy and the extra city level controls in the regression. The coefficients are also all negative, although as in panel A of Table 4 only two of them are statistically significant.
6. We also examined the effects of unaffordable housing on three- and four-year employment growth in California cities (results available upon request). In addition, we conducted parallel empirical analysis using census data on US metropolitan areas and US counties over ten-year periods (Chakrabarti and Zhang, 2010). The results from these additional analyses suggest that the longer-term effects of unaffordable housing are larger than the two-year effects reported here.

References

Bartik TJ (1991) *Who Benefits from State and Local Economic Development Policies?* Kalamazoo: W.E. Upjohn Institute for Employment Research.

Bover O, Muellbauer J and Murphy A (1989) Housing, wages and UK labour markets. *Oxford Bulletin of Economics and Statistics* 51: 97–136.

Chakrabarti R and Zhang J (2010) Unaffordable housing and local employment growth. Working paper 10–3, New England Public Policy Center at the Federal Reserve Bank of Boston, USA.

Glaeser EL and Gyourko J (2003) The impact of zoning on housing affordability. *Economic Policy Review* 9(2): 21–39.

Glaeser EL and Ward BA (2009) The causes and consequences of land use regulation: Evidence

from Greater Boston. *Journal of Urban Economics* 65: 265–278.

Glaeser EL, Gyourko J and Saks R (2005a) Why have housing prices gone up? *American Economic Review Papers and Proceeding* 95: 329–333.

Glaeser EL, Gyourko J and Saks R (2005b) Why is Manhattan so expensive? Regulation and the rise in housing prices. *Journal of Law and Economics* 48: 331–369.

Glaeser EL, Gyourko J and Saks R (2006) Urban growth and housing supply. *Journal of Economic Geography* 6: 71–89.

Graves PE (1980) Migration and climate. *Journal of Regional Science* 20: 227–237.

Hwang M and Quigley JM (2006) Economic fundamentals in local housing markets: Evidence from US metropolitan regions. *Journal of Regional Science* 46: 425–453.

Ihlanfeldt K (2007) The effect of land use regulation on housing and land prices. *Journal of Urban Economics* 61: 420–435.

Johnes G and Hyclak T (1994) House prices, migration, and regional labor markets. *Journal of Housing Economics* 3: 312–329.

Johnes G and Hyclak T (1999) House prices and regional labor markets. *Annals of Regional Science* 33: 33–49.

NOAA National Climatic Data Center (no date). *Climate data online*. Available at: http://cdo.ncdc.noaa.gov/CDO/cdo.

Poelhekke S (2006) Do amenities and diversity encourage city growth? A link through skilled labor. Working paper no. 2006/10. Florence: European University Institute.

Quigley JM and Raphael S (2005) Regulation and the high cost of housing in California. *American Economic Review Papers and Proceedings* 95: 323–329.

RAND California (no date) Housing sales prices and transactions in California counties and cities. Available at: http://ca.rand.org/stats/economics/houseprice.html.

Rappaport J (2007) Moving to nice weather. *Regional Science and Urban Economics* 37: 375–398.

Roback J (1982) Wages, rents, and the quality of life. *Journal of Political Economy* 90: 1257–1278.

Saiz A (2010) The geographic determinants of housing supply. *Quarterly Journal of Economics* 125: 1253–1296.

Saks RE (2008) Job creation and housing construction: Constraints on employment growth in metropolitan areas. *Journal of Urban Economics* 64: 178–195.

US Energy Information Administration (2012) State electricity profiles: Data for 2010. Available at: http://www.eia.doe.gov/cneaf/electricity/st_profiles/sept08ca.xls.

Zabel J (2012) Migration, housing market, and labor market responses to employment shocks. *Journal of Urban Economics* 72: 267–284.