Land-Use Restrictions: Implications for House Prices, Inequality, and Mobility*

Don Jayamaha†

October 23, 2019

Click here for the latest version

Abstract

I investigate the extent to which land-use restrictions, through their impact on house prices, act as a barrier to labor mobility. To do so, I develop a multi-region heterogeneous agents model of migration and housing, where land-use restrictions act as a friction affecting the productivity of land and the housing supply. Using the structure of the model along with data on regional prices, output, and housing densities, I estimate a panel measure of implied restrictions. Consistent with the existing measure of restrictions, the model-implied measure suggests that restrictions are most stringent in regions with high incomes and house prices. Further, the measure shows that the regions that were most restricted in the past have become even more restrictive over time. Calibrating the model to 2014, I show that the variation in regional productivities and land-use restrictions generate the income and house price gaps observed in the data. Performing a counterfactual exercise, I find that lowering the level of restrictions in California back to its level in 2000 results in a large reallocation of labor. The state's population rises by 45%, while the income gap and house value gap between California and the rest of the U.S falls by 3.7% and 2.7%, respectively. I also study the importance of borrowing constraints and moving costs in hindering labor mobility and find that conditional on the observed income and house price gaps, neither plays a significant role.

*I am indebted to my advisors Virgiliu Midrigan, Diego Perez, and Raquel Fernandez, for their guidance throughout this project. I am also grateful for the comments and suggestions provided by Jaroslav Borovicka, Katarina Borovickova, Tim Cogley, Venky Venkateswaran, Abdoulaye Ndiaye, Skand Goel, Hargungeet Singh, Laszlo Tetenyi, Robert McDowall, James Graham, as well as seminar and conference participants at NYU.

†New York University, Economics Department. Email: don.jayamaha@nyu.edu

1 Introduction

In recent decades, we have observed large and increasing income differentials across the many regions of the United States. Traditionally, we have considered the movement of labor to be the primary mechanism through which regional differences are smoothed out, as people leave poorer regions and move into more prosperous regions.[1] Since the 1980's however, this has not been the case. As has been well documented in the literature, we have seen a significant decline in labor mobility as evidenced by the steady decline in internal migration across the country.[2] Numerous works, including Nieuwerburgh and Weill (2010) and Ganong and Shoag (2017), have pointed out that high income regions have disproportionately higher house prices, alluding to the prohibitively high costs of housing acting as a barrier to entry. In this paper, I investigate the extent to which land-use restrictions are a potential source of the higher house prices and study the degree to which they prevent the mobility of labor.

Land-use restrictions vary significantly across the U.S., both in their stringency and in their particular form. The restrictions can be thought of as arising for two broad reasons. Firstly, natural limitations such as poor soil quality, a very steep gradient, or being located in an earthquake zone as is the case for parts of California, might make land unsuitable to hold large structures. Secondly, imposed land-use regulations, such as density requirements, floor area requirements, restrictions on the number of units allowed, and limitations on building permits, also limit the extent to which land can be developed. While the former is an innate feature of the land or surrounding geography, and hence cannot be changed, the latter is directly controllable by policy. Regardless, both broad groups of restrictions limit the productivity of land by restricting the amount of buildings and housing that can be put up, and consequently limit the supply of housing, which in turn raise the price of housing. However, since the impact of a given restriction varies so significantly over time and space, we do not have a good measure of restrictions that is representative and also comparable.

In order to get around this data limitation, I develop a multi-region model of migration where the level of land-use restrictions in a region directly affects the degree to which land

[1] Prominent papers documenting the adjustment of labor in response to regional differences include Blanchard and Katz (1992), and Barro and Martin (1991).

[2] Examples include Kaplan and Schulhofer-Wohl (2017) and Molloy, Smith, and Wozniak (2011).

can be used by the construction sector and production sector, to build housing and produce output, respectively. Combining the structure of the model with regional data on wages, house prices, available land, housing stock, and output, I estimate state-level panel measures of land-use restrictions and productivity. Comparing the model-implied measure of restrictions to the available measure of land-use regulations, detailed in Gyourko et al. (2008), I show that the level restrictions in a state is highly correlated to its level of regulation. This suggests that regulations may account for a significant share of the variation in restrictions across states.

Consistent with Hsieh and Moretti (2019) and Herkenhoff et al. (2018) who perform similar exercises, as well as other studies focusing on individual regions, I find that states with high levels of land-use restrictions also tend to have higher incomes, house prices, and productivities.[3] The states of California and Massachusetts which are known to have high restrictions are particularly striking examples of this relationship. The panel measure further suggests that most states have experienced a tightening of restrictions, and that in fact the states experiencing the largest increase in restrictions were those that were most restricted to begin with.

Since this project focuses on the impact of land-use restrictions on an agent's migration decision, through its affect on house prices, I model a rich household sector of the economy. Agents are heterogeneous in ability and wealth, make a dynamic migration choice as well as a continuous housing choice, and have access to a risk-less asset which they can use to borrow against their house. This marks a key departure from the existing literature, which has looked at migration decisions without all the ingredients that are fundamental to this model. Given the importance of housing in the agent's migration choice, it is essential to allow a continuous housing choice. Excluding this feature would force agents to purchase equally sized houses when they move from a poorer region to a higher income region, thereby overestimating the prohibitiveness of housing costs. This is not a feature of the data as agents move into smaller units when they migrate into higher house price regions. Similarly, the ability to save in a risk-less asset and borrow against one's house is essential in modeling a

[3] Studies include: Glaeser and Gyourko (2018) and Whittemore (2012) which look at cities, and Levine (1999) which focuses on California.

realistic housing and migration choice, as we see agents using savings to cover moving costs as well as taking out mortgages when purchasing houses.

Calibrating the model to two regions, California and the rest of the U.S. as at 2014, I find that the spatial variation in land-use restrictions and productivities generate the income and house value gaps observed in the data. The high productivity and severe land-use restrictions in California keep wages in the region high and push up house prices by restricting housing supply. The prohibitively high house prices prevent agents from moving into California despite the higher wage. Agents sort along two dimensions. The returns to the higher wage are increasing in an agent's ability, making California more attractive to higher skilled agents. Thus agents sort by ability. Furthermore, only high net worth agents are able to absorb the moving costs as well as the house price differential faced when moving, as they have to sell their less valuable house and purchase a considerably more expensive house in California. Consequently, the model also generates sorting by wealth.

Performing the key counterfactual exercise of the paper, I loosen the level of restrictions in California down to its level in 2000 keeping all else unchanged. I find that lowering restrictions has a significant impact on the agent's migration choice as the minimum ability and wealth thresholds required to move reduce significantly. Consequently, California's population increases by 45% as agents who previously found California's house prices prohibitively high now move in. The increase in California's housing supply and large reallocation of labor reduces the income and housing gaps between the two regions by 3.7% and 2.7%, respectively.

Further, the lower restrictions also raise aggregate output via two channels. Firstly, the reallocation of labor shifts a larger share of the population to the more productive region, leading to an increase in efficiency. Secondly, lower restrictions effectively increase the supply of usable land, which is a factor of production for the consumption goods producing firm, thereby raising the firm's output.

Lastly, I exploit the heterogeneity and richness of the households in the model to study the importance of credit conditions and moving costs for labor mobility. I find that conditional on the observed income and house price gaps, agents do not find it optimal to move even if they were allowed to increase their borrowings against their house. Further, while the moving cost does impact the migration rate in the steady-state of the model, it does not significantly affect the population share in each region.

The remainder of the paper is organized as follows. Section 2 discusses how this paper relates to the existing literature. Section 3 describes the model and highlights the features that allow me to separately identify land-use restrictions and productivities. Section 4 presents the model implied measures and provides some intuition regarding the patterns they depict. Section 5 studies the baseline model and presents the results from the counterfactual exercise of lowering land-use restrictions in California. Section 6 discusses extensions to the baseline model and immediate next steps, while section 7 concludes.

