The Effect of Land Use Regulation on Housing Prices: Theory and Evidence from California *

Desen Lin[†] Susan Wachter[‡]

April 4, 2019

Abstract

Land use regulation may affect housing prices through housing supply and demand, but the empirical literature conflates both effects and finds wide variation in the estimated impact. We disentangle three channels through which regulation may affect housing prices: the production channel through housing supply, the amenity channel through housing demand, and the general equilibrium (GE) channel that captures price feedback effects on location choice. We develop a GE model with households' choices on consumption and location and with housing developers' choice on housing production. Our theoretical model delivers a closed-form solution to the equilibrium prices and a simple form of the estimation equations. Using property transaction-assessment data from 1993 to 2017 in California and a regulatory index compiled from the Wharton Residential Land Use Survey (Gyourko, Saiz and Summers, 2008), we structurally estimate and disentangle the supply and demand-side effects. We find that the regulatory impact on housing prices through the production channel is much stronger than the amenity channel (4.38% vs 0.32% if referenced to the average city in California) and is heterogeneous across cities. The relationship still holds, even when the GE effects are included in the two channels (3.24% vs 0.27%). The total effect of regulation will be 4 times larger, if referenced to the average regulation in the US. Our estimations point out the key roles of structural characteristics of housing and macroeconomic conditions in the prediction of housing prices. Estimations without quality adjustment underestimate land regulation's impact on prices. Additionally, we examine the within-MSA regulatory interdependence and find significant and positive spillover effects of regulation on housing prices. Estimations without spillover consideration underestimate the regulatory impact on prices.

Keywords: housing prices, land use regulation, general equilibrium, spillover effect, California
JEL: R10, R13, R31, R52, R58

* We thank Joe Gyourko, David E Rappoport for constructive advice and comments. Susan Wachter acknowledges financial support from the Zell-Lurie Real Estate Center at The Wharton School of the University of Pennsylvania. We thank the Zillow Group for sharing data. All errors are our own.
† Department of Economics, University of Pennsylvania. Email: desenlin@sas.upenn.edu
‡ Department of Real Estate, The Wharton School, University of Pennsylvania. Email: wachter@wharton.upenn.edu

1. Introduction

Since the US Supreme case of Euclid v. Ambler (1926), land use regulation has been central to the debates of housing affordability and economic growth.[1] While land use regulation makes housing less affordable by tightening supply constraints (Glaeser, Gyourko and Saks, 2005a, 2005b; Saiz, 2010), it may also increase environmental amenities and thereby raise housing demand (Hamilton, 1975; Fischel, 1990; Gyourko and Molloy, 2015). Land use regulation may even go beyond localities and exert impacts on multiple jurisdictions (Pollakowski and Wachter, 1990). The empirical literature finds a wide variation in the estimated impacts of land use regulation on housing prices (Quigley and Rosenthal, 2005). In part, the variation in estimated impacts is due to conflating supply and demand-side effects.

We contribute to the literature on land use regulation's impact on housing prices in several ways. First, we base our empirical analysis in a general equilibrium framework with household mobility across geographical markets. We include households' decisions over consumption and location together with developers' housing production decisions. We incorporate multiple transmission channels of regulation on prices that result from this general equilibrium framework. Land use regulation and per capita income are key pricing factors, with the quadratic and interactive effects micro-founded in the model. Our theoretical model delivers a closed-form solution to the equilibrium prices and a simple form of the estimation equations.

We characterize and disentangle three channels through which land use regulation may affect housing prices. The first channel of the regulatory impact goes through housing supply and the effect is local. We call it the *production channel*, because regulation increases the cost of housing supply and the local housing prices. The second channel goes through housing demand and the effect is also local. We call this the *amenity channel*, because regulation boosts amenity values and the housing demand, leading to an increase in the local housing prices. There is a third channel related to the household location choice. We call it the *general equilibrium (GE) channel*, because it captures the price feedback by taking household mobility into account. Tighter regulation that makes housing more expensive will drive housing demand to the neighboring cities. We also calculate *net production and amenity channels* that incorporate the GE effects.

Our empirical analysis is based on structural estimations using Generalized Method of Moments (Hansen, 1982). We use the structural estimates to quantify how different channels respond to the regulatory change by city level over the period 1993 through 2017. We base our measure of regulatory constraint using the Wharton Residential Land Use Survey (Gyourko, Saiz, and Summers, 2008). We estimate city-level land use regulation effects after controlling for individual property-based housing characteristics, metro level per capita income and national credit supply. The average marginal effect of

---

[1] The case of Village of Euclid, Ohio v. Ambler Realty Co. set the precedent of new zoning practice and served to bolster local zoning ordinances nationwide. For the details of Euclid v. Ambler, see Fluck (1986).

regulation on housing prices through all channels is about 3% in California (with the average California city as the regulatory reference point) which consists of 4.38%, 0.32% and -1.73% from the production, the amenity and the GE channels respectively (or 3.24% and 0.27% from the net production and amenity channels). The housing price responses through the amenity and the GE channels to a unit increase in regulation are small (0.32% and -1.73% respectively). If referenced to the average level of regulation in the US, the total regulatory effect will be 4 times larger. The heterogeneous regulatory impacts across MSAs are mainly driven by the response through the production channel. San Francisco, San Jose, Los Angeles and San Diego MSAs have the largest production effect, more than 50% larger than the MSA average response through the production channel (3.22%). The net production and amenity channels that incorporate the GE effects are smaller. The price responses through the net production channel of San Francisco, San Jose, Los Angeles and San Diego MSAs are 4.58%, 4.57%, 3.55% and 3.29% respectively, while the response through the net amenity channel is 0.27%. Among the studies with standardized regulatory measures, our estimates of the net production effect are larger than those of Quigley, Raphael and Rosenthal (2008) on the regulatory impact on prices in San Francisco Area (1.2%-2.2% in OLS and 3.8%-5.3% in IV estimations).

Our empirical estimations point out the key roles of structural characteristics of housing and macroeconomic conditions in the prediction of housing prices. Instead of hedonic price indices, we use property transaction data together with housing characteristics in the empirical analysis for housing quality adjustment. Our method has smaller standard errors in estimation than the index approach. We show that aggregate analysis using the housing prices without quality adjustment underestimates the marginal impact of land use regulation by about 33%. We find macro variables are empirically important to predicting time-series movement of housing prices. For example, a one percentage point increase in real GDP per capita is associated with 1.3% increase in housing prices and. a one percentage point increase in the growth of household mortgage debt and the 30-year fixed rate mortgage rate leads to 2.88% and -2.59% change in housing prices respectively.

In addition to identifying and measuring regulatory effects decomposed into supply and demand effects for metro areas, we take a more granular view to explicitly examine the within-MSA interdependence of land use regulation among cities. We define the difference between neighboring and home regulatory indices as a relative restrictiveness index, whose marginal contribution to the housing prices measures the spillover effect of regulation. We examine 4 major MSAs selected for their data coverage. We show a robust finding that the regulatory and the spillover effects on housing prices are significant and positive. For the relative restrictiveness indices that weigh the neighboring regulatory impacts in different model specifications, leaving the spillover effect out of the analysis tends to underestimate the regulatory effect on housing prices in the home city.

We aim to establish a direct mapping from the theoretical model to the empirical estimates. We use multiple data sources for the data counterparts in our theory. We obtain residential property transaction data in California based on a property-level transaction-assessment dataset from the Zillow Group. We control a comprehensive set of housing characteristics (the number of bedrooms/bathrooms, distance to CBD, indicator of single family residential/condominium, square foot, property age). The data on GDP per capita is calculated based on the regional/MSA dataset from Moody's Analytics. To deal with the endogeneity issue of the GDP per capita, we construct demographic variables from American Community Survey as instruments, including mean household age, share of high education and share of high-tech jobs. In addition to the housing characteristics, we control macro variables including the growth of household mortgage debt and the real 30-year fixed rate mortgage rate for goodness of fit along the time-series dimension.

We examine 179 cities in the metro areas from 185 cities surveyed in California in the Wharton Residential Land Use Survey (Gyourko, Saiz and Summers, 2008). We use principal factor analysis to quantify the intensity of land use regulation by creating a unidimensional index of regulation intensity, and to quantify the factorial contribution of the underlying sub-indices. The housing sample matched to the Wharton survey include more than 5 million transactions, ranging from 1993 to 2017 in 25 MSAs.

The organization of the paper is as follows. Section 2 reviews the literature. Section 3 sets up a spatial equilibrium model of land use regulation with endogenous housing prices. Section 4 describes the data and summary statistics. Section 5 maps the model to the structural estimation. Section 6 discusses the decomposition of the regulatory effects through the production and the amenity channels. Section 7 estimates the city-level spillover effects, followed by the conclusion in Section 8.

2. Literature Review

Spatial equilibrium models in urban economics date back to the pioneering work by Rosen (1979) and Roback (1982) and is enriched by Glaeser and Gottlieb (2009). There are two types of theoretical models related to the impact of land use regulation on housing prices, Brueckner (1990) and Engle, Navarro and Carson (1992) propose amenity models with negative population externality in the utility. Regulation mitigates negative externalities and boosts housing prices. The analysis holds only under the small-city assumption, with no role of the supply constraint of land. On the other hand, Brueckner (1995) does propose a supply-restriction model that puts constraints on the developable land and emphasizes supply constraints in housing price determination. Our theoretical model has both demand- and supply-side regulatory implications on housing prices.

In the discussion of the geographical interdependence of land use regulation and housing prices across regions, Pollakowski and Wachter (1990) first introduce the concept of the spillover effect in the housing market. Tighter growth controls in the neighboring area will increase the home housing prices through housing demand. Our theoretical model formalizes the idea by means of the location choices of households.[2] Our empirical method improves the estimations of Pollakowski and Wachter (1990), in the sense that we use a more comprehensive regulatory measure and make housing quality adjustment in the estimation that is infeasible in the construction of the price indices in previous study.[3] [4] We show that the city-level spillover effect is strong but the regulatory implications are quite heterogeneous across metro areas.

Empirical models vary greatly in their data and methods and their results are not directly comparable as pointed out by Quigley and Rosenthal (2005). The regulatory factors used in empirical analyses vary across studies with different linear or non-linear scales, making the evaluation of the marginal effects of regulation on housing prices more case-by-case.

Most studies develop a regulatory index and find positive and significant regulatory impact on housing prices. Pollakowski and Wachter (1990) construct a regulatory index as the weighted sum of land in various zoning categories. Jackson (2016, 2018) apply the same method to California Land Use Survey (Mawhorter and Reid, 2018). Quigley and Raphael (2005), Ihlanfeldt (2007) and Glaeser and Ward (2009) focus on California, Florida and the Great Boston respectively and define the city regulation index as the total number of adopted regulatory controls. Kok, Monkkonen and Quigley (2014) study the San Francisco Bay Area and use a normalized regulatory index in their analysis. Glaeser, Gyourko and Saks (2005b) examine define a regulatory tax measure as the markup of the housing price over the marginal cost for NYC. Malpezzi (1996) and Malpezzi, Chun and Green (1998) use the Wharton survey of Planning and Policy (Linneman, Summers, Brooks, and Buist, 1990; Buist, 1991) to construct a simple sum of standardized sub-indices as the regulatory measure.[5] Gyourko, Saiz and Summers (2008) update the original Wharton survey. They conduct a national survey with responses from 2,649 jurisdictions and use a principal factor analysis to construct a single regulatory index. Many subsequent studies use the Wharton Land Use Survey and study the regulatory impact in housing and land markets. Saiz (2010) estimates the housing supply elasticity as a function of physical constraints

[2] Consistent with Pollakowski and Wachter (1990), our empirical result echoes the finding on the significant and positive impact of the relative regulatory restrictiveness on the housing prices.

[3] Pollakowski and Wachter (1990) use the transaction prices as the dependent variable and control the real per capita income, distance to Federal Triangle, Gravity Employment Index, real mortgage rate, real construction cost index and percentage of vacant land. They conduct a pooled cross-section time-series regression with 17 areas and 24 quarters to construct the real housing price indices. In comparison, we match the property transactions with the structural characteristics of housing and take into account the property-level heterogeneity that we find crucial to the estimates of the regulatory impacts.