2 Related Literature

Migration: Halket and Vasudev (2014), Bilal and Rossi-Hansberg (2018), Eckert and Kleineberg, Kennan and Walker (2011), Roback (1982).

Regional divergence: Nieuwerburgh and Weill (2010), Eeckhout et al. (2014), Diamond (2016).

Housing supply: Herkenhoff, Ohanian, and Prescott (2018), Hsieh and Moretti (2019), Ganong and Shoag (2017), Glaeser and Gyourko (2018), Parkhomenko (2018), Jackson (2016), Turner, Haughwout, and van der Klaauw (2014), Saiz (2010), Favilukis, Mabille, and Nieuwerburgh (2018), Levine (1999).

3 Model

The model consists of two distinct parts. Firstly, I model a rich dynamic heterogeneous household sector, building on the standard Bewley (1977) framework, and expanding the economy to allow for multiple regions, migration, regional housing markets, and a continuous housing choice. This enables me to more carefully investigate the individual's migration choice as well as it's interaction with the agent's ability draw, asset holdings, and housing choice.

Secondly, I model a static non-durable good producing sector and construction sector in each region. The functional form assumptions in these sectors allow me to separately identify regional productivities and land-use restrictions, as well as back them out as functions of observables, as will be discussed in section 3.5.

3.1 Environment

The economy consists of N regions and a continuum of measure one of infinitely lived ex-ante identical agents who can move across regions each period. Time is discrete and agents maximize their expected life-time utility over an aggregate consumption good $\tilde{C}_t$ and the amenity value associated with their region of residence. The aggregate good is made up of non-housing consumption c_t and the individual's housing stock h_t as follows,

$$\tilde{C}_t = {c_t}^{\chi}(h_t - \bar{h})^{1-\chi},$$

where χ is the share of non-housing consumption in utility. The agent has log utility over this aggregate good. Each agent has a stochastic endowment of efficiency units of labor, $\epsilon_t \in E$. The shocks are i.i.d across agents and follow a Markov process with transition probability $\pi(\epsilon', \epsilon) = P(\epsilon_{t+1} = \epsilon' | \epsilon_t = \epsilon)$. Agents pay a fixed cost to relocate between regions, and must own their house. Agents can also borrow or save in a one-period risk-less asset at an exogenously given interest rate r. The problem of an agent in a given region n is given by,

$$\bar{V}(a, h, n, \epsilon) \quad = \quad \max \; log(\tilde{C}) + z_n + \beta \mathbb{E}[\bar{V}(a', h', n', \epsilon') | \epsilon] \tag{1}$$

s.t.

$$c + a' + P_{n'}h' + \mathbb{1}[n' \neq n]\kappa = (1+r)a + w_n\epsilon + P_n h(1-\delta) \tag{2}$$

$$a' \geq -\theta P_{n'}h', \tag{3}$$

where a refers to the agent's asset holdings, h is the housing stock, z_n is the amenity value of living in region n, κ is the fixed cost of relocating, w_n is the wage rate per efficiency unit in region n, P_n is the price of housing in region n and δ is the depreciation rate of housing. The agent's budget constraint is given by (2), while (3) is the agent's borrowing constraint, with θ being the maximum loan-to-value ratio.

To simplify the agent's problem I assume that households can choose their housing stock in period t after the realization of their ability shock ϵ_t. The agent's financial position at the beginning of period t is then summarized by net worth

$$b = a + P_n h,$$

and the agent's state in period t is (b, n, ϵ). The agent's problem in recursive form is then given by,

$$V(b,n,\epsilon) \;=\; \max_{b',n',c,h} \; log(\tilde{C}) + z_n + \beta\mathbb{E}[V(b',n',\epsilon')|\epsilon] \tag{4}$$

s.t.

$$c + b' + \mathbb{1}[n' \neq n]\kappa + uP_n h = (1+r)b + w_n\epsilon \tag{5}$$

$$P_n h \leq \frac{b}{1-\theta}, \tag{6}$$

where $u = (1+r) - (1-\delta)$ is the user cost of housing, which is increasing in the interest rate and the rate of depreciation of housing. With this timing assumption, allowing agents to make their housing quantity choice after realizing their ability shock, housing is now chosen in a static manner along with consumption. The only dynamic choices are that of net worth b_{t+1} and next period's location n_{t+1}.

The housing choice is then simply characterized by,

$$h_t = \min\left[\left(\frac{1-\chi}{\chi}\right)\left(\frac{1}{u_t P_{t,n_t}}\right)c_t, \frac{b_t}{(1-\theta)P_{t-1,n_t}}\right], \tag{7}$$

where the second term is simply the housing choice of a constrained agent.

3.2 Production Sector

A representative competitive final good producing firm operates in each region and uses labor and structures to produce the trade-able consumption good using Cobb-Douglas technology as follows:

$$Y_n(L_{ny}, X_{ny}) = A_n L_{ny}^{\alpha} (\tau_{ny} X_{ny})^{1-\alpha},$$

where A_n is the regional productivity, L_{ny} is the firm's labor demand, and X_{ny} is the quantity of land used by the production firm. τ_{ny} captures the degree to which land-use restrictions inhibit production in region n and so can be thought of as the effective productivity of a unit of land in region n.

The production firm's problems is then given by,

$$\max_{L_{ny}, X_{ny}} A_n L_{ny}^{\alpha} (\tau_{ny} X_{ny})^{1-\alpha} - w_n L_{ny} - q_n X_{ny},$$

where q_n is the rental cost of a unit of land.

3.3 Construction Sector

Within each region there is a continuum of construction firms that combine labor and land to produce new housing units G_n using Cobb-Douglas technology as follows,

$$G_n(L_{nh}, X_{nh}) = L_{nh}^{\xi} (\tau_{nh} X_{nh})^{1-\xi},$$

where τ_{nh} captures the degree to which land-use restrictions inhibit construction in region n. The problem of the representative construction firm is then simply given by,

$$\max_{L_{nh}, X_{nh}} P_n L_{nh}^{\xi} (\tau_n X_{nh})^{1-\xi} - w_n L_{nh} - q_n X_{nh}.$$

3.4 Equilibrium Definition

Let S denote the state space, λ be the distribution of agents over states, and λ_n be the distribution of agents in region n.

A stationary recursive competitive equilibrium consists of a value function $V(\epsilon, b, n)$, policy functions for the household $c(\epsilon, b, n), b'(\epsilon, b, n),\ n'(\epsilon, b, n),\ a(\epsilon, b, n),$ and $h(\epsilon, b, n)$, production firm choices $\{L_{ny}, X_{ny}\}_{n\in N}$, construction firm choices $\{L_{nh}, X_{nh}\}_{n\in N}$, prices $r, \{P_n\}_{n\in N}, \{q_n\}_{n\in N}, \{w_n\}_{n\in N}$, and a stationary measure λ, such that:

- given prices $r, \{P_n\}_{n\in N}$, and $\{w_n\}_{n\in N}$, the household policy functions solve the household's problem, and V is the associated value function;
- given $\{q_n\}_{n\in N}$, and $\{w_n\}_{n\in N}$, the production firm in each region chooses its inputs optimally, i.e.
  1. $w_n = \alpha A_n L_{ny}^{\alpha-1}(\tau_n X_{ny})^{1-\alpha}$;
  2. $q_n = (1-\alpha)A_n L_{ny}^{\alpha}(\tau_n X_{ny})^{-\alpha}$;
- given $\{P_n\}_{n\in N}$, $\{q_n\}_{n\in N}$ and $\{w_n\}_{n\in N}$, the construction sector chooses its inputs optimally, i.e.
  1. $w_n = \xi P_n L_{nh}^{\xi-1}(\tau_n X_{nh})^{1-\xi}$;
  2. $q_n = (1-\xi)P_n L_{nh}^{\xi}(\tau_n X_{nh})^{-\xi}$;
- all available land in each region $n \in N$ is utilized, i.e. $X_{ny} + X_{nh} = X_n$;
- the labor market in each region $n \in N$ clears:

$$L_n = \int_S \epsilon \, d\lambda_n;$$

- the housing market in each region $n \in N$ clears:

$$H_n^* = \int_S h(.) \, d\lambda_n;$$

- the law of motion for housing in each region $n \in N$ satisfies:

$$H_n' = H_n(1-\delta) + G_n,$$

  since we are considering a stationary equilibrium we have $H_n' = H_n = H_n^*$ and so $G_n = \delta H_n^*$, which states that the quantity of newly constructed structures is equal to the amount of depreciated strcutures from the previous period;
- the mass of agents migrating out of a region is equal to the mass of agents migrating into a region, so that the popuation in each region $n \in N$ is constant;
- the invariant probability distribution λ satisfies :

$$\lambda = \int_S Q((\epsilon, b, n), \mathcal{E} \times \mathcal{B} \times \mathcal{N}) d\lambda,$$

where $Q(.)$ is the transition function defined by,

$$Q((\epsilon, b, n), \mathcal{E} \times \mathcal{B} \times \mathcal{N}) = \mathbb{1}[b'(\epsilon, b, n) \in \mathcal{B}] \ \mathbb{1}[n'(\epsilon, b, n) \in \mathcal{N}] \sum_{\epsilon \in \mathcal{E}} \pi(\epsilon', \epsilon).$$

3.5 Identifying Land-Use Restrictions and Regional Productivity

The specification of the production and construction sectors, in the spirit of Herkenhoff, Ohanian, and Prescott (2018), allows me to separately identify land-use restrictions and regional productivities as functions of observables. This result relies on three key assumptions.

Firstly, I assume Cobb-Douglas technology using labor and land in both sectors. While the Cobb-Douglas technology is standard and well studied in this literature, I abstract from capital in production function in order to preserve the tractability of the model. Secondly, I require a mapping between the level restrictions in each sector. Since evidence from Gyourko, Saiz, and Summers (2008) suggests that residential land-use restrictions are strongly correlated with commercial land-use restrictions, I simply impose symmetric land-use restrictions so that $\tau_{nh} = \tau_{ny}$, henceforth τ_n. Lastly, I assume regions do not differ in the productivity of their construction sectors so that differences in TFP A_n only turn up in the goods producing sector.

Given this specification, the optimality conditions of the production sector and construction sector, coupled with the market clearing conditions, allow me to derive the following expressions for regional land-use restrictions (8) and regional productivity (9). A detailed description of the derivation is left to Appendix A.

$$\tau_n \quad = \quad \frac{1}{X_n(1-\xi)} \left[\frac{w_n}{\xi P_n} \right]^{\frac{\xi}{1-\xi}} \left[(1-\xi)\delta H_n + \frac{(1-\alpha)Y_n}{P_n} \right], \tag{8}$$

$$A_n = \frac{Y_n}{L_{ny}^{\alpha} (\tau_n X_{ny})^{1-\alpha}}. \tag{9}$$

Equation (8) pins down land-use restrictions as a function of a region's endowment of land X_n, wage rate w_n, house price P_n, housing stock H_n, and output Y_n, all of which are observable in the data. Similarly, L_{ny} and X_{ny} in equation(9) are simply functions of the same observables as well.

4 Model Implied Values

In this section, with the expressions for land-use restrictions and regional productivity in hand, I feed in state level data into equations 8 and 9 to back-out model-implied values across states and over time. Comparing the implied values to the existing measure of land-use restrictions, I confirm that τ matches the qualitative patterns already documented in the literature. Further, I study how land-use restrictions vary with regional incomes, house prices, and productivities, and also investigate how restrictions have evolved over time.

4.1 Data

I perform this exercise at the state level for two primary reasons. Firstly, the state is the smallest geographic unit for which representative data covering the whole U.S. is available for each of the necessary variables, going back before the year 2000. Secondly, the model and research question focuses on individuals who are forced to work and live in the same region. Consequently, I require the location of the agent's labor market to be equivalent that of her housing market, and this is not necessarily the case when considering smaller units such as cities, where an agent could work in the city but reside outside of it.

The data required come from many sources. I obtain urban land area at the state level from the US Department of Agriculture. The Census Bureau recognizes two types of urban areas, (i) Urbanized Areas (UA) of 50,000 or more people, and (ii) Urban Clusters (UC) of at least 2,500 and less than 50,000 people. While Urban land represents only a small share of the U.S. land mass, 3% in 2012, it accounts for over 80% of the total population. Due to a change in the definition and measurement of urban land, between 1997 and 2000, the values before and after this period are not directly comparable, and so my analysis will focus only on the years following 2000. Further, since the data is available only for selected years, I use linear interpolation to obtain the urban area for each state in the years of interest.

I obtain annual wage and house price data from the Census and the American Community Survey using the Integrated Public Use Microdata Series, provided by Ruggles et al. (2019). In calculation of the median wage by state, I consider only those of working age (18-65), who were employed and worked more than 26 weeks in the year in question. I construct adjusted wages, controlling for differences in educational attainment and industry concentration across states, so that all of the cross-sectional variation in wages is coming from the state fixed effect. I leave a more detailed description of the computation of adjusted wages to appendix B. The median house price is constructed using only owner-occupied single-family housing units and I deflate both wages and house prices using a national deflater.

Housing stock data comes from the Census, which uses use the most recent decennial census to form the base for the annual housing unit estimates. Building permits, estimates of non-permitted construction, mobile home shipments, and estimates of housing loss, are the used to estimate the change in the housing stock. Lastly, I use the state level GDP from the Bureau of Economic Analysis (BEA) to compute state level output shares and deflate this measure using a national deflater with base year 2017 in order to obtain real output shares.

4.2 Comparison to Available Measure of Land-Use Regulations

Thus far, I have not referenced the source of the heterogeneity in land-use restrictions, which as detailed in section 1, could result from (i) physical limitations in a region's endowment of land, or (ii) imposed land-use regulations which inhibit the productivity of a unit of land. However, having now derived the model-implied measure of land-use restrictions τ, I compare it to a measure of imposed land-use regulations, in order to better understand the source of the variation in restrictions.

The best available and most comprehensive measure of land-use regulations in the literature is the Wharton Residential Land-Use Regulation Index (WRLURI), which is based on surveys sent to municipalities across the country. The survey asks local municipalities 15 questions regarding such items as local political pressure, density restrictions, local zoning approval, approval delays, and exactions. I aggregate the WRLURI, which is available at the MSA level, up to the state. Figure 1 depicts this state level WRLURI against the model-implied measure of restrictions τ, for 2010. Note, a higher WRLURI value indicates tighter regulations, while a lower value of τ represents higher restrictions.

Clearly, I observe a strong relationship between both measures, suggesting that a significant share of the differences in land-use restrictions may be due to imposed regulations.