[4] Our regulatory index and the measure of relative regulatory measures are more comprehensive than Pollakowski and Wachter (1990). 3 of the 8 underlying factors, the open space index and the supply restriction index and the local zoning approval index are close to 3 regulatory measures in the previous studies (the percent of vacant land, the development ceiling, the zoning index respectively), although the mapping is not identical. We use multiple measures of the spillover effects to confirm the robustness of our results.

[5] Malpezzi (1996) tries factor analysis as an alternative data reduction method. Because the aggregate score by the simple sum and by factor analysis are highly correlated, Malpezzi (1996) reports the results using the simple sum. Jackson (2016, 2018) do the same as Malpezzi (1996) on a different dataset.

and regulatory measures from the Wharton survey. Turner, Haughwout and Van Der Klaauw (2014) use the Wharton survey to identify the local regulatory effect on the land transaction prices at the boundaries of adjacent jurisdictions with different regulation. Quigley, Raphael and Rosenthal (2008) uses the Wharton survey instruments that are adapted to California to study the housing markets in the San Francisco Bay Area.

Our methodology is close to Gyourko, Saiz and Summers (2008) and Quigley, Raphael and Rosenthal (2008) that use principal factor analysis to define the first common factor as the regulatory index, achieving data reduction of multiple sub-indices. Because we focus on property sales in California, we use the sub-indices with within-state variation in the Wharton Residential Land Use Regulation Index (WRLURI) (Gyourko, Saiz and Summers, 2008) to construct the regulatory measure.[6]

3. Model

3.1 Household Problem

Consider an economy with a unit mass of households. Household *i* values the non-durable consumption *c* and housing consumption *h*. We assume that the household's preference has a Cobb-Douglas form. A household makes two sets of choices on consumption and the location. Given staying in city *j* and housing rent r_j, Household *i* solves the standard consumption choice problem.[7]

$$v_j^i(r_j) = \max_{c,h}(1-\alpha)\ln c + \alpha \ln h + \beta_{ij} \quad s.t. \ r_j h + c \le Y_i Z_j A_j, \text{ where } A_j = Z_j^{\phi-1}\tau_j^{\eta} \quad (1)$$

The indirect utility of household *i* in city *j* can be written as a function of housing rent, r_j. We assume that the household income consists of three components: an idiosyncratic household income Y_i, a city-specific income Z_j, and amenity value A_j.[8] We assume that two income components, Y_i and Z_j, are independently distributed and are multiplicative for tractability of analysis.

We assume amenity value A_j is a function of city income Z_j and the regulation intensity τ_j. The value ϕ-1 controls the income elasticity of amenity demand. If $\phi > 1$, the amenity value increases with

[6] For the discussion on WRLURI, see Gyourko, Saiz and Summers (2008) and Gyourko and Molloy (2015).

[7] We assume that the expenditure on housing consumption is linear in the housing rent. There are models in the literature with non-linear pricing to take into account housing quality (Landvoigt, Piazzesi and Schneider, 2015). We make the assumption not only because linear pricing is simple and tractable for analysis, but also because we are able to use the housing transaction-assessment matched data with detailed structural characteristics to control housing quality in the model estimation.

[8] Glaeser and Gyourko (2006) point out the importance of spatial heterogeneity of amenities in housing price dynamics. Similar to their work, we incorporate the impact of amenities in the household utility. We incorporate the amenity in the model as a multiplier of the household income. With log preference, it is mathematically equivalent to a model where amenity creates an additive utility flow $\ln(A_j)$.

city income. The parameter η control the regulation elasticity of amenity demand. Amenities in the model serve as the demand shifter of both non-durable and housing consumption.[9]

The parameter α measures the share of housing consumption relative to non-durable consumption in total expenditure. B_{ij} denotes the city utility flow to an individual household; this captures personal preference of location and any hidden benefit unobservable to econometricians. Conditional on living in city j, the optimal housing consumption choice and the indirect utility function are

$$h_i^* = \frac{\alpha Y_i Z_j^{\phi} \tau_j^{\eta}}{r_j} \tag{2}$$

$$v_j^i(r_j) = \alpha \ln \alpha + (1-\alpha)\ln(1-\alpha) - \alpha \ln r_j + y_i + \phi z_j + \eta \ln \tau_j + \beta_{ij} \tag{3}$$

where $y_i = \ln(Y_i)$ and $z_j = \ln(Z_j)$. summarizes the city-specific value. The location choice of household i is thus a discrete choice problem. If household i moves to the city j among a set of cities S instead of an alternative city k in the choice set. Then, the utility given city j must yield the highest value.

$$\begin{gathered} v_j^i(r_j) \geq \max_{k \neq j} v_k^i(r_k) \Leftrightarrow \\ \phi z_j + \eta \ln \tau_j - \alpha \ln r_j + \beta_{ij} \geq \phi z_k + \eta \ln \tau_k - \alpha \ln r_k + \beta_{ik}, \text{ for all } k \neq j \end{gathered} \tag{4}$$

We assume that β_{ij} is unobservable to econometricians and it is identically and independently Type-I Extreme-Value distributed across cities. That is, when a household makes a location choice, they can make decisions based on the realization of the city income, the growth controls relevant to amenity value, the housing price and a private signal β_{ij} about the utility flow from city j.[10] The difference $\beta_{ij} - \beta_{ik}$ has a Logit distribution, because the private signal is Type-I Extreme-Value distributed. The share of households located in city j is thus as follows.

$$q_j(r) = \frac{Z_j^{\phi} \tau_j^{\eta} r_j^{-\alpha}}{\sum_{k \in S} Z_k^{\phi} \tau_k^{\eta} r_k^{-\alpha}}, \; r = \{r_k\}_{k \in S} \tag{5}$$

We can interpret the share as a standardized city index that households create to make location choices based on observables. As we normalize the mass of household to unity, the share of household living in city j coincides with the moving probability of a household to city j.

[9] Glaeser, Kolko and Saiz (2001) regress the log housing price on the log per capita income and define the residuals as the amenity indices. Our modeling of the amenities inherits similar idea and define the non-linear piece of city income after taking out the linear component as the amenity value.

[10] In estimation, we will use the log GDP per capita of the MSA where city j is located as the data counterpart of z_j. Quigley and Rosenthal (2005) says that although land use regulation is local, growth is regional. As a result, we only allow income variation across MSAs in the empirical analysis.

3.2 Housing Developer Problem

The production of housing service needs land L as the input. In each city, we assume there is a local housing developer, operating a production technology with decreasing return to scale. The assumption motivates an upward sloping housing supply curve. The housing developer pays a license fee F_j charged by the local government to operate the business and pays the city-specific marginal cost c_j for each unit of land. c_j captures both the construction cost of materials and labors, and the shadow cost tied to local land use regulation. We assume that the housing produced each period is fully depreciated. The housing developer in city j solves the profit maximization problem.

$$\max_{L,H} r_j H - c_j L - F_j \quad s.t.\ H = A_0 L^{\theta},\ c_j = \tau_j c_0 \tag{6}$$

where $A_0 > 0$ is the aggregate productivity and $\theta < 1$ controls the curvature of production technology. c_0 is the construction cost associated with materials and labor, identical across cities.[11]

The parameter $\tau_j > 0$ is the intensity of land use regulation. The concept is similar to Glaeser, Gyourko and Saks (2005b). The more regulated the land market in city j is, the higher τ_j will be. Concretely, the parameter is a reduced-form index of regulation, taking into account the shadow costs of land supply elasticity, time length of permit approval, density and supply restriction etc. The regulation intensity τ_j can be interpreted as an aggregate of multiple measures of land use regulation.

$$\tau_j = \prod_s (\tau_j^s)^{\rho_s} \tag{7}$$

where τ_j^s is an underlying regulation factor and $\rho_s > 0$ is the corresponding factor weight.[12] The profit maximization of a local housing developer leads to a land supply curve with a positive slope in city j.

$$H_j(r_j) = A_0^{\frac{1}{1-\theta}} \left(\frac{\theta r_j}{c_j} \right)^{\frac{\theta}{1-\theta}} \tag{8}$$

Housing supply is thus increasing the productivity and housing rent but is decreasing in the cost of land. The local government will set the license fee F_j to charge away any positive profit, so in equilibrium, the local housing developer has zero profit.

[11] This assumption is without loss of generality. We think the assumption is reasonable, because the construction industry is extremely competitive (Saiz, 2010). Gyourko and Molloy (2015) in their Figure 1 shows that the real construction cost is stable compared with the strong movement of the real housing prices. We can relax the assumption to allow for time-varying construction cost. What is essential is that the construction cost is that there is no cross-sectional variation among cities and that it is exogenous to local changes in housing demand over time.

[12] We assume the relationship between the unidimensional measure and the underlying factors of land use regulation follows a product form. The log form of equation (7) will correspond to the predicted score regression in the principal factor analysis that we use to construct a unidimensional index from multiple measures of land use regulation.

3.3 Exogenous Processes

To bridge the housing price with housing rent, we assume that the conventional user cost relationship between housing price p_j and housing rent r_j in city j holds.

$$\ln p_j = \ln r_j - \ln u_j \tag{9}$$

where u_j is the user cost. In the model, we take the log city income z_j, and the log user cost $\ln(u_j)$ as exogenous and time-varying. We assume that the log income z_{jt} and the log user cost at time t have independent normal distributions.

$$z_{jt} \sim N(\mu_z, \sigma_z^2),\ \ln u_{jt} \sim N(\mu_u, \sigma_u^2),\ \forall j, t \tag{10}$$

3.4 Equilibrium Conditions and Housing Prices

There are two equilibrium conditions needed to satisfy to close the model. First, each household with random utility flow unobservable to econometricians should move to the city delivering the highest utility. The optimal consumption and location choices have been encoded into the moving probability $q_j(r)$. Second, the housing price of each city is an endogenous object. We clear the housing markets in all cities and solve the prices simultaneously. The market clearing condition (11) demand that we equate the housing demand by aggregating the individual demand (2) and the housing supply (8) in each city.

$$q_j(r)\frac{\alpha Y_0}{r_j} Z_j^{\phi} \tau_j^{\eta} = H_j(r_j) \text{ for all } j \in S \tag{11}$$

where $Y_0 = E(Y_i)$ is the expected household income. The house demand in city j is thus the product of the share q_j of households moving to city j and the aggregated house demand in city j.[13] As is shown in the equilibrium condition, the housing markets are inter-linked. The market clearing condition of city j depends on the housing prices in other cities. Households have freedom to move and will choose their location depending on city-specific income and private utility flow. The impact of local land use regulation will spill over to the other cities through the location choice of households.

We prove in the appendix that for an arbitrary number of cities $n \geq 2$, there exists a unique set of moving probabilities and housing prices that clear the housing markets in n cities. In the following analysis, we focus on the location choice with binary options ($n = 2$), cities j and k. We examine the price determinants of a particular city j, and the city k is interpreted as the outside moving option of city j. It simplifies the mapping from the model to the data and makes the illustration of mechanism straightforward.