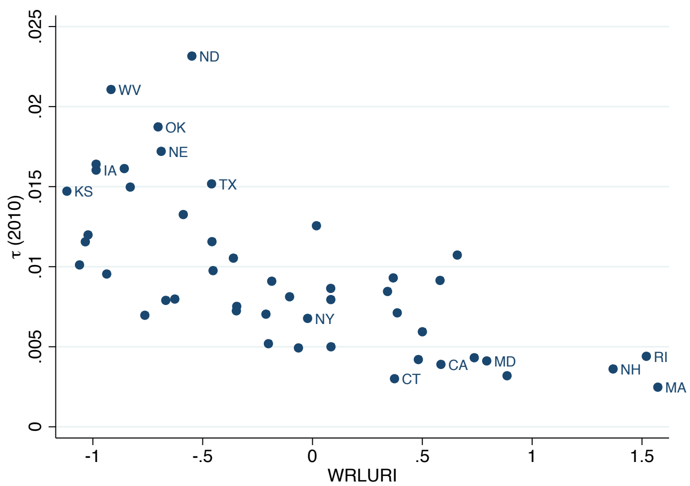


Figure 1: Model-implied restrictions vs. WRLURI

Ranking states by the severity of restrictions, as measured by τ, and the severity of regulations, as measured by WRLURI, we have a rank-rank correlation of 0.78. Given the inability to disentangle regulations from the physical limitations of land, I cannot directly quantify the extent to which regulations explain differences in restrictions. However, the strong correlation between land-use restrictions and regulations, which are directly controlled and hence changeable by policy, suggests that studying the implications of changing the level of restrictions is a worthy endeavor.

Notably, the states comprising the North East as well as California stand out as being the most restricted while the Central states are the least restricted. This is consistent with findings from numerous empirical studies such as Glaeser and Gyourko (2018), Whittemore (2012), and Levine (1999), which focus on cities, particularly those in California and the North East. Further, the model-implied measure τ, is also consistent with measures derived by Hsieh and Moretti (2019) and Herkenhoff et al. (2018), who perform a similar exercise.

4.3 Incomes, House Prices, and Productivities

Figure 2 plots land-use restrictions against state-level adjusted wage incomes and house prices, as at 2014. We see that the highest income states, which consist of California and those in North East, also have the highest levels of restrictions. Similarly, the states with

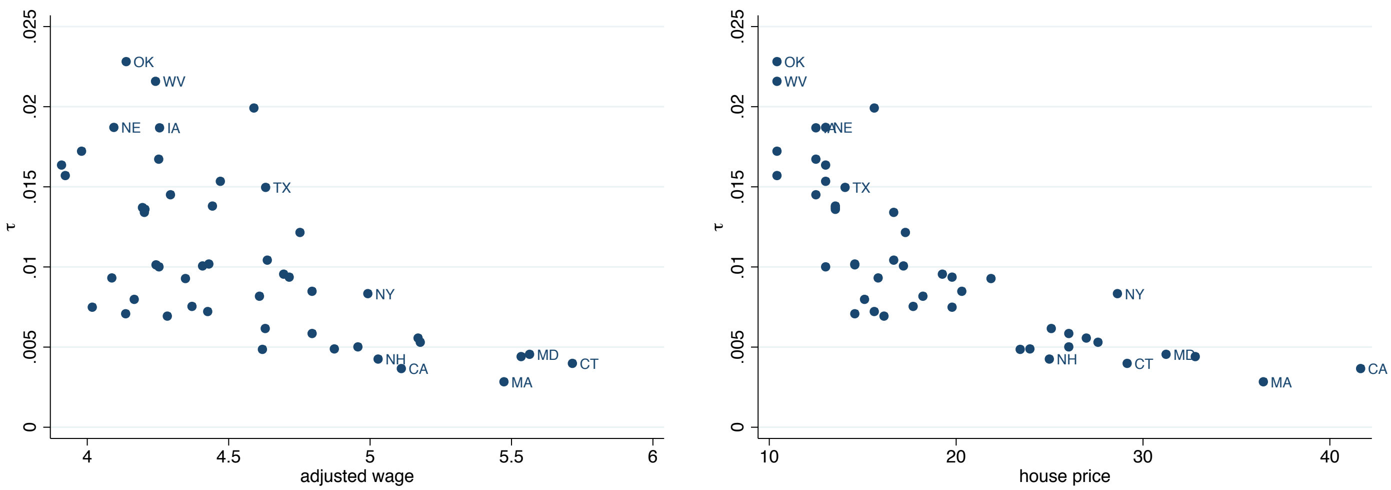


Figure 2: Land-use restrictions vs. wage income and house prices.

the highest house values also seem to have the highest levels of restrictions. This pattern once again highlights the mechanism of interest in this paper, where differences in land-use restrictions exacerbate differences in house prices, and prevent the re-allocation of labor due to the proportionately higher costs of living in high income in regions.

In addition to land-use restrictions, I also back-out a model-implied measure of regional productivity by feeding in data to equation 9. Figure 3 plots land-use restrictions vs. productivity for 2014. We see here an almost linear relationship where the most productive states, once again consisting of California and those in North East, also have the tightest restrictions.

While both, restrictions and productivity have implications for regional prices, housing density, and output, the ability to separately identify them relies on their specific effect on a given variable. To better understand this consider the following example. Suppose there are two identical regions. If region 1 were to experience a tightening of restrictions this would result in a fall in the housing supply in region 1, a subsequent increase in house prices, and so a shift of some of the population out of region 1 and into region 2, due to the increase in housing costs. This population shift will in turn put upward pressure on wages in region 1 and downward pressure on wages in region 2. Consequently, we finally have higher wages and house prices in region 1, along with a lower density of housing (since the housing stock has fallen while the amount of land remains unchanged).

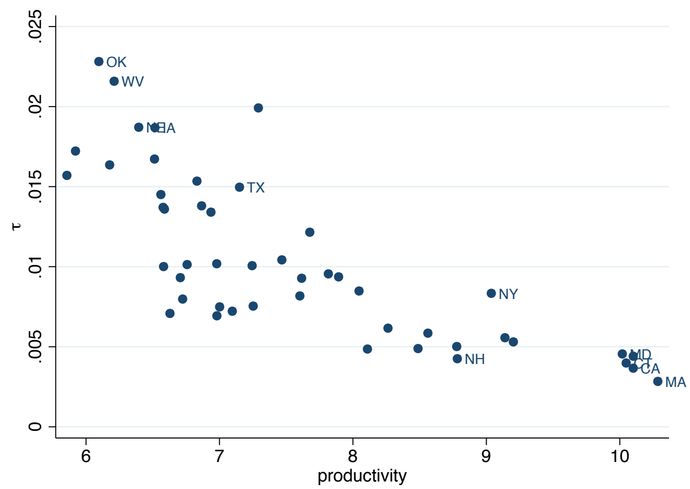


Figure 3: τ and productivity in 2014

Now consider what would happen if region 1 experienced an increase in productivity rather than a tightening of restrictions. Region 1 would enjoy an increase in wages, which increases the region's population share as people migrate to take advantage of the higher income. The higher population would increase the demand for housing which in turn would raise the house price in the region. Subsequently, as in the previous case, region 1 would have a higher wage and house price, however now region 1 would have an increase in housing density. Thus, the identification crucially depends on the cross-sectional variation in wages, house prices, and densities.

4.4 Land-Use Restrictions Over Time

The fact that the identifying equations express restrictions as a function of observable data also enables me to analyze how restrictions have evolved over time. Figure 4, plots the change in restrictions from 2000 to 2014, against their level in the year 2000. Firstly, we notice that most states have experienced a tightening of restrictions as evidenced by the negative growth rate. Secondly, we see that the states experiencing the largest increase in restrictions (fall in τ), were those that were most regulated to begin with, namely California and those in the North East. To the extent that we believe that physical limitations of regional land endowments have not changed in this period, this decline in restrictions

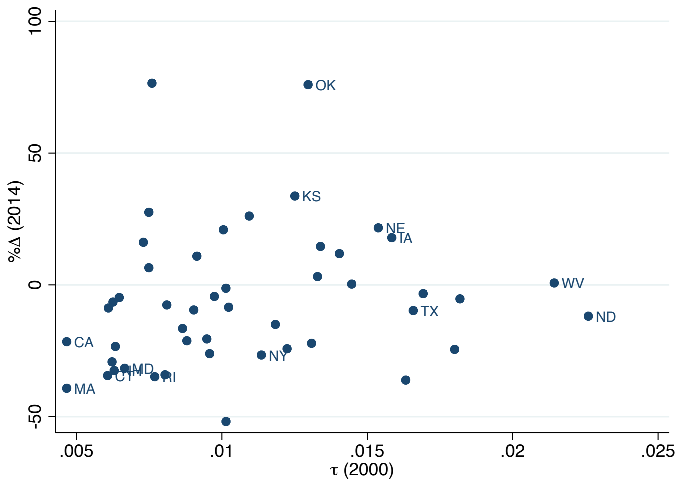


Figure 4: Change in restrictions over time

is indicative of an increase in imposed regulations in these regions. Consequently, when conducting the counterfactual exercise in the following section, I will lower the level of restrictions in the high income region to its own level in 2000. Thus, this counterfactual maps to a policy reforming land-use restrictions, rather than to removing innate limitations in a region's land, which is less feasible.

5 Quantitative Exercise

Having thus far exploited the firms' side of the model to obtain measures of land-use restrictions and productivity, I now analyze the full model focusing on the household decision and stationary distribution. Following a description of the calibration strategy, I study the baseline distribution of the model and the key mechanisms driving the results. Lastly, I perform the counterfactual exercise of lowering land-use restrictions in the high productivity region and investigate the implications of land-use restrictions.

5.1 Calibration

The model period is one year. Given the heterogeneity in the model, the number of regions is limited to 2 in order to preserve tractability. I set $N = \{0, 1\}$, where region 0 corresponds to California and region 1 corresponds to the rest of the United States. I focus on California for 2 reasons. Firstly, as observed in the previous section California epitomizes the 'good' regions of the U.S., having high incomes, house prices, and productivity, while also having very stringent land-use restrictions. Secondly, given the severity of the housing affordability crisis in California and the much publicized political debate around land-use restrictions in the state, the implications of loosening land-use restrictions in California are particularly relevant for policy.

The baseline model is calibrated to match the data as at 2014. The parameters of the model can be divided into two groups. The first consists of parameters that are externally calibrated, taking values established in the literature or given by data. The second consists of those that are internally calibrated in order to match the moments of the model.

5.1.1 Externally Calibrated Parameters

The logarithm of the idiosyncratic ability process faced by agents follows an AR(1) process,

$$\log \epsilon_t = \rho \log \epsilon_{t-1} + e_t,$$

where $|\rho| < 1$ and $e_t \sim \mathcal{N}(0, \sigma_e^2)$. The AR(1) process is covariance-stationary with mean 0 and $\mathrm{var}(\log \epsilon_t) = \frac{\sigma_e^2}{1-\rho^2}$. I approximate discretely this continuous process through a finite-state Markov process following Tauchen (1986). Table 1 below presents the list of externally calibrated parameters along with their values.

Notably, the land share in the construction sector is chosen to be larger than the corresponding share in the production sector. The land share in the production sector is chosen based on findings from Davis and Heathcote (2007) who show that land accounted for 35-45% of the value of the aggregate housing stock between 1975 and 2006. The Cobb-Douglas utility function and preference weight on consumption is chosen to be consistent with papers such as Piazzesi et al. (2007) that show that the non-housing expenditure share does not vary significantly over time and lies approximately between 0.8 and 0.86 over their period of interest. The interest rate is a parameter since I do not close the asset market and abstract from capital in the firm sectors. Consequently, agents can borrow and save at the exogenously given interest rate. The remainder of the economy-wide parameters are standard in the literature.

Table 1: Externally calibrated parameters

Parameters	Value	Description
Economy-wide parameters		
β	0.95	discount factor
α	0.9	labor share in production sector
ξ	0.6	labor share in construction sector
χ	0.84	preference weight on consumption
δ	0.02	depreciation rate
θ	0.8	maximum loan-to-value ratio
ρ	0.9	persistance of log ability process
r	0.02	exogenous interest rate
Regional parameters		
X_n	0.073 , 0.927	land share
τ_n	.0037 , .0099	land-use restrictions
A_n	10.1 , 7.57	regional productivity

The regional measures of land-use restrictions and productivity come from identifying equations (8) and (9), while the regional endowment of urban land comes from USDA data as previously detailed in section 4.1.

5.1.2 Internally Calibrated Parameters

The list of internally calibrated parameters and their targeted moments are summarized in table 2 below. The regional amenity is value is chosen to match the population share in each region. The moving cost is chosen to match the migration rate in California. Note, since I am studying the steady-state of the model, the flow of agents into each region is equivalent to the outflow of agents from each region. However, given these two flows are not equivalent in the data, I compute the average between the in-migration rate and the out-migration in California and use this as the target moment. I focus only on migration to and from other states, excluding international migration. The subsistence level of housing

Table 2: Internally calibrated parameters

Parameter	Value	Description	Target
z_n	0.39 , 0.1	amenity value	population share
κ	7.7	moving cost	migration rate
$\bar{h}$	0.48	subsistence level of housing	median $\frac{housevalue}{income}$
σ_e	0.22	s.d. of log ability process	inc p75/p25

in the utility function is chosen to match the ratio of the median home price to median income, while the variance of the logarithm of the ability process is chosen to match the ratio between the 75th and 25th percentiles of the income distribution.

5.2 Baseline Model

Calibrating California to have higher productivity, tighter land-use restrictions, and a higher amenity value, results in the following. The higher productivity in California contributes to a higher wage in the region. This higher wage coupled with the higher amenity value raises the demand for housing, given the attractiveness of the region. However, the higher level of land-use restrictions reduces the capacity of the production sector and more importantly hinders the construction sector's ability to produce houses. Thus, the limited supply of housing leads to a significantly higher house price. The prohibitively high cost of housing keeps out individuals who would have otherwise moved to California, and this lower labor supply puts further upward pressure on the region's wages. This results in the equilibrium prices outlined in the table below.

Table 3: Regional Prices

Price	Region 0	Region 1
w	4.43	4.16
P	63.78	24.32

Household Behavior. In order to clearly understand the behavior of the households in the model I simulate the sample path of a household for a given sequence of ability draws.