---

[13] Because Y_i and Z_j are assumed independently distributed, we can integrate over the household demand and get the housing demand in city j. Because the individual housing demand is linear in Y_i, only the first moment is needed for aggregation.

$$\ln p_j = (1-\theta)(\ln q_j - \ln b_j) - \ln u_j$$

$$\text{where } q_j = \frac{(Z_j^{\phi}\tau_j^{\eta})^{1-\lambda} b_j^{\lambda}}{(Z_j^{\phi}\tau_j^{\eta})^{1-\lambda} b_j^{\lambda} + (Z_k^{\phi}\tau_k^{\eta})^{1-\lambda} b_k^{\lambda}},\ \lambda = \frac{\alpha(1-\theta)}{\alpha(1-\theta)+1},\ b_j = \frac{A_0^{\frac{1}{1-\theta}}}{\alpha Y_0 Z_j^{\phi}\tau_j^{\eta}}\left(\frac{\theta}{c_j}\right)^{\frac{\theta}{1-\theta}} \quad (12)$$

To understand the determinants of the housing price in city *j*. we express ln(p_j) explicitly as follows.

$$\begin{aligned} \ln p_j = &\{\theta \ln c_0 + [\theta + \eta(1-\theta)]\ln \tau_j\} - (1-\theta)\ln\left[1 + e^{(2\lambda-1)\phi(z_j - z_k) + [\frac{\theta}{1-\theta}\lambda + \eta(2\lambda-1)](\ln\tau_j - \ln\tau_k)}\right] \\ &+ [(1-\theta)(\ln Y_0 + \phi z_j) - \ln A_0 - \ln u_j] + [(1-\theta)\ln\alpha - \theta\ln\theta] \end{aligned} \quad (13)$$

There are four terms that determines the log housing price. The first two terms are associated with the land use regulation through the cost of land. The first term summarizes the production channel that show that a higher local housing price reflects higher marginal cost of land due to tighter land use regulation. The second term shows the general equilibrium effect of the housing markets. Regulatory change may induce households to make new location choices. Leading to reallocation of housing demand and price adjustment of multiple cities. The first two terms indicate two opposite forces of regulation intensity on the housing price. If we apply first-order Taylor approximation to the second term,

$$\begin{aligned} &(1-\theta)\ln\left[1 + e^{(2\lambda-1)\phi(z_j - z_k) + [\frac{\theta}{1-\theta}\lambda + \eta(2\lambda-1)](\ln\tau_j - \ln\tau_k)}\right] \\ &\approx \tfrac{1}{2}\{(1-\theta)(2\lambda-1)\phi(z_j - z_k) + [\theta\lambda + \eta(1-\theta)(2\lambda-1)](\ln\tau_j - \ln\tau_k)\} + (1-\theta)\ln 2 \end{aligned} \quad (14)$$

Combining the first and the second terms, we find that the effect of regulation intensity through the production channel is dominant, leading to a positive relationship between regulation intensity and housing price in aggregate.

The third term, (1-θ)(lnY_0+ϕz_j)–ln(A_0)–ln(u_j), summarizes the expected household income, the productivity of housing production, and the local user cost. Higher income will increase housing demand and increase the housing price in city *j*, while higher productivity will increase the housing supply and decrease the housing price. Given the rent in city *j*, higher user cost implies a lower housing price.

For the first part of the last term, (1-θ)ln(α), determines the marginal rate of substitution between housing and non-durable consumption. A higher marginal value from housing increases demand, and thus the housing price. The last term, θln(θ), characterizes the production technology. When θ is smaller, the retained profit of the housing developer will be higher, which is consistent with a higher equilibrium housing price.

Besides using the Taylor approximation in (14) to simplify the log housing price equation (13), we make a normalization assumption on city *k* which is the outside moving option of city *j*. We normalize τ_k = 1, indicating constant regulation intensity of outside moving option for any city.[14] To be

[14] As is shown in (13), the normalized value will only affect the level of log housing price.

consistent, we do the same normalization when we measure regulation intensity in the data. For the city income of the outside moving option, we assume it is the mean income of all cities.

$$z_k = \frac{1}{n}\sum_{l \in S} z_l \tag{15}$$

Because z_k is constant for each cross-section, it plays a role similar to the year fixed effect in the log housing price equation. Together with (13), we can use the Taylor approximation and the normalization to determine the housing price differential across cities. For two cities *j* and *j'*,

$$\begin{aligned} \ln p_j - \ln p_{j'} &\approx [\theta(1-\tfrac{1}{2}\lambda) + \eta(1-\theta)(\tfrac{3}{2}-\lambda)](\ln \tau_j - \ln \tau_{j'}) \\ &+(1-\theta)(\tfrac{3}{2}-\lambda)\phi(z_j - z_{j'}) - (\ln u_j - \ln u_{j'}) \end{aligned} \tag{16}$$

The cross-city price differential consists of three terms. The first term indicates that the regulation intensity differential between cities *j* and *j'* and has a positive impact on the cross-city housing price differential. The second term emphasizes a positive correlation between city income differential and cross-city housing price differential. The general equilibrium effect mitigates the first and the second term by a fraction λ. The last term captures the differential of user costs across cities and has a negative impact on the housing price differential.

4. Data

We use multiple sources of data. The land use regulation data are derived from the Wharton residential land use regulation survey. The housing data come from the Zillow Transaction and Assessment Dataset. The regional data is based on the dataset compiled by Moody Analytics and American Community Survey.

4.1 Land Use Regulation Data

To measure the land use regulation intensity in the data, we rely on the sub-indices underlying the Wharton Residential Land Use Regulation Index (WRLURI) compiled by Gyourko, Saiz and Summers (2008).[15] WRLURI is a cross-sectional survey and is estimated at the jurisdiction levels (cities hereafter). We focus on the cities in California state that are covered by WRLURI, because the quality of land use regulation data and the housing data in California is better than that in other states.[16] Moreover, jurisdictions in California enjoy remarkable autonomy of land use regulation, creating

[15] Data on WRLURI is available online (http://real.wharton.upenn.edu/~gyourko/landusesurvey.html).

[16] The number of cities covered by the land use regulation survey in California is the second highest among all states, only 2 cities fewer than Pennsylvania. The housing data discussed below has more comprehensive coverage and longer time length in California than in other states.

geographical variations of policies (Fischel and Fischel, 1995). Throughout our analysis, we assume that the land use regulation is constant over time.[17]

There are 185 cities in California that responded to the Wharton Land Use Survey. While WRLURI covers only a limited number of jurisdictions (Turner, Haughwout and Van Der Klaauw, 2014), the survey data covers 43 out of 103 principal cities marked by the Census Bureau, including the top 6 cities with the highest population in California (Los Angeles, San Diego, San Jose, San Francisco, Long Beach and Fresno).[18] The survey topics range from zoning and project approval to supply and density restriction that are aggregated into 11 sub-indices as the bases of WRLURI. Not all sub-indices are city-dependent with a state. We thus use only 8 sub-indices that vary across jurisdictions to construct a unidimensional measure of regulation intensity, including the local political pressure index (LPPI), local zoning approval index (LZAI), local project approval index (LPAI), density restriction index (DRI), open space index (OSI), exactions index (EI), supply restriction index (SRI), approval delay index (ADI).[19]

Similar to Gyourko et al (2008), we apply the principal factor analysis to the 8 sub-indices above and define the predicted score of the first factor as the measure of land use regulation intensity. We use the regression method to derive the score. We normalize the score to zero mean and unit variance and define the standardized value as the California Land Use Regulation Index (CALURI). The model counterparts are $\ln(\tau_j)$ for CALURI and $\ln(\tau_j^s)$ for the sub-index s.

In Figure 1, we show the spatial distribution of regulation intensity in California across 185 cities. Noticeably, several cities within Los Angeles-Long Beach-Anaheim Metropolitan Statistical Area are highly ranked in terms of regulation intensity. In Figure 2, we show the kernel density of CALURI. Compared with the standard normal density, the distribution of CALURI is more concentrated near the mean. CALURI has a fat right tail, indicating a non-trivial number of highly land use regulated cities. In the appendix, we list the estimated CALURI by MSA and city. In Figure 3, we compare CALURI with WRLURI. We show that CALURI is highly positively correlated with WRLURI and the simple sum of the 8 sub-indices underlying CALURI, so the method of constructing the index is not driving the unidimensional measure of regulation intensity.

[17] We recognize how stringent the assumption of constant regulation intensity is but given the cross-sectional nature of the Wharton Land Use Survey and the slow-moving nature of the land use regulation reform, we believe our results will not be driven by the assumption.

[18] The principal cities within metropolitan and micropolitan statistical areas uses the 2006 US Census definition to align with the survey year. The ranking of the city population in California comes from US Census. For the number of principal cities covered by each metro area, see the appendix Table A2.

[19] The three sub-indices for dropout are the state political involvement index (SPII), the state court involvement index (SCII), and local assembly index (LAI) that is available only in New England. For the definitions of the sub-indices, see Gyourko et al (2008).

4.2 Housing Data

For the housing data, we rely on the Zillow Transaction and Assessment Dataset (ZTRAX).[20] The entire ZTRAX dataset contains more than 370 million public records from across the US and includes information on deed transfers, mortgages, property characteristics, and geographic information for residential and commercial properties (Graham, 2018).

Particularly, we are interested in the transaction prices in the deed transfers and the housing characteristics in the property assessment in California. We restrict the data to observations with non-foreclosed sales of residential properties that have detailed documentation of housing characteristics. We use the following housing characteristics: the transaction date, the property use, the number of bedrooms, the number of bathrooms, the age of the property, the property size and the distance to the nearest core cities. We encode the age of the property, the property size and the distance to the nearest core cities that are not directly observable in ZTRAX. The age of the property is calculated as the difference of the transaction year and the built year. There are multiple fields measuring different aspects of the size of a property, so we define the maximum value in those fields as the property size. For properties located in a city in a Core-Based Statistical Area (CBSA), we calculate the great-circle distance in miles to the center of the leading principal city listed in the name of a CBSA. If there are multiple leading principal cities in the CBSA title, we use the distance to the center of the nearest leading principal cities. Other housing characteristics are available in ZTRAX, but they are either optionally reported or sparsely populated. The details of data filtering and construction of variables are documented in the appendix. We use the city name of a sales transaction as the key to match ZTRAX to the land use data. 184 out of 185 cities responded to the Wharton Land Use Survey have at least one transaction record in ZTRAX (with Crescent City as the only exception).

4.3 Regional data

We calculate gross domestic products (GDP) per capita based on the city income data comes from Moody's Analytics. Moody Analytics compile GDP of 402 US metropolitan statistical areas or metropolitan divisions from Current Employment Statistics, Bureau of Economic Analysis and County Business Patterns, and collect data on the metropolitan population from US Census Bureau. Both the GDP and the population are annual basis. Ideally, we would use city-level income and population, but we use the MSA-level data instead city-level data are not available in general or long enough.[21] Although land use regulation is local, growth is regional (Glickfeld and Levine, 1992; Quigley and

[20] More information on accessing the data can be found at http://www.zillow.com/ztrax. ZTRAX database is provided by the Zillow Group. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group or any of its affiliates.

[21] Moody's data at the MSA level traces back to 1990 and allow us to use observations from all sample years in ZTRAX Also those metropolitan statistical areas, by definition, are socioeconomically tied to the principal cities by commuting.

Rosenthal, 2005; Quigley and Swoboda, 2007). Assuming the city income component of a household to be constant within an MSA sounds reasonable, while we still allow for city-specific characteristics to determine the location choice of households.[22] Moody Analytics only covers the city income and the population in the metropolitan statistical areas instead of micropolitan statistical areas. 179 out of 185 cities responded to the Wharton Land Use Survey are matched to an MSA in Moody's data.[23]

To account for possible endogenous concerns of GDP per capita, we additionally collect other regional data as instrumental variables. The lag term of the log GDP per capita is a natural instrumental variable. In addition, we have 3 candidate instrumental variables on MSA demographics: the share of high education including college and graduate education for at least 1 year, the age of the population, and the share of high-tech jobs. Data on the share of high education and the average age of the population come from the American Community Survey (ACS) Micro data from IPUMS USA. Because ACS data starts from 2000, we fit the time trend and extrapolate the data for each MSA before 2000. Data on the share of high-tech jobs from 1990 to 2017 is compiled by Moody's Analytics, based on Bureau of Labor Statistics and Bureau of Economic Analysis.[24]

4.4 Macroeconomic data

In addition, we control for variables related to macroeconomic conditions. The data covers the period that witnesses the strong boom and bust in residential mortgage and housing prices from 2001 to 2007 in California (Choi, Hong, Kubik, and Thompson, 2016). The time series variation of housing prices may heavily depend on lending conditions. We take this concern into account by introducing two macro variables: the growth rate of household mortgages in the US and the US 30-year average fixed-rate mortgage rate. Higher growth rate of mortgage lending is expected to increase housing demand by easing household borrowing, while a lower mortgage rate achieves the same effect by making borrowing cheaper. The macro variables serve to improve the goodness of fit along the time dimension.

---

[22] Note that using MSA-level income from the data to proxy the regional component z_j in the model doesn't mean that city-specific feature is not important in households' decisions. The data counterpart of a city j is mapped to a city or a town in the data. In the model, the utility of a household depends on city-specific utility flow and house prices.

[23] 6 cities we drop in the analysis fall into 6 micropolitan statistical areas. They are: Fortuna city in Eureka-Arcata-Fortuna, μMSA; Lakeport City in Clearlake, μMSA; Susanville City in Susanville, μMSA; Ukiah City in Ukiah, μMSA; Corning City in Red Bluff, μMSA; Crescent City in Crescent City, μMSA.

[24] High-tech jobs are defined from the following NAICS industries (NAICS code): Pharmaceutical and Medicine Manufacturing (3254), Computer and Peripheral Equipment Manufacturing (3341), Communications Equipment Manufacturing (3342), Semiconductor and Other Electronic Component Manufacturing (3344), Navigational, Measuring, Electromedical, and Control Instruments Manufacturing (3345), Medical Equipment and Supplies Manufacturing (3391), Software Publishers (5112), Wired Telecommunications Carriers (5171), Wireless Telecommunications Carriers (except Satellite) (5172), Satellite Telecommunications (5174), Other Telecommunications (5179), Other Information Services (5191), Data Processing, Hosting, and Related Services (5182), Computer Systems Design and Related Services (5415), Scientific Research and Development Services (5417), Other Professional, Scientific, and Technical Services (5419), Medical and Diagnostic Laboratories (6215)

In Figure 4, we show the time paths of the macro variables. We collect the data on the US household mortgage debt from Z.1 Financial Account Table from the Board of Governor of Federal Reserves and calculate the annual growth rate. The data on US 30-Year average fixed-rate mortgage rate comes from Primary Mortgage Market Survey by Freddie Mac.

4.5 Summary Statistics

In Table 1, we show the geographical coverage of our matched land use sample of property transactions in 179 cities. Property sales in 963 cities are not matched to a city in the land use regulation data, but our matched sample covers 5.3 million residential transactions in 39 out of 58 California counties and 25 out of 26 metropolitan statistical areas in California from 1993 to 2017 in ZTRAX.

In Table 2, we report the summary statistics of CALURI, together with the 8 underlying sub-indices and WRLURI originally estimated by Gyourko, Saiz and Summers (2008). The city-level regulation indices are weighted by the number of property transactions in the cities. CALURI has a positive mean 0.27, a median -0.01, and a standard deviation 1.23. Because CALURI is normalized to zero mean and unit variance, the weighted statistics are consequences of the property transactions concentrated in more regulated and more populated cities in our sample.

In Table 3, we show the distribution of residential property uses. 76% of the property transactions are single-family residential, followed by 21% of condominium transactions. Compared with the distribution of the unmatched sample, we have a lower share of single-family units and a higher share of condominiums in the land use sample (84% and 13% in the unmatched sample respectively).

In Table 4, we report the summary statistics of the housing characteristics we control in the empirical analysis. The sales prices have been inflation adjusted to 2006. The average sales price is $370,000 dollars. The average size of a residential property is 1,700 square feet. We also show the sales price per square foot mean and median as $221 and $181. The average age of a residential property is 30 years. There are 2 bathrooms and 3 bedrooms on average in a residential property. The mean and the median distance of a property to the nearest core city in a metropolitan statistical area is 28 miles and 8 miles, respectively.[25]

In Table 5, we show summary statistics of the instrumental variables. The average share of high education is 36% in an MSA, while 6.84% of the total employment are high-tech jobs. The average age of an individual is 35 years ago. In Table 6, we report the correlation of the real GDP per capita with its lag term and 3 demographic instrumental variables, 0.99, while its correlation with the share of the share of high education, the population age, and the share of high-tech job 0.823, 0.753 and 0.651, respectively.

[25] Compared with the unmatched sample, the average property in the land use sample is more expensive in terms of the sales price per square foot and is smaller in size. It has slightly older age and a shorter distance to the nearest core cities. The number of bath rooms and bedrooms are close in the matched and unmatched samples.

5. Structural Estimation and Results

5.1 Estimation Method

We apply a first-order Taylor approximation to the equilibrium condition of housing price (13) and express it in a linear form for estimation.

$$\begin{aligned}\ln p_{ijmt} &= \beta_0 + [\theta(1-\tfrac{1}{2}\lambda) + \eta(1-\theta)(\tfrac{3}{2}-\lambda)]CALURI_j \\ &\quad + (1-\theta)(\tfrac{3}{2}-\lambda)\phi z_{mt} + \tfrac{1}{2}(1-\theta)(2\lambda-1)\phi z_{0t} + X_{ijmt}\gamma + M_t\nu + \varepsilon_{ijmt} \qquad (17)\end{aligned}$$

$$\text{where } z_{0t} = \sum\nolimits_m g_{mt} z_{mt}, \sum\nolimits_m g_{mt} = 1, \lambda = \tfrac{\alpha(1-\theta)}{\alpha(1-\theta)+1}$$

The log real housing price as the dependent variable has 4 subscripts that uniquely identify an observation of property transaction: property *i*, city *j*, MSA *m*, and year *t*. β_0 is the constant term. z_{mt} is the log real GDP per capita of MSA *m* where property *i* is located. z_{0t} is the log of population-weighted mean GDP per capita of California, with g_{mt} to be the population share of MSA *m* in year *t*. To take into account the structural characteristics of residential properties, we control a vector of housing characteristics X_{ijmt}.[26] To control for the time-varying macro conditions, we use a set of macro variables M_t, with the vector of the corresponding coefficients stored in υ.

The number of parameters is more than that of the moment conditions. We need one more condition to achieve the just identification of the model. We thus exogenously estimate a relationship between η and ϕ, using the correlation of regulation and the log per capita income. We log-linearize the identity of the amenity demand in (1) and transform it into the following auxiliary regression.

$$z_m = -\frac{\eta}{\phi-1}\ln\tau_j + cons + controls_m + residuals_{jm} \qquad (18)$$

where amenity is treated as the residuals.[27]

[26] The housing characteristics include the property use, the number of bedrooms, the number of bathrooms, the property age, the log property size, and the log miles to the nearest core cities. We recode the property use into three main categories: single-family residential, condominium and others. The number of bedrooms and the number bathrooms are recoded into 5 levels (0, 1, 2, 3, 4+), while the age of property is divided into 8 levels (0, 1-5, 6-10, 11-20, 21-30, 31-40, 41-50, > 50). Recoding the numbers and the age into the discrete bins allows us to control the non-linear effects on the housing price.

[27] In the auxiliary regression, we use the MSA-level data from year 2006. The regulation intensity is aggregated to the MSA level using the probability weight from Gyourko, Saiz and Summers (2008). We add demographic variables as controls. The demographic controls include the tech-job share, the mean age of MSA, the college share, the minority share, the net migration, the employment, the index of cost of doing business and the population. We show the definition of the controls, the model specification and the result of the auxiliary regression in the appendix. We use logarithmic transformations to the property size and the distance to the core cities. The details of the auxiliary model are reported in the appendix. We find the estimated coefficient of ln(τ_j) to be -0.0033, leading to an additional condition for the main estimation: η= 0.0033(ϕ-1). For robustness, we also use the city-level income data aggregated from the tract-level income in 2009-2014 5-year ACS. We find the estimated coefficient is 10 times bigger but statistically insignificant and still economically small. Our estimations won't qualitatively change, when we use the condition with the alternative estimate.

Besides the coefficients of housing characteristics and macro variables, we need to estimate 3 structural parameters (θ, λ, ϕ) using 3 instruments ($CALURI_j$, z_{mt}, z_{0t}). It is more convenient here to treat λ instead of α as a primitive parameter. Our estimation strategy is to use Generalized Method of Moments (GMM) to estimate the structural parameters (Hansen, 1982). GMM won't improve the estimation of the just-identified model, but the estimation method can be naturally extended to the models with additional instrumental variables to deal with endogeneity of per capita income.

5.2 Estimation Results

In Table 7, we report the estimation results. The estimation of the coefficients is based on GMM or GMM-IV estimations. In the appendix, we report the estimation of the structural parameters ($\theta, \lambda, \phi, \alpha$).

5.2.1 GMM Estimators

Estimations of Model 1 and Model 2 are based on the model specification without and with the vector of housing characteristics, respectively. When housing characteristics are controlled, Model 2 shows that a unit increase (a standard deviation increase) in the regulation intensity (CALURI) increases the housing price by 2.93%. A 1% increase in per capita income increases the local housing price by 1.326%, while a 1% increase in the population-weighted mean per capita income of California increases the local housing price by 0.352%. Model 1 underestimates the marginal effect of regulation intensity by 33%, The reason is that housing characteristics are correlated with the regressors in Model 1. In our data, regulation intensity is negatively correlated with the property size, the number of bedrooms and the number of bathrooms, and positively correlated with the property age.

One caveat at interpreting the marginal effect of regulation is that the regulatory reference point is the average California city, instead of the average city in the US. As we show in Table 2, the frequency-weighted mean and median of WRLURI are 0.8 and 0.55 respectively, much higher than the weighted mean and median of CALURI (0.27 and -0.01 respectively). The regulation of the average California city is much tighter than that of the average city in the US (see Figure 3). If we mistake the regulatory reference point in Table 7 for the average regulation in the US and improperly extend the California estimates to other US cities, we are going to underestimate the national regulatory impact on housing prices.[28] In Table A5 in the appendix, we replicate our estimations in Table 7, but instead use WALURI

[28] There are two sources of underestimating the national level regulatory impact by using the estimates with CALURI and the California sample. The first source is due to the greater standard deviation of CALURI than WRLURI (1.23 vs 0.79 respectively from Table 2). All else equal, if we scale down a regulatory index (e.g. CALURI) by multiplying a factor $x < 1$, we will equivalently scale up the regulatory impact by a factor of $1/x > 1$ in estimation. The second source is related to the non-linear relationship between CALURI and WALURI. In Figure 3, WALURI roughly increases in CALURI at an increasing rate. The convex relationship indicates that specifications with WALURI will yield a higher estimate of the regulatory impact than those with CALURI.

as the regulatory measure. We show that the estimated regulatory impact with the national average as the reference point is 4 times larger (11.7% vs 2.93%) than the California-based regulatory impact.

5.2.2 GMM-IV Estimators

The city income Z_j in the structural model is exogenous, but the per capita income which is its data counterpart can be endogenous. To deal with the endogeneity, we use the lag terms of GDP per capita and the population-weighted mean GDP per capita in California to instrument the contemporaneous variables in Model 3. In Model 4, we build on Model 3 to include 3 demographic variables (the share of high education, the population age, and the share of high-tech jobs) as additional instrumental variables. The GMM-IV estimators of the regulation intensity, the log GDP per capita and the population-weighted mean GDP per capita of California are not substantially different across Model 3 and Model 4 (0.0290, 1.311 and 0.369 for Model 3; 0.0297, 1.291 and 0.432 for Model 4).[29]

By comparing Model 2 and Model 4, we see the difference between GMM and GMM-IV estimators. Treating per capita income as exogenous in Model 2 also underestimates the marginal effect of land use regulation intensity, albeit by a small amount. In Model 4, a unit increase in the regulation intensity increases the housing price by 2.97%, compared with 2.93% in Model 2.[30] If we use the average US regulation as the reference point, we show in Appendix Table A5 that the regulatory impact on housing price is 12.4%. In Appendix, we report the structural parameter estimates of the models in Table 7.

5.2.3 Factorial Contribution of Land Use Regulation to Housing Prices

Our analysis relies on CALURI as a unidimensional measure of land use regulation intensity, but we can also quantify the marginal contribution of an underlying factor to the housing prices with one more step. Note that CALURI is the predicted score of the first common factor of 8 sub-indices, derived from the regression method of the principal factor analysis. We can recover the contribution of the sub-indices by regressing CALURI on the standardized sub-indices without a constant term.[31]

$$\begin{aligned} CALURI_j &= 0.418 * LPPI_j^{std} + 0.351 * LZAI_j^{std} + 0.412 * LPAI_j^{std} + 0.118 * DRI_j^{std} \\ &+ 0.255 * OSI_j^{std} + 0.151 * EI_j^{std} + 0.147 * SRI_j^{std} + 0.133 * ADI_j^{std} \end{aligned} \tag{19}$$

[29] We also test the model specifications by including one of the three, or two of the three demographic variables as additional instrumental variables. The estimations results are quantitatively similar. The results are available upon request.
[30] For other marginal effects, Model 2 will underestimate the mean per capita income of California on housing price and will overestimate the marginal impact of the log per capita income. A 1% increase in per capita income increases the local housing price by 1.291% in Model 4, compared with 1.326% in Model 2. A 1% increase in the population-weighted mean per capita income of California increases the local housing price by 0.432% in Model 4, compared with 0.352% in Model 2.
[31] A constant term is not needed because both CALURI and the sub-indices have been standardized to zero mean.

where the superscript *std* means that a sub-index is normalized to zero mean and unit variance. The relationship is exact without an error term, because CALURI, by definition, is a rescaled fitted value of the predicted score regression. The factor weights do not sum to one, because CALURI as the predicted score does not necessarily yield unit variance and we have normalized CALURI in the analysis.