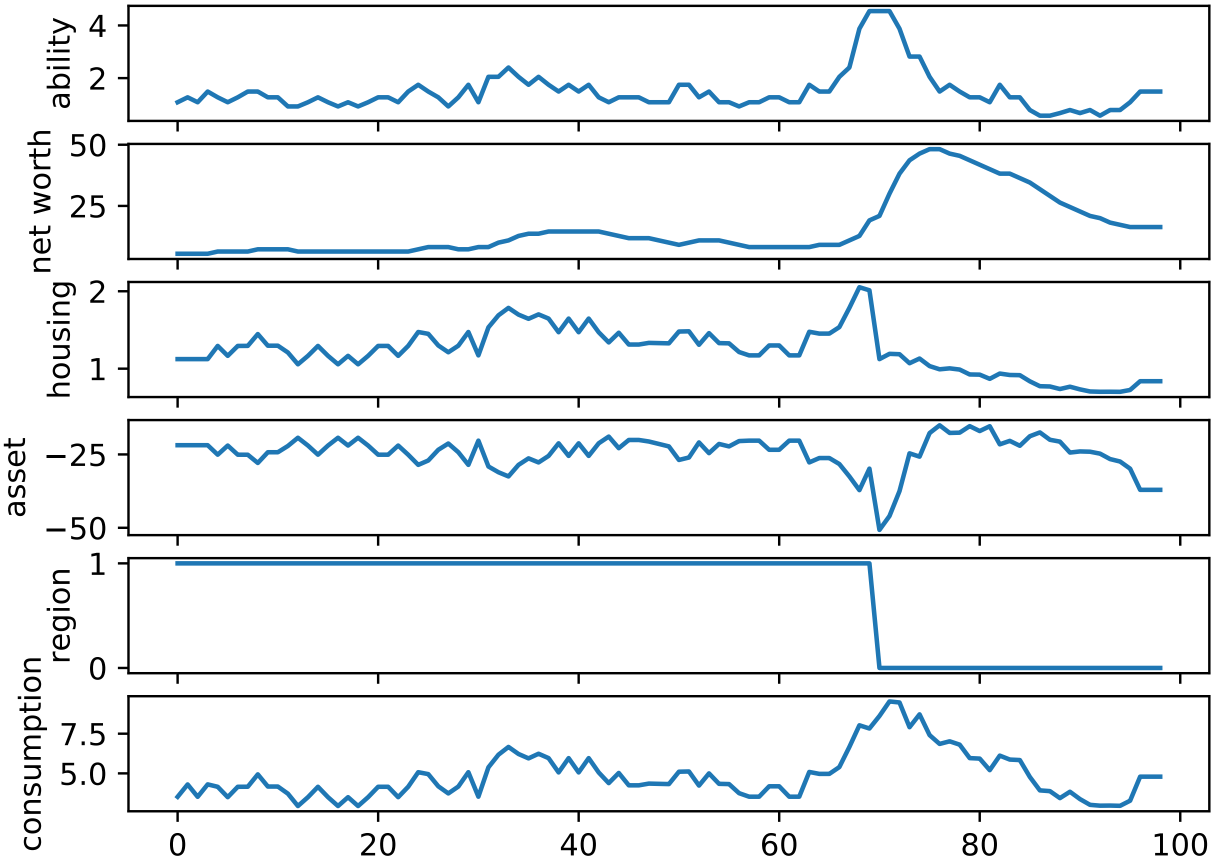


Figure 5: Household Simulation. First 2000 periods have been dropped.

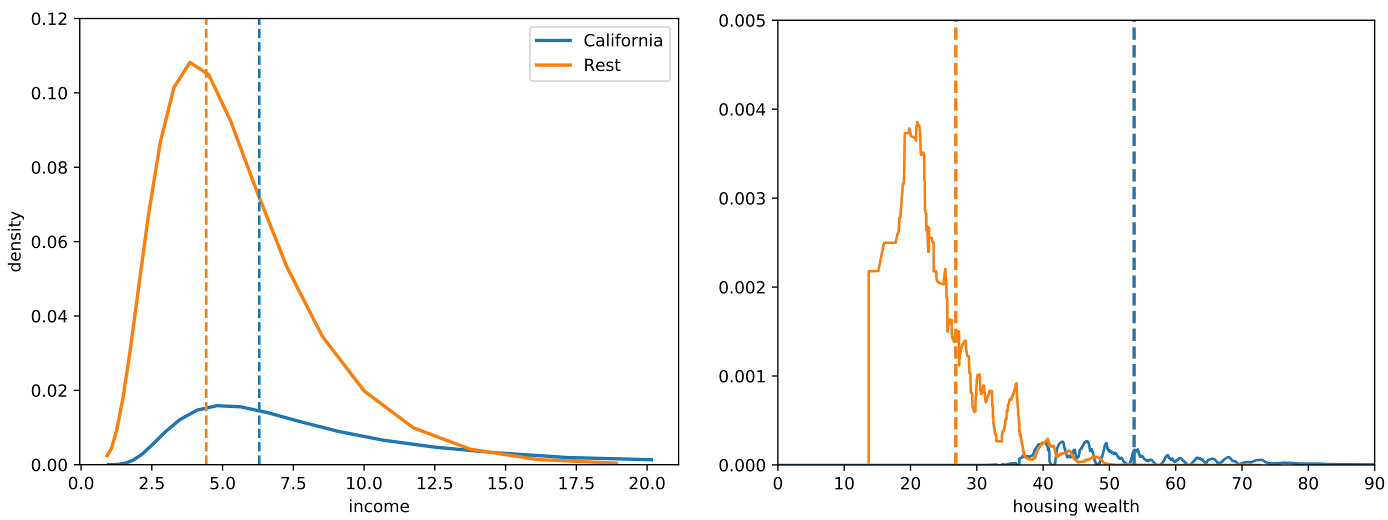
The simulation is summarized in figure 5 below. The agent is stuck in region 1 at first, while she has low ability and net worth. Over time as she receives better ability draws she starts paying off her debt and increasing the quantity of housing she owns, thereby increasing her net worth. Contingent on sufficient net worth and a large enough ability draw, she uses her higher income and net worth to pay the moving cost, cover the higher housing cost, and move to region 0.

The equilibrium prices in each region, given in table 3, clearly illustrate the trade-off, between a higher wage and a higher house price, faced by the agent as she contemplates moving to region 0. When the agent moves, she sells her larger but less valuable house in region 1 and is forced to buy a much smaller house in region 0, given the significantly higher price of housing. In order to absorb this housing cost and the cost of moving, she borrows against her house. This is seen in the sharp drop in her asset position. The agent stays in region 0 while her draws are sufficiently high and takes on debt to smooth against low ability draws. Eventually, as she receives a sequence of low ability draws and runs down her net worth she will move back to region 1.

As evident from this simulation there are two sources of sorting in this environment. Firstly, since an agent's income is given by the product of her ability and the regional wage, ϵw_n, the increase in income from moving from region 1 to 0 is higher for higher ability agents. Thus, agents sorts on ability. Secondly, since only high net worth agents are able to absorb the moving cost and afford the significantly higher house price in region 0, agents also sort on net worth.

Stationary Distribution. Studying the stationary distribution of the economy, depicted in figure 6, the income and house value gaps between the two regions are clearly observable. The higher income in California comes from the fact that the region has both, a higher equilibrium wage and a population that is on average more skilled than those in the rest of the country. Consistent with the data we observe a proportionately larger house value gap. Note, the higher house values in California come purely from the higher per unit house price in the region, as in fact the average house in California is smaller than the average house in the rest of the country.

Figure 6: Income and Housing Wealth Distribution



Notes: The dotted lines plot the mean income (left panel) and mean house value (right panel) in each region.

While I have not explicitly targeted the income and house value gaps between regions, the model generated gaps match the differentials observed in the data well. This is evident from table 4. Notably, the model captures the fact the house value gap is significantly larger than the income gap. However, the model slightly under estimates the house value gap, while slightly over estimating the income gap.

Table 4: Untargeted Moments

	Data	Model
income gap	1.13	1.42
house value gap	2.19	2.00

5.3 Counterfactual Exercise

In order to analyze the quantitative implications of land-use restrictions, I now study the key counterfactual exercise of this paper. I lower the level of restrictions in California (raise τ_0) to it's level in 2000, holding all else including the amenity values, productivities, and the level of restrictions in the rest of the U.S., at their 2014 levels. The values of τ are outlined in table 5, and this exercise represents a loosening of restrictions in California by 27%.

Table 5: Change in restrictions

	California	Rest
baseline τ	.0037	.0099
counterfactual τ	.0047	.0099

The lower level of restrictions in California have two direct effects. First, local wages rise, since the marginal product of labor in increasing in τ. Second, the increased usability of land increases the housing stock in the region and reduces house prices. This results in a large movement of people into California from the rest of the country. As evidenced by table 6, the rest of the U.S. also enjoys a higher wage, due to a lower labor supply, and benefits from lower house prices, resulting from a decrease in the demand for housing.