The marginal contribution of sub-indices on the housing prices is the product of the marginal effect of CALURI reported in Table 7 and the factor weights in the predicted score regression (19). The factor weights in (19) are mapped to the estimated parameters of $\{\rho_s\}$ in (7). In Table 8, we report the marginal effects of the sub-indices for the model specifications in Table 7.

Local political pressure, local project approval and local zoning approval are the leading factors contributing 21.06%, 20.76% and 17.68% respectively to CALURI. In aggregate, CALURI attaches almost 60% of weight to these three factors. In terms of the marginal effect on housing prices, a unit increase (1 standard deviation increase) in these three sub-indices leads to an increase in the housing price by 1.24%, 1.22% and 1.04% respectively. The availability of open space contributes 12.85% to CALURI, and a unit increase leads to an increase in the housing price by 0.76%. Exactions, supply restriction, approval delay and density restriction consist of the remainder of the contribution (7.61%, 7.41%, 6.70% and 5.94%) and lead to a price increase respectively of 0.45%, 0.44%, 0.40% and 0.35%.

5.3 Foundation and Estimation of the Non-Linear Effects on the Log Housing Prices

We show that land use regulation and the log per capita income have positive impacts on the log housing price. Our estimations yield the average marginal effects. It is natural to ask whether the constant marginal effect is only local, and whether the model ignores any non-constant linear or non-linear effect consideration. Model 4 in Table 7 is treated as the benchmark model in this section where we address this question.

We micro-found the impact of non-constant marginal effects by extending the benchmark to the model that (1) the cost of housing supply may vary with the amenity level; (2) the income elasticity of amenity demand is not a constant. These two extensions represent the supply and demand channels through which amenity and city income can affect the local housing prices.

5.3.1 Foundation of the Interactive Effect

We establish how our estimation equation relies on the assumption that the measured impact of regulation on housing production is correlated with the amenity level. Motivated by the finding in the literature, we generalize the log marginal cost of housing production with the following multiplicative form.[32] [33]

$$\ln c_j = (\delta_1 z_j + \delta_0)\ln \tau_j + \ln c_0 \tag{20}$$

The parameters δ_1 and δ_0 control the sensitivity of the marginal cost. With $\delta_0 = 1$ and $\delta_1 = 0$, we go back to the benchmark case. When $\delta_1 > 0$ (we show it is the case), the housing supply exhibits a higher price impact in cities with high income and amenity demand. In estimation, we impose a parametric restriction to focus on the following class of the models that include the benchmark model as a special case.

$$\delta_1 E_t(z_{0t}) + \delta_0 = 1 \tag{21}$$

The relationship indicates that the term in the parenthesis in (20) is unity on average. For a property located in an MSA with the log per capita income equal to $E_t(z_{0t})$, *ceteris paribus*, the marginal effect of regulation intensity will be identical in the estimation equations with and without an interactive term. For computation, there are two parameters with one degree of freedom. The new estimation equation will be similar to (17), but with an additional interactive term of CALURI and the log per capita income.

5.3.2 Foundations of the Quadratic Effect

We extend the assumption of constant income elasticity of amenity. The extension results in the quadratic term of the log per capita income in estimation.[34]

The power term ϕ–1 in the benchmark model has the interpretation of the income elasticity of amenity. The amenity adjusted household income can be written as exp(ϕz). We extend the linear form to the quadratic form in the power term.[35]

$$\exp(\phi_0 + \phi_1 z + \phi_2 z^2) = Z \exp(\phi_0) Z^{\phi_2 z + \phi_1 - 1} \tag{22}$$

where the last term on the right side is the amenity value and $2\phi_2 z + \phi_1 - 1$ will be the income elasticity of amenity demand. With $\phi_0 = 0$ and $\phi_2 = 0$, we go back to the benchmark case. When $\phi_2 > 0$ (we show it is the case), the income elasticity of amenity demand is higher for wealthier cities.

[32] Glaeser, Gyourko and Saks (2005a) find that the likelihood to build new housing units, an inverse measure of time cost, is lower in wealthier communities. Homeowners in the wealthy communities may use time to influence local planning (Gyourko and Molloy, 2015). Fischel (2001) brings about the homevoter hypothesis that homeowners in wealthy communities have stronger incentive to protect local amenities capitalized in housing values.

[33] If the impact of the log amenity comes into the marginal cost in an additive form. The parameters δ_1 and δ_0 will remain unidentified in estimation.

[34] We leave out the quadratic effect of the regulation intensity in the section, because we don't find the quantitatively important quadratic effect along the dimension. Moreover, the regulation intensity is an index we construct from sub-indices. We take the stand that the index construction should pick up the high-order effects, if there is any.

[35] The extension of the amenity demand captures two things, the residual linear effect after taking out the unity linear impact of the per capita income, and any non-linear effect of per capita income.

In the estimation, we impose two parametric restrictions to focus on the following class of models that include the benchmark model as a special case.

$$\begin{aligned} \phi_{avg} - 1 &= 2\phi_2 E_t(z_{0t}) + \phi_1 - 1 \\ \phi_{avg} E_t(z_{0t}) &= \phi_0 + \phi_1 E_t(z_{0t}) + \phi_2 [E_t(z_{0t})]^2 \end{aligned} \tag{23}$$

The value ϕ_{avg} - 1 captures the average income elasticity of amenity demand according to the first restriction. The second restriction indicates that when the city income is equal to $E_t(z_{0t})$, the elasticity of amenity demand is identical in the benchmark and the extended model.

To focus on the marginal effect of regulation and per capita income on housing prices, we further make an assumption that a household uses the average elasticity ϕ_{avg}-1 in the location choice problem, which is the same as the benchmark case. The assumption allows us to focus on the quadratic term of the log per capita income as an additional term in the estimation equation (17).[36]

5.3.3 Estimation of the Non-Linear Effects on the Log Housing Prices

Our extended estimation equation with the interactive and quadratic effects takes the following form.

$$\begin{aligned} \ln p_{ijmt} = {} & \beta_0 + [\theta(\delta_0 - \tfrac{1}{2}\lambda) + \eta(1-\theta)(\tfrac{3}{2} - \lambda)]CALURI_j + \theta\delta_1 z_{mt} \cdot CALURI_j \\ & + (1-\theta)[\phi_1 - \tfrac{1}{2}(2\lambda - 1)\phi_{avg}]z_{mt} + (1-\theta)\phi_2 z_{mt}^2 + \tfrac{1}{2}(1-\theta)(2\lambda-1)\phi_{avg} z_{0t} \\ & + X_{ijmt}\gamma + \varepsilon_{ijmt} \end{aligned} \tag{24}$$

We report the estimation of four model specifications in Table 9 and the parameter estimates in the appendix. Model 4 is the benchmark case (δ_0 = 1 and δ_1 = 0; ϕ_0 = 0 and ϕ_2 = 0). Model 5 builds on Model 4 with the interactive effect (ϕ_0 = 0 and ϕ_2 = 0), while Model 6 builds on Model 4 with the quadratic effect (δ_0 = 1 and δ_1 = 0). Model 7 incorporates both effects in the benchmark model.

In Model 5, we find the interactive term has a positive coefficient, so the marginal effect of regulation on the log housing prices is not constant. For an average property in year 2006 located in an MSA whose log per capita income is one standard deviation above (below) the mean, a unit increase in the regulation leads to 3.47% (2.57%) increase in the housing price in Model 5, compared with a uniform 2.97% increase in Model 4. The significant impact of the interactive term supports the hypothesis that there is a direct and positive impact of the city income on the marginal cost of housing production.

The way we model the interactive effect by allowing the cost to housing production to vary with amenity and city income provides one explanation for the positive interactive effect, but there are alternative explanations. Wealthier and bigger cities may have more complex sets of the growth control policies that cannot be fully incorporated in the survey with limited dimensions. If the omitted growth control policies are positively correlated with our regulatory measure and the omitted variable bias is

[36] What the approximation assumption leaves out is the interactive effect of the GDP per capita and the mean GDP per capita of California, and the quadratic effect of the latter term.

more severe in the wealthier and bigger cities, then we will see a larger upward bias of the regulatory effect for the wealthier and bigger cities.

In Model 6, we find the quadratic effect of the log per capita income is positive; the marginal effect of the log per capita income has a positive and increasing impact on the log housing prices. For an average property in year 2006 located in an MSA whose log per capita income is one standard deviation above (below) the mean, 1% increase in the per capita income leads to 1.99% (1.17%) increase in the housing price in Model 6, compared with a uniform 1.29% increase in Model 4. The significant impact of the quadratic term supports the hypothesis that the income elasticity of amenity demand is not constant but positively correlated with the income.

Model 7 reports the coefficients with the interactive and quadratic effects that are both significant. The marginal effect of land use regulation is thus corrected for the quadratic effect of the log per capita income. For an average property in year 2006 located in an MSA whose log per capita income is one standard deviation above (below) the mean, a unit increase in the regulation intensity leads to 5.08% (1.98%) increase in the housing price in Model 7, compared with a uniform 2.97% increase in Model 4. With the quadratic effect considered, the marginal effect of land use regulation is found more disperse geographically in Model 7 than in Model 5.

In Figure 5(a), we visualize the relationship of the log housing price, CALURI and the log GDP per capita in the benchmark model.[37] The tighter the regulation is or the higher the per capita income is, the higher the housing price.[38] In Figure 5(b), we show the same relationship with the interactive and quadratic effects. There is wide dispersion of the marginal effect of land use regulation by city income. When we approach the corner where the land use regulation is tight and the log GDP per capita is high, the increasing steepness shows the importance of the non-linear effect.

We use the top 6 most populated MSAs in California as an example to show the price dynamics.[39] The leading principal cities of these MSAs are Los Angeles, San Francisco, Riverside, San Diego, San Jose and Fresno. Figure 6 compares the dynamics of the actual price and the estimated price based on

[37] We simulate the grid points of CALURI and the log GDP per capita that are normal distributed with the mean and the standard deviation estimated from the data. The grid of each dimension is truncated at 1.64σ above and below the variable mean, so the grid points fall into the 90% confidence intervals along each dimension. We thus look at the space where a majority of the grid points lie.

[38] To construct Figure 5(a), we evaluate the parameters of CALURI and the log GDP per capita at the parameters estimated from the linearized Model 4 in the exact model solution (13). We find that the surface in Figure 5(a) is very close to a hyperplane, indicating that the estimation equation (17) based on the first-order Taylor approximation is precise enough to capture the marginal impact of land use regulation and the per capita income on the housing prices in the benchmark model.

[39] The population ranking is based on the Moody's data in 2006. We exclude Sacramento--Roseville--Arden-Arcade MSA, because the land use data from Gyourko, Saiz and Summers (2008) is not available from the leading principal city (Sacramento). As a result, our choice of the top 6 most populated MSAs are Los Angeles-Long Beach-Anaheim MSA, San Francisco-Oakland-Hayward MSA, Riverside-San Bernardino-Ontario MSA, San Diego-Carlsbad MSA, San Jose-Sunnyvale-Santa Clara MSA, Fresno MSA.

the structural estimates from Model 7. The subplots are sorted by the MSA population in 2006 in descending order. The estimated prices from our empirical model trace the actual prices closely.[40]

6. Decomposing the Regulatory Effects: Production and Amenity Channels

6.1 Measuring Production and Amenity Channels

The effect of land use regulation on housing prices can be decomposed into three channels. The first channel goes through the housing supply. We call this the production channel. The second channel goes through the housing demand. We call it the amenity channel, because the regulation protects the amenity value and increases housing demand, leading to an increase in the local housing prices. There is a third channel related to the household location choice; this is the general equilibrium (GE) channel associated with the feedback effect of housing prices on housing choice. Tighter regulation that makes housing more expensive will drive housing demand to neighboring cities.

We disentangle these three channels using our structural estimates. We decompose the responses of housing prices through these different channels to a land use regulatory change. We can rewrite the estimation equation (24) by separating the impacts of regulation as follows.

$$\begin{aligned}
\ln p_{ijmt} &= prod_{jmt} + amen_{jmt} + ge_{jmt} + [other\ terms]_{ijmt} \\
prod_{jmt} &= \theta(\delta_0 + \delta_1 z_{mt}) \cdot CALURI_j \\
amen_{jmt} &= \eta(1-\theta) \cdot CALURI_j \\
ge_{jmt} &= -\tfrac{1}{2}[\theta\lambda + \eta(1-\theta)(2\lambda - 1)] \cdot (CALURI_j - CALURI_k)
\end{aligned} \tag{25}$$

where *prod*, *amen* and *ge* stand for the production, the amenity and the GE channels respectively.[41] We define three channels in this way, because they achieve the normalization with zero mean; if the land use regulatory measure and the per capita income is evaluated at their means ($CALURI_j = CALURI_k = 0$), *prod*, *amen* and *ge* will yield zero values.

The GE channel is closely related to the spillover effect in Pollakowski and Wachter (1990) which emphasizes the interdependence of land use regulation and housing prices across regions. Our structural model picks up the effect as part of the GE channel. Pollakowski and Wachter (1990) and our model predict that tighter land use regulation in the neighboring regions increases local housing prices, as

[40] The housing boom and bust in the 2000s in the Los Angeles, San Francisco and San Jose MSAs are very well captured by our structural model. Note that the estimated prices of Fresno MSA are not as good as those in other MSAs. Our estimated price dynamics in Fresno MSA capture the shape along the time dimension, but not the level. We think the main reason is that our GMM-IV structural estimates based on cross-sectional time-series data are not indexed by MSA and year, so more weights will be assigned to bigger MSAs including Los Angeles and San Francisco MSAs.

[41] Note that *ge* is not identical to the Taylor approximated term (14); only the effect related to CALURI in (14) is included in the empirical measure of the GE channel.

regulation intensity of neighboring cities $CALURI_k$ is positively correlated with the GE channel and housing prices.

The production and the amenity channels do not include the second-order effects due to the price feedback, but separately identifies them in the GE effect. We thus conduct another decomposition that adds back the price feedback effects to construct the net production and the net amenity channels.

$$\begin{aligned} \ln p_{ijmt} &= prod_{jmt,ge} + amen_{jmt,ge} + [other\ terms]_{ijmt} \\ prod_{jmt,ge} &= \theta(\delta_0 + \delta_1 z_{mt}) \cdot CALURI_j - \tfrac{1}{2}\theta\lambda \cdot (CALURI_j - CALURI_k) \\ amen_{jmt,ge} &= \eta(1-\theta) \cdot CALURI_j - \tfrac{1}{2}\eta(1-\theta)(2\lambda - 1) \cdot (CALURI_j - CALURI_k) \end{aligned} \tag{26}$$

In Figure 7, we provide a graphical illustration of the housing price responses to the regulation increase through the (net) production and amenity channels. The housing supply curve will shift to the left through the production channel, while the housing demand curve will shift to the right through the amenity channel. The response through the net production (amenity) channel is the price response through the production (amenity) channel net of the GE effect.

6.2 Responses of Production and Amenity Channels to Land Use Regulation

Table 10 reports the responses of housing prices through each channel in response to one unit increase in land use regulation. We construct the counterfactual prices that only one channel in (25) or (26) responds to the regulatory change. The response is measured by the percentage deviation of the counterfactual price from the estimated price. We report the result by MSA, because our measure of the per capita income only varies at the MSA level.[42]

From Columns 1-3 in Table 10, the response of housing prices through the production channel is in general larger than the responses through the amenity or the GE channel. Tight regulation has the first-order effect to increase housing prices directly through the housing supply (3.22% on average). In comparison, the response through the GE channel is much smaller (-1.73%), because the effect comes from the demand feedback of housing prices. The response through the amenity channel has in general the smallest impact on housing prices (0.32%). If the production and the amenity channels take the GE effects into account, we see in Columns 4-5 in Table 10 that both effects become smaller.

The total response of housing prices to a unit increase in CALURI combines the responses of all channels. San Francisco area (4.84%), San Jose area (4.84%), Los Angeles area (3.82%) and San Diego area (3.53%) show the largest response of housing prices to a unit increase of regulation. These 4 MSAs have higher per capita income than the average MSA in California. The strong response of the housing prices in these MSAs is mainly attributed to the production channel. In these 4 MSAs, the responses

[42] We aggregate the city regulatory measure to the MSA level using the probability weight provided by Gyourko, Saiz and Summers (2008) as before.

through the net production channel are more than 50% larger than the MSA average response through the net production channel (1.55%). The price will increase through the net production channel by 4.58% in San Francisco, 4.57% in San Jose, 3.55% in Los Angeles, and 3.26% in San Diego. Our estimated response of the net amenity channel is constant by construction across MSAs. A unit increase in CALURI lead to 0.27% increase through the net amenity channel.

We can also use these results to simulate the impact of changes in regulatory regimes. Our regulatory index at the city level ranges from to -3.23 to 3.38. Los Angeles City scores the highest, while Hillsborough town scores the lowest in terms of CALURI in our sample. Using these measures to set up a counterfactual: If Los Angeles City were to relax its land use regulation to the lowest level among cities, *ceteris paribus*, housing prices could be as much as 25% lower. The production, the amenity and the GE channels contribute to -34.60%, -2.12% and 11.44% respectively.[43] [44]

Our estimated effects in San Francisco MSA are comparable to the estimated marginal effects in Quigley, Raphael and Rosenthal (2008) (QRR), because both works have a single standardized regulatory index and the questionnaire are similar. More importantly, the local survey conducted by QRR is based on the questionnaires of Qyourko, Saiz and Summers (2008) but is adapted to California jurisdictions. Table 11 shows the comparison of QRR's analysis in several aspects to ours.[45] QRR's OLS estimates of the marginal effect of regulation on housing prices range from 1.2% to 2.2% and their IV estimates range from 3.8% to 5.3%. Our GMM-IV estimators are close to QRR's IV estimators. The marginal effect through the production, the amenity, and GE channels are 6.25%, 0.32% and -1.73% respectively. If the we factor the GE effects into the first two channels, we find 4.58% and 0.27% for the production and amenity channels. We find that the total marginal effect of regulation on prices in San Francisco MSA is 4.84%.

7. The Spillover Effect of Land Use Regulation on Housing Prices

Pollakowski and Wachter (1990) using a database for a single county, Montgomery, Maryland find that the relative restrictiveness of regulation between neighboring and home cities has a positive spillover effect on the housing prices in the home city. Our dataset is larger. We use a more granular sample from California to confirm the existence and the positive impact of the city-level spillover effects. We find

[43] For the net channels incorporated with GE effects, the contributions to the price decrease are -23.48% and -1.79% for the net production and amenity effects.

[44] To calculate these price change through different channels, we use the estimated responses of Los Angeles MSA in Table 10. We multiply the responses by the size of regulatory change, 3.38 - (-3.23), to estimate the decline of housing prices attributed to different channels.

[45] QRR focus on the pricing data from the cross-section data of 2000 Census from 86 cities in San Francisco Bay area, while we have transaction cross-section time-series data from 25 cities in San Francisco MSA from 1997 to 2017. There are 10 sub-indices underlying the single index in QRR, compared to 8 sub-indices behind CALURI in our work.

that the home regulatory impact on home housing prices is stronger, once the neighboring regulatory impact is controlled. Consistent with Pollakowski and Wachter (1990), we show the previous finding holds in more recent data and more widely in the metro areas.

7.1 Measuring the Spillover Effect

While the spillover effect establishes the price interdependence of neighboring housing markets through regulation, it is different from the home regulatory effect through the general equilibrium channel in the previous analysis. The latter captures the second-order price feedback effect through the production or the amenity channels due to the spatial reallocation of housing demand. The spillover effect may capture any direct regulatory impact from the neighboring cities, in addition to the price feedback channel.

We define the relative restrictiveness index (RRI) as the difference between neighboring and home regulatory indices whose marginal effect measures the spillover effect in the section.

$$RRI_j = CALURI_{-j} - CALURI_j \tag{27}$$

We specify the functional form of the neighboring regulatory index of city *j* as the weighted average of the regulatory indices in California and consider 2 weighting measures of the neighboring indices that weigh on the city proximities.

$$CALURI_{-j} = \sum_{k \neq j} weight_{jk} \cdot CALURI_k \tag{28}$$

$$\begin{gathered} \text{inv. sq. distance: } weight_{jk} = x_{invdist2} / d_{jk}^2 \\ \text{gravity: } weight_{jk} = x_{gravity} z_j z_k / d_{jk}^2 \end{gathered} \tag{29}$$

where $x_{invdist2}$, and $x_{gravity}$ are constants to make sure that the sum of the weights is equal to 1.[46] The second case generalizes the first one and takes a gravitational form. The gravity model puts weight on the per capita income of the home and neighboring cities, adjusted by the distance.

7.2 Additional Data

Our estimations in the previous sections exclude the discussion of the spillover effect due to limited data availability.[47] As the spillover effect is local among the neighboring cities that are geographically close,

[46] Alternatively, we also test the inverse distance to weigh the neighboring indices. Compared with the case of inverse squared distance, the alternative case puts less weight on the neighboring cities closer to the home city.

[47] We make the decision to use more data to produce more precise estimates and to exclude the spillover effect in the estimations in the previous sections where the per capita income varies only at the metro level. The choice may raise the concern of downward bias of the home regulatory impact. As will be shown in this section, we find a negative correlation between the regulatory index and the relative restrictiveness index. However, we find the issue is minor in the previous estimates for the following reason.

The decomposition of the regulatory effects in (25) shows that the negation of the GE channel takes a form similar to our definition of the spillover effect, so previous estimates do partially take into account the effect of relative restrictiveness index. The difference is that the neighboring regulatory index $CALURI_k$ is not varying by city but one needs to control city-level variation of the per capita income to identify the spillover effect in the metro areas. There is no series of per capita income that covers the whole sample period from 1993 to 2017 at such granular level. An additional data issue is the low response rate of the Wharton Land Use Survey in some MSAs. Among the most populated MSAs, only San Diego-Carlsbad MSA has a response rate of jurisdictions that exceeds 50% (11 out of 18 cities).[48] The construction of RRI which relies on geo-spatial information may be severely biased towards the cities responding to the Survey.

To overcome the data issue, we additionally collect census tract data from the tract-block Summary File of the 2014 American Community Survey (ACS) 5-year estimates. The 5-year survey spans from 2010 to 2014 but the estimates do not represent any single year in the range.[49] We calculate the city-level per capita income by averaging the tract-level median income per capita and using the tract population as the weight.

To match the time frame of the income data, the empirical analysis in the section will use the property transactions in California in 2014. We thus exclude the variables that don't exhibit cross-sectional variations to prevent collinearity problem.[50] The independent variables include the first and second order terms of the log per capita income and structural characteristics of housing in the benchmark estimation (Model 4 in Table 7). We select four major MSAs that are the least likely to suffer the data issue of low response rates in the Survey and have not too small numbers of cities within the metro area (Los Angeles-Long Beach-Anaheim MSA, San Francisco-Oakland-Hayward MSA, San Diego-Carlsbad MSA, Oxnard-Thousand Oaks-Ventura MSA, with LA, SF, SD and VT respectively for short notations).[51]

In Figure 8, we show the distribution of the CALURI and RRI. RRI under different weight measures in (29) show similar distributional patterns, with the bell shapes and the two-sided fat tails. In Figure 9, we show the scatter diagrams of CALURI and RRI by city. There is a strong negative

normalized to 0 for all cities. Because the mean of CALURI is zero by construction, the assumption of zero neighboring regulatory index is thus a special case where equal weight is assigned to all cities, regardless of the distance. The assumption turns out to make the regulatory estimates more robust for MSAs with low survey response rates.

[48] In the appendix, we report the response rate of cities by CBSA (MSA and μMSA) in the Wharton Land Use Survey.