Table 6: Implications of lower restrictions

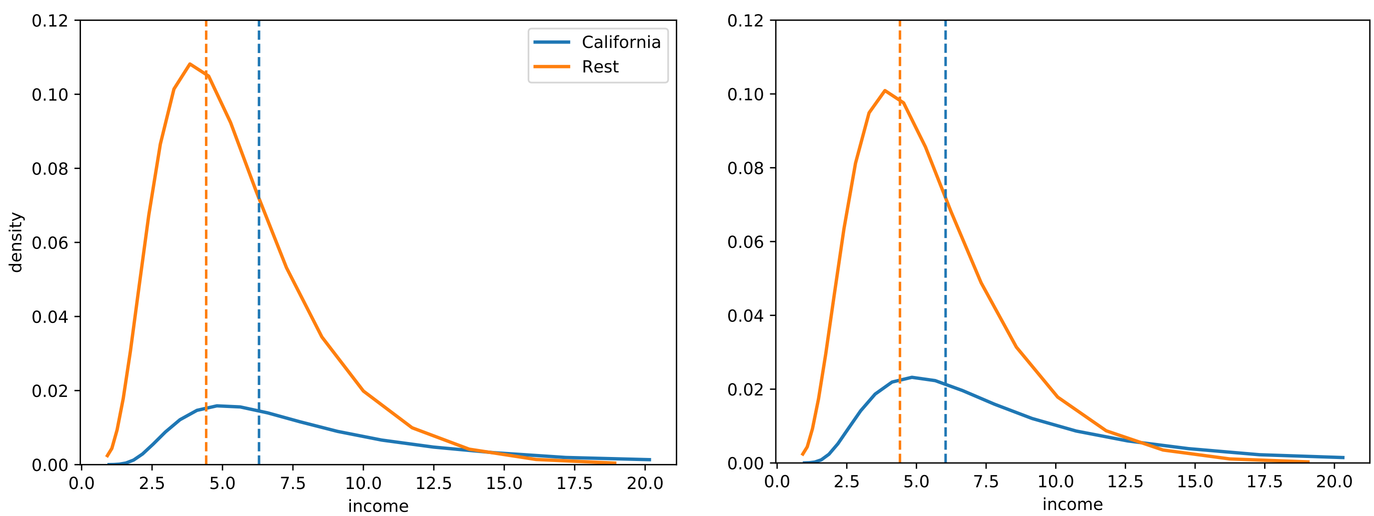
Variable	Baseline		Counterfactual	
	Cal	Rest	Cal	Rest
w	4.433	4.162	4.466	4.190
P	63.78	24.32	62.37	23.82
mass	12.9%	87.1%	18.9%	81.1%
regional output	1.154	3.940	1.367	3.770
aggregate output	5.094		5.127	

Figure 7 compares the income distribution in the counterfactual model to that in the baseline model. While the population change in each region is clearly evident, the skill composition of workers also changes. The new residents of California are on average less skilled than those previously in the region. This further lowers the average income in the region. As evidence by table 7, the combination of converging wages and a lower skill composition in California, lead to a fall in the income gap between the regions by 3.69%, when compared to the baseline.

Figure 7: Income Distribution

(a) Baseline

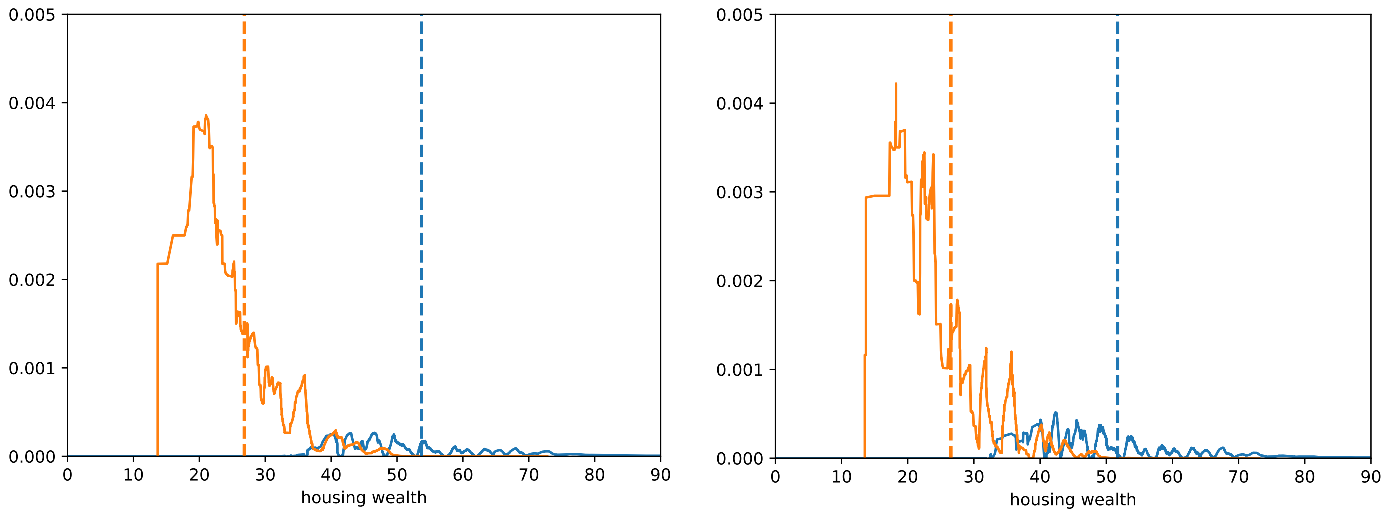
(b) Counterfactual



Notes: The dotted lines plot the mean income.

Figure 8: Housing Wealth Distribution

(a) Baseline (b) Counterfactual



Notes: The dotted lines plot the mean housing wealth.

Figure 8 compares the housing wealth distribution in both cases. We see that the new entrants to California on average own significantly smaller houses than those previously in the region. Thus, the inflow of agents into California also reduces the mean level of housing wealth in the region and consequently, lowers the housing wealth gap between the regions.

Table 7: Regional Gaps

	Baseline	Counterfactual	% change
population in Cal	12.9%	18.8%	↑ 45.0%
income gap	1.42	1.37	↓ 3.69%
housing wealth gap	2.00	1.95	↓ 2.66%

As noted in table 6, the counterfactual economy also enjoys higher aggregate output. This comes about via two channels. The first comes through the reallocation of labor, as a larger share of the population now lives and works in the more productive region. The second comes from the fact that lowering restrictions effectively increases a factor of production, subsequently enabling the consumption goods producing firm to produce more from a given unit of land while also enjoy a lower rental cost of land.

6 Discussion and Extensions

Borrowing Constraints

Moving Costs

Residential Land-Use vs. Commercial Land-Use

Agglomeration

Congestion

Transition Path

6.1 Policy Relevance

7 Conclusion

References

Robert J Barro and Xavier S Martin. Convergence across states and regions. *Brookings Papers on Economic Activity*, 1:107 – 182, 1991.

Truman Bewley. The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory*, 16:252 – 292, 1977.

Adrien Bilal and Esteban Rossi-Hansberg. Location as an asset. *Working Paper*, 2018.

Olivier Jean Blanchard and Lawrence F Katz. Regional evolutions. *Brookings Papers of Economic Activity*, 23:1 – 76, 1992.

Morris Davis and Jonathan Heathcote. The price and quantity of residential land in the united states. *Journal of Monetary Economics*, 54:2595 – 2620, 2007.

Rebecca Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106:479 – 524, 2016.

Fabian Eckert and Tatjana Kleineberg. Can we save the american dream? a dynamic general equilibrium analysis of the effects of school financing on local opportunities. *Working Paper*.

Jan Eeckhout, Roberto Pinheiro, and Kurt Schmidheiny. Spatial sorting. *Journal of Political Economy*, 122:554 – 620, 2014.

Jack Favilukis, Pierre Mabille, and Stijn Van Nieuwerburgh. Affordable housing and city welfare. *Working Paper*, 2018.