[49] The first wave of the tract level data is 2009 ACS 5-year estimates, but we use the wave of 2014 ACS 5-year estimates to exclude any unobservable consequence of the Great Recession on the housing market. 2014 ACS 5-year estimates is the wave that is closest to the time of the Wharton Land Use Survey with no single year falling into the Great Recession.

[50] The excluded independent variables in the section are the growth rate of the household mortgage debt, the real 30-year fixed-rate mortgage rate, and the log of population-weighted mean GDP per capita of California.

[51] To choose MSAs, we set the following criteria: (1) there are at least 10 cities in an MSA covered by the Wharton Land Use Survey; (2) an MSA has more than 1 principal city based on the definition in the historical delineation files of metropolitan and micropolitan statistical areas (2006) from the Census Bureau; (3) more than 50% of the leading principal cities (listed in the name of an MSA) are covered by the Survey. Three MSAs survive the criteria: Los Angeles-Long Beach-Anaheim MSA, San Francisco-Oakland-Hayward MSA, and San Diego-Carlsbad MSA (For San Francisco MSA, it is long known as San Francisco–Oakland–Fremont MSA until 2013). We additionally add Oxnard-Thousand Oaks-Ventura MSA as another case. Based on the appendix Table A2, the share of cities and the share of principal cities covered by the Survey are both high among MSAs.

correlation between CALURI and RRI and the negative relationship is robust under different weighting measures (-0.92, -0.93 respectively). We separately mark the cities in the four selected MSAs (LA, SF, SD, VT) and show that the negative correlation still holds within each metro area.

7.3 Estimation of the Spillover Effect

In Tables 12a-d, we report the estimated home and neighboring regulatory effects for the four selected metro areas (Los Angeles MSA, San Francisco MSA, San Diego MSA, Oxnard MSA respectively). We report three model specifications in each table. Similar to the method adopted by Pollakowski and Wachter (1990), we use Ordinary Least Square in the estimations.[52] Model 1 in Table 12 includes the home regulatory impact but excludes the spillover term in the estimations, while Models 2 and 3 add the relative restrictiveness indices under 2 different weighting measures in (29). Our cross-sectional estimations can explain 43%-61% of the log price variations, depending on the model specifications and the MSAs.

By estimating the regulatory impacts using the city-level per capita income, we find in Model 1 that the marginal effect of regulation on housing prices are qualitatively similar to the estimated regulatory effects shown in Table 10. This specification has the interpretation of equal weights assigned to all cities available in the Wharton Land Use Survey, regardless of the geographical distance (see footnote 47).

Models 2 and 3 build on Model 1 by considering the neighboring regulatory impact and apply the inverse squared distance and the gravity weight respectively as the weighting measures to the neighboring cities. General results hold for all models. The home and neighboring regulatory effects will be both significantly positive for all of the 4 selected MSAs.

The comparison of Model 2 to Model 1 shows that the marginal impacts of land use regulation will be bigger if the relative restrictiveness index is controlled in the log housing price equations. The result follows naturally from the fact that CALURI and RRI are negatively correlated and omitting RRI in the estimation in Model 1 will downward bias the estimated coefficients of CALURI. We see large spatial variation in the estimated regulatory and spillover impacts. In Model 2, the home and neighboring regulatory effects on the log housing prices (referenced to the average city in the metro area) are 14.7% and 8.78% in Los Angeles MSA, 6.00% and 4.10% in San Francisco MSA, 25.7% and 10.3% in San Diego MSA, and 6.07% and 8.05% in Oxnard MSA.

If we build on Model 2 and further take the per capita income of the neighboring cities into account in Model 3, the estimated regulatory and spillover effects are both larger in all of 4 selected MSAs. In

[52] Instrumenting the per capita income with city-level demographic variables (using the mean population age and share of high education aggregated from the tract level) won't qualitatively change the estimated regulatory and spillover effect of the selected MSAs.

Model 3, the regulatory and the spillover effects on the log housing prices are 18.2% and 12.4% in Los Angeles MSA, 7.15% and 5.12% in San Francisco MSA, 26.2% and 10.6% in San Diego MSA, and 6.79% and 8.79% in Oxnard MSA.

8. Conclusion

In this paper, we develop a general equilibrium framework to determine the impact of land use regulation on housing prices in cities in California over the years 1993 to 2017. We use housing transaction prices and housing characteristics along with data on macro credit supply and regional per capita income together with the Wharton Residential Land Use Survey (Gyourko, Saiz and Summer, 2008) to identify the impacts of land use regulation on housing prices.

We identify the separate channels through which land use regulation can impact housing prices. Specifically, we characterize the production channel which measures the increasing cost of housing production and the amenity channel which measures the increase in environmental attractiveness of communities with greater land use regulation. While the empirical literature discusses these channels, the literature does not measure these effects in a general equilibrium framework. In addition, we show the general equilibrium effects of mitigating housing price impacts through households' location choice response to higher prices. Our estimated effects show that Los Angeles is the city whose housing prices are most impacted by regulation. In our calculations, if land use regulation in LA were to be decreased to the level observed in the least regulated cities, housing prices would decline by approximately 25%. Besides, we take a more granular view to examine the regulatory interdependence among cities and to estimate the spillover effects of regulation. We define the relative restrictiveness indices as the difference between the neighboring and home regulatory effects and report robust finding on the significant and positive spillover effects on housing prices.

Reference

American Institute of Planners. (1976). *Survey of State Land Use Planning Activity*. Report to the U.S. Department of Housing and Urban Development. Washington

Albouy, D., & Ehrlich, G. (2015, January). The distribution of urban land values: Evidence from market transactions. In *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association* (Vol. 108, pp. 1-32). National Tax Association.

Albouy, D., & Ehrlich, G. (2018). Housing productivity and the social cost of land-use restrictions. *Journal of Urban Economics*. 107, 101-120.

Brueckner, J. K. (1987). The structure of urban equilibria: A unified treatment of the Muth-Mills model. *Handbook of regional and urban economics*, 2, 821-845.

Brueckner, J. K. (1990). Growth controls and land values in an open city. *Land Economics*, 66(3), 237-248.

Brueckner, J. K. (1995). Strategic control of growth in a system of cities. *Journal of Public Economics*, 57(3), 393-416.

Buist, H. (1991). The Wharton Urban Decentralization Database. University of Pennsylvania. Unpublished paper.

Choi, H. S., Hong, H., Kubik, J., & Thompson, J. (2016). Sand states and the US housing crisis. Working paper.

Engle, R., Navarro, P., & Carson, R. (1992). On the theory of growth controls. Journal of Urban Economics, 32(3), 269-283.

Fischel, W. A., & Fischel, W. A. (1995). *Regulatory takings: Law, economics, and politics*. Harvard University Press.

Fischel, F. A. (2001). *The Homevoter Hypothesis* (Vol. Cambridge, MA.).

Fluck, T. A. (1986). *Euclid v. Ambler: a retrospective*. Journal of the American Planning Association, 52(3), 326-337.

Glaeser, E. L., & Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of economic literature*, 47(4), 983-1028.

Glaeser, E. L., Gyourko, J., & Saks, R. E. (2005a). Why have housing prices gone up? *American Economic Review*, *Papers and Proceedings*. 95(2), 329-333.

Glaeser, E. L., Gyourko, J., & Saks, R. (2005b). Why is Manhattan so expensive? Regulation and the rise in housing prices. *The Journal of Law and Economics*, 48(2), 331-369.

Glaeser, E. L., Gyourko, J., & Saks, R. E. (2005c). Urban growth and housing supply. *Journal of economic geography*, 6(1), 71-89.

Glaeser, E. L., & Gyourko, J. (2006). Housing dynamics (No. w12787). National Bureau of Economic Research.

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of economic geography*, 1(1), 27-50.

Glaeser, E. L., & Ward, B. A. (2009). The causes and consequences of land use regulation: Evidence from Greater Boston. *Journal of Urban Economics*, 65(3), 265-278.

Goetzmann, W. N., Spiegel, M., & Wachter, S. M. (1998). Do cities and suburbs cluster?. *Cityscape*, 193-203.

Glickfeld, M., & Levine, N. (1992). Regional growth, local reaction: A review of empirical evidence on the effectiveness and efficiency of local government land use regulation. Cambridge MA: *Lincoln Institute of Land Policy*.

Graham, J. (2018), House prices and consumption: a new instrumental variables approach. Working paper.

Gyourko, J., & Molloy, R. (2015). Regulation and housing supply. In *Handbook of regional and urban economics* (Vol. 5, pp. 1289-1337). Elsevier.

Gyourko, J., Saiz, A., & Summers, A. (2008). A new measure of the local regulatory environment for housing markets: The Wharton Residential Land Use Regulatory Index. *Urban Studies*, 45(3), 693-729.

Hamilton, B. W. (1975). Zoning and property taxation in a system of local governments. *Urban studies*, 12(2), 205-211.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 1029-1054.

Ihlanfeldt, K. R. (2007). The effect of land use regulation on housing and land prices. *Journal of Urban Economics*, 61(3), 420-435.

Jackson, K. (2016). Do land use regulations stifle residential development? Evidence from California cities. *Journal of Urban Economics*, 91, 45-56.

Jackson, K. (2018). Regulation, land constraints, and California's boom and bust. *Regional Science and Urban Economics*, 68, 130-147.

Kahn, M. E. (2011). Do liberal cities limit new housing development? Evidence from California. *Journal of Urban Economics*, 69(2), 223-228.

Kok, N., Monkkonen, P., & Quigley, J. M. (2014). Land use regulations and the value of land and housing: An intra-metropolitan analysis. *Journal of Urban Economics*, 81, 136-148.

Landvoigt, T., Piazzesi, M., & Schneider, M. (2015). The housing market (s) of San Diego. *American Economic Review*, 105(4), 1371-1407.

Linneman, P., Summers, A., Brooks, N., & Buist, H. (1990). The state of local growth management. Wharton Real Estate Center Working Paper, 81.

Malpezzi, S. (1996). Housing prices, externalities, and regulation in US metropolitan areas. *Journal of Housing Research*, 209-241.

Malpezzi, S., Chun, G. H., & Green, R. K. (1998). New place‐to‐place housing price indexes for US Metropolitan Areas, and their determinants. *Real Estate Economics*, 26(2), 235-274.

Pollakowski, H. O., & Wachter, S. M. (1990). The effects of land-use constraints on housing prices. *Land economics*, 66(3), 315-324.

Quigley, J. M., & Rosenthal, L. A. (2005). The effects of land use regulation on the price of housing: What do we know? What can we learn? *Cityscape*, 69-137.

Quigley, J. M., Raphael, S., & Rosenthal, L. A. (2008). Measuring land-use regulations and their effects in the housing market.

Quigley, J. M., & Swoboda, A. M. (2007). The urban impacts of the Endangered Species Act: A general equilibrium analysis. *Journal of Urban Economics*, 61(2), 299-318.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy*, 90(6), 1257-1278

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics*, 74-104.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3), 1253-1296.

Sarah Mawhorter and Carolina Reid (2018). Terner California Residential Land Use Survey. Berkeley, CA: University of California, Berkeley.

Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, & Matthew Sobek. IPUMS USA: Version 8.0 American Community Survey. Minneapolis, MN: IPUMS, 2018.

Turner, M. A., Haughwout, A., & Van Der Klaauw, W. (2014). Land use regulation and welfare. *Econometrica*, 82(4), 1341-1403.

Tables

Table 1. Sample Coverage by Geographical Cities

	City	County	CBSA	Count
Land Use Sample	179	39	25	5,318,379
Unmatched Sample	963	47	25	7,403,052

Table 2: Summary Statistics of Land Use Regulation Indices

	Mean	Median	Std.Dev	Pct.25	Pct.75
LPPI	0.47	0.11	1.08	-0.31	1.09
LZAI	1.87	2	0.61	1	2
LPAI	1.69	1	0.98	1	2
DRI	0.15	0	0.35	0	0
OSI	0.87	1	0.33	1	1
EI	0.93	1	0.26	1	1
SRI	0.19	0	0.77	0	0
ADI	9.04	8.06	4.51	5.67	12.13
CALURI	0.27	-0.01	1.23	-0.41	0.6
WRLURI	0.8	0.55	0.79	0.16	1.5

Note: local political pressure index (LPPI), local zoning approval index (LZAI), local project approval index (LPAI), density restriction index (DRI), open space index (OSI), exactions index (EI), supply restriction index (SRI), approval delay index (ADI). California Land Use Regulation Index (CALURI), Wharton Residential Land Use Regulation Index (WRLURI). Frequency weights of the property transactions are used. Source: Gyourko, Saiz and Summer (2008) and authors' calculation.