Peter Ganong and Daniel Shoag. Why has regional income convergence in the u.s. declined? *Journal of Urban Economics*, 102:76 – 90, 2017.

Edward Glaeser and Joseph Gyourko. The economic implications of housing supply. *Journal of Economic Perspectives*, 32:3 – 30, 2018.

Joseph Gyourko, Albert Saiz, and Anita A. Summers. A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies*, 45, 2008.

Jonathan Halket and Santhanagopalan Vasudev. Saving up or settling down: Home ownership over the life cycle. *Review of Economic Dynamics*, 17:345 – 366, 2014.

Kyle F. Herkenhoff, Lee E. Ohanian, and Edward C. Prescott. Tarnishing the golden and empire states: Land-use restrictions and the u.s. economic slowdown. *Journal of Monetary Economics*, 93, 2018.

Chang-Tai Hsieh and Enrico Moretti. Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11:1–39, 2019.

Kristoffer Jackson. Do land use regulations stifle residential development? evidence from california cities. *Journal of Urban Economics*, 91:45 – 56, 2016.

Greg Kaplan and Sam Schulhofer-Wohl. Understanding the long-run decline in interstate migration. *International Economic Review*, 58:57 – 94, 2017.

John Kennan and James R. Walker. The effect of expected income on individual migration decisions. *Econometrica*, 79:211 – 251, 2011.

Ned Levine. The effects of local growth controls on regional housing production and population redistribution in california. *Urban Studies*, 36:2047 – 2068, 1999.

Raven Molloy, Christopher L Smith, and Abigail Wozniak. Internal migration in the united states. *Journal of Economic Perspectives*, 25:1 – 24, 2011.

Stijn Van Nieuwerburgh and Pierre-Olivier Weill. Why has house price dispersion gone up? *Review of Economic Studies*, 77:1567 – 1606, 2010.

Andrii Parkhomenko. Housing supply regulation: Local causes and aggregate implications. *Working Paper*, 2018.

Monika Piazzesi, Martin Schneider, and Selale Tuzel. Housing, consumption and asset pricing. *Journal of Financial Economics*, 83:531 – 569, 2007.

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

Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. Ipums usa: Version 9.0 [dataset]. *Minneapolis, MN*, 2019.

Albert Saiz. The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125:1253 – 1296, 2010.

George Tauchen. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, 20(2):177–181, 1986.

Matthew A. Turner, Andrew Haughwout, and Wilbert van der Klaauw. Land use regulation and welfare. *Econometrica*, 82:1341 – 1403, 2014.

Andrew H. Whittemore. Zoning los angeles: a brief history of four regimes. *Planning Perspectives*, 27:393–415, 2012.

Appendices

A Derivation of identifying equations

The derivation of identifying equations (8) and (9) is as follows.

- From the production firm's first order conditions obtain $L_{ny} = \frac{\alpha Y_n}{w_n}$ and $X_{ny} = \frac{(1-\alpha)Y_n}{q_n}$.
- From the construction firm's first order conditions obtain $L_{nh} = \frac{\xi P_n h_n}{w_n}$ and $X_{nh} = \frac{(1-\xi)P_n h_n}{q_n}$.
- Using the Cobb-Douglas function for the construction technology $G_n = L_{nh}^{\xi}(\tau_n X_{nh})^{1-\xi}$, solve for τ_n to obtain,

$$\tau_n = \left[\frac{G_n}{(L_{nh})^{\xi}}\right]^{\frac{1}{1-\xi}} \frac{1}{X_{nh}}.$$

- Plugging in $X_{nh} = \frac{(1-\xi)P_n G_n}{q_n}$, obtain

$$\begin{aligned} \tau_n &= \left[\frac{G_n}{(L_{nh})^{\xi}}\right]^{\frac{1}{1-\xi}} \frac{q_n}{(1-\xi)P_n G_n}, \\ &= \frac{1}{1-\xi}\left[\frac{G_n}{L_{nh}}\right]^{\frac{\xi}{1-\xi}} \frac{q_n}{P_n}. \end{aligned}$$

- Adding the two expressions $X_{ny} = \frac{(1-\alpha)Y_n}{q_n}$ and $X_{nh} = \frac{(1-\xi)P_n G_n}{q_n}$, and solving for q_n, obtain $q_n = \frac{1}{X_n}[(1-\xi)P_n G_n + (1-\alpha)Y_n]$. Plugging this into the above expression for τ_n, obtain

$$\begin{aligned} \tau_n &= \frac{1}{X_n(1-\xi)}\left[\frac{G_n}{L_{nh}}\right]^{\frac{\xi}{1-\xi}} \frac{(1-\xi)P_n G_n + (1-\alpha)Y_n}{P_n}, \\ &= \frac{1}{X_n(1-\xi)}\left[\frac{G_n}{L_{nh}}\right]^{\frac{\xi}{1-\xi}} \left[(1-\xi)G_n + \frac{(1-\alpha)Y_n}{P_n}\right]. \end{aligned}$$

- Substituting for L_{nh}, obtain

$$\tau_n = \frac{1}{X_n(1-\xi)}\left[\frac{w_n}{\xi P_n}\right]^{\frac{\xi}{1-\xi}} \left[(1-\xi)G_n + \frac{(1-\alpha)Y_n}{P_n}\right].$$

- Lastly, using the equilibrium condition that $G_n = \delta H_n$ we have equation (8),

$$\tau_n \;=\; \frac{1}{X_n(1-\xi)}\left[\frac{w_n}{\xi P_n}\right]^{\frac{\xi}{1-\xi}}\left[(1-\xi)\delta H_n + \frac{(1-\alpha)Y_n}{P_n}\right].$$

Consequently, we have τ_n as a function of X_n, w_n, P_n, H_n, and Y_n, all of which are observable. Using data on these variables, I can back out a time series of land-use restrictions for each region. With τ_n in hand, I invert the production technology to obtain the model implied productivity A_n in each region,

$$A_n = \frac{Y_n}{L_{ny}^{\alpha}(\tau_n X_{ny})^{1-\alpha}},$$

where L_{ny}, X_{ny}, τ_n are functions of the observables described above.

B Construction of adjusted wages

From the Census and ACS I obtain microdata on wages as well the individual's level of education, industry of employment, and state of residence. With this is hand, I run the following regression for each year,

$$\log(w_i) = \alpha + \beta_1\ educ_i + \beta_2\ ind_i + \beta_3\ state_i, \tag{10}$$

where $educ_i$ is a dummy variable for individual $i's$ educational attainment which can take three values depending on whether the individual,

i did not complete high school,

ii completed high school and some college, or

iii completed at least 4 years of college.

ind_i is a dummy variable for individual $i's$ industry of employment which can take seven values, and $state_i$ is a dummy variable for the individual's state of residence. Note, all coefficient's are statistically significant at the 1% level. I then use the coefficients to fix the educational attainment and industry composition in all states. That is, I compute

$$\hat{educ} = \sum_{j=1}^{3} \beta_{1,j}\ educ\_share_j,$$

where j indexes the education bin, and $educ\_share_j$ is the share of the U.S. population that have the level of educational attainment associated with education bin j. Similarly, I compute

$$\hat{ind} = \sum_{j=1}^{7} \beta_{2,j}\ ind\_share_j,$$

where j indexes the industry bin, and ind_share_j is the share of the U.S. population that are employed in the industry associated with industry bin j. I then compute the state level adjusted log wage as follows,

$$\log(w_s) = \alpha + \hat{educ} + \hat{ind} + \beta_3\ state_i,$$

and finally take the exponential of $\log(w_s)$ to obtain the adjusted state level wage. Consequently, all of the cross-sectional variation in adjusted state wages are coming from the state fixed effect. I repeat this exercise for each year to obtain a panel of adjusted state wages.