Table 3. Distribution of Residential Property Use

	Land Use Sample		Unmatched Sample	
Property Type	Frequency	Percent	Frequency	Percent
Single Family Residential	4,045,001	31.80	6,200,178	48.74
Townhouse	13,401	0.11	31,418	0.25
Cluster Home	39,918	0.31	45,049	0.35
Condominium	1,133,241	8.91	951,460	7.48
Cooperative	859	0.01	323	0.00
Row House	336	0.00	702	0.01
Planned Unit Development	84,951	0.67	159,699	1.26
Inferred Single Family Residential	672	0.01	14,223	0.11
Total	5,318,379	100.00	7,403,052	100.00

Note: the total sample is the non-foreclosed residential sales transactions in California from 1993 to 2017. Source: ZTRAX and authors' calculation. ZTRAX database is provided by Zillow Group. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group or any of its affiliates.

Table 4. Summary Statistics of Property Characteristics

	Mean	Median	Std.Dev	Pct.25	Pct.75
Land Use Sample					
Sales Price	369,615	282,102	620,425	169,943	453,920
Sq.Ft.	1,699.40	1,503.00	858.78	1,162.00	2,011.00
Price/Sq.Ft	221.27	181.26	518.6	115.82	283.93
Age of Property	30	26	24.56	9	46
No.of Bathroom	2	2	0.81	2	2
No.of Bedrooms	3.03	3	1.04	2	4
Miles to Core Cities	28.08	8.14	240.19	4.44	14.5
Unmatched Sample					
Sales Price	352,330	270,609	643,300	165,749	427,337
Sq.Ft.	1,778.34	1,574.00	1,048.22	1,217.00	2,128.00
Price/Sq.Ft	199.64	164.88	761.11	108.91	250.08
Age of Property	27.8	24	23.13	8	44
No.of Bathroom	2.05	2	0.8	2	2
No.of Bedrooms	3.16	3	0.95	3	4
Miles to Core Cities	52.34	10.99	362.95	5.83	20.65

Note: Sales Price and Price/Sq.Ft are inflation adjusted to Jan. 2006 US dollars, using the Consumer Price Index for All Urban Consumers: Housing (FRED: CPIHOSNS). Source: ZTRAX and authors' calculation. ZTRAX database is provided by Zillow Group. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group or any of its affiliates.

Table 5. Summary Statistics of Instrumental Variables

	Mean	Median	Std.Dev	Pct.25	Pct.75
share of high education (%)	35.92	35.2	8.02	29.12	42.10
population age	34.48	34.3	2.22	32.72	36.27
share of high-tech jobs (%)	6.84	5.37	5.90	2.94	8.11

Note: variables are weighted by the MSA population. Source: American Community Survey, Moody's Analytics.

Table 6. Correlation Matrix: Instrumental Variables

	GDP pca	L.GDP pca	high educ %	high-tech %	pop. age
GDP pca	1.000				
L.GDP pca	0.992	1.000			
high educ %	0823	0.820	1.000		
high-tech %	0.651	0.627	0.706	1.000	
pop. age	0.753	0.762	0.905	0.405	1.000

Note: all variables are in log form. Correlation is weighted by the MSA population. Source: American Community Survey, Moody's Analytics.

Table 7. Benchmark Estimation: Coefficients

	Model 1 GMM	Model 2 GMM	Model 3 GMM-IV	Model 4 GMM-IV
CALURI	0.0195***	0.0293***	0.0290***	0.0297***
	(0.000)	(0.000)	(0.000)	(0.000)
log GDP per capita	1.231***	1.326***	1.311***	1.291***
	(0.001)	(0.001)	(0.001)	(0.001)
log Avg. GDP per cap	0.496***	0.352***	0.369***	0.432***
	(0.005)	(0.004)	(0.004)	(0.004)
Bedroom: 1		-0.120***	-0.120***	-0.129***
		(0.003)	(0.003)	(0.003)
Bedroom: 2		-0.291***	-0.293***	-0.300***
		(0.003)	(0.003)	(0.003)
Bedroom: 3		-0.389***	-0.391***	-0.405***
		(0.003)	(0.003)	(0.003)
Bedroom: 4+		-0.453***	-0.455***	-0.471***
		(0.003)	(0.003)	(0.003)
Bathroom: 1		0.134***	0.135***	0.107***
		(0.006)	(0.006)	(0.006)
Bathroom: 2		0.209***	0.211***	0.169***
		(0.006)	(0.006)	(0.006)
Bathroom: 3		0.161***	0.165***	0.115***
		(0.006)	(0.006)	(0.006)
Bathroom: 4+		0.303***	0.308***	0.264***
		(0.007)	(0.007)	(0.007)
log sq.feet		1.084***	1.084***	1.107***
		(0.001)	(0.001)	(0.001)
log miles to core cities		-0.0262***	-0.0261***	-0.0314***
		(0.000)	(0.000)	(0.000)
SFR		-0.0576***	-0.0595***	-0.0709***
		(0.001)	(0.001)	(0.001)
condominium		0.0217***	0.0219***	0.0233***
		(0.001)	(0.001)	(0.001)
Age: 1-5		0.132***	0.133***	0.115***
		(0.001)	(0.001)	(0.001)
Age: 6-10		0.0847***	0.0848***	0.0695***
		(0.001)	(0.001)	(0.001)
Age: 11-20		0.0652***	0.0656***	0.0530***
		(0.001)	(0.001)	(0.001)
Age: 21-30		0.0576***	0.0585***	0.0413***
		(0.001)	(0.001)	(0.001)
Age: 31-40		0.108***	0.110***	0.0937***
		(0.001)	(0.001)	(0.001)
Age: 41-50		0.127***	0.129***	0.117***
		(0.001)	(0.001)	(0.001)
Age: > 50		0.143***	0.147***	0.137***
		(0.001)	(0.001)	(0.001)
growth rate of mortgage debt	3.072***	3.025***	3.024***	2.881***
	(0.008)	(0.006)	(0.006)	(0.006)
30-year FRM rate	-4.003***	-3.190***	-3.153***	-2.591***
	(0.049)	(0.040)	(0.042)	(0.042)
Constant	5.806***	-1.825***	-1.830***	-2.115***
	(0.021)	(0.021)	(0.022)	(0.022)
Observations	5,259,215	5,259,215	5,259,215	5,259,215

Note: robust standard errors in the parentheses. * p<0.10, ** p<0.05, *** p<0.010. The base levels of the factor variables are: no bedroom, no bathroom, property use other than single-family and condominium, new property (age is zero). The lag terms of log real GDP per capita and log mean GDP per capita in California are used as IVs of their contemporaneous terms in Models 3-4; the share of high education, the population age and the share of high-jobs are additional IVs of Model 4.

Table 8. Marginal Effect of Sub-indices on log Housing Price

	Model 1 GMM	Model 2 GMM	Model 3 GMM-IV	Model 4 GMM-IV	Contribution (%) sum to 100%
LPPI	0.0082	0.0122	0.0121	0.0124	21.06
LZAI	0.0068	0.0103	0.0102	0.0104	17.68
LPAI	0.0080	0.0121	0.0119	0.0122	20.76
DRI	0.0023	0.0035	0.0034	0.0035	5.94
OSI	0.0050	0.0075	0.0074	0.0076	12.85
EI	0.0029	0.0044	0.0044	0.0045	7.61
SRI	0.0029	0.0043	0.0043	0.0044	7.41
ADI	0.0026	0.0039	0.0039	0.0040	6.70

Note: local political pressure index (LPPI), local zoning approval index (LZAI), local project approval index (LPAI), density restriction index (DRI), open space index (OSI), exactions index (EI), supply restriction index (SRI), approval delay index (ADI). All sub-indices have been standardized to zero mean and unit variance. The marginal effect of a sub-index is the marginal effect of CALURI on the log housing prices multiplied by the sub-index weight in the predicted score regression. The control variables and the estimation method can be found in Table 7.

Table 9. Estimation with Non-Linear Effects: Coefficients

	Model 4 GMM-IV	Model 5 GMM-IV	Model 6 GMM-IV	Model 7 GMM-IV
CALURI	0.0297***	-0.0577***	0.0341***	-0.267***
	(0.000)	(0.004)	(0.000)	(0.004)
log GDP per	1.291***	1.293***	-6.440***	-6.758***
capita	(0.001)	(0.001)	(0.030)	(0.031)
Avg.log GDP per	0.432***	0.426***	0.356***	0.343***
capita	(0.004)	(0.004)	(0.004)	(0.004)
CALURI*log GDP		0.0221***		0.0760***
per capita		(0.001)		(0.001)
log GDP per			1.008***	1.049***
Capita squared			(0.004)	(0.004)
Observations	5,259,215	5,259,215	5,259,215	5,259,215

Note: robust standard errors in the parentheses. * p<0.10, ** p<0.05, *** p<0.010. The lag terms of the log real GDP per capita and log mean GDP per capita in California are used as IVs of their contemporaneous terms; the share of high education, the population age and the share of high-jobs are additional IVs of Models 4-7. The control variables and the estimation method can be found in Table 7.

Table 10. Counterfactual Experiments: Responses to +SD CALURI (% deviation)

	production	amenity	GE	production with GE	amenity with GE	total
Bakersfield	3.22	0.32	-1.73	1.55	0.27	1.81
Chico	3.23	0.32	-1.73	1.56	0.27	1.82
Fresno	3.12	0.32	-1.73	1.45	0.27	1.72
Hanford-Corcoran	0.04	0.32	-1.73	-1.63	0.27	-1.37
Los Angeles-Long Beach-Anaheim	5.23	0.32	-1.73	3.55	0.27	3.82
Madera	1.64	0.32	-1.73	-0.04	0.27	0.23
Merced	0.35	0.32	-1.73	-1.32	0.27	-1.06
Modesto	2.02	0.32	-1.73	0.34	0.27	0.61
Napa	4.74	0.32	-1.73	3.07	0.27	3.33
Oxnard-Thousand Oaks-Ventura	3.37	0.32	-1.73	1.70	0.27	1.96
Redding	4.20	0.32	-1.73	2.52	0.27	2.79
Riverside-San Bernardino-Ontario	1.67	0.32	-1.73	0.00	0.27	0.27
Sacramento-Roseville-Arden-Arcade	4.89	0.32	-1.73	3.21	0.27	3.48
Salinas	2.99	0.32	-1.73	1.32	0.27	1.58
San Diego-Carlsbad	4.94	0.32	-1.73	3.26	0.27	3.53
San Francisco-Oakland-Hayward	6.25	0.32	-1.73	4.58	0.27	4.84
San Jose-Sunnyvale-Santa Clara	6.25	0.32	-1.73	4.57	0.27	4.84
San Luis Obispo-Paso Robles-Arroyo Grande	4.26	0.32	-1.73	2.59	0.27	2.86
Santa Cruz-Watsonville	3.57	0.32	-1.73	1.90	0.27	2.16
Santa Maria-Santa Barbara	4.52	0.32	-1.73	2.85	0.27	3.11
Santa Rosa	3.91	0.32	-1.73	2.23	0.27	2.50
Stockton-Lodi	2.04	0.32	-1.73	0.37	0.27	0.64
Vallejo-Fairfield	2.14	0.32	-1.73	0.47	0.27	0.73
Visalia-Porterville	0.80	0.32	-1.73	-0.87	0.27	-0.60
Yuba City	1.12	0.32	-1.73	-0.55	0.27	-0.29
mean	3.22	0.32	-1.73	1.55	0.27	1.81

Note: the numbers reported by MSA are the time average percentage deviations of the counterfactual prices from the estimated prices, for the period from 1993 to 2017. The estimated parameters from Model 7 are used to construct the counterfactual prices. The price dynamics of MSAs in bold type (most populated MSAs in 2006 with leading principal cities available in the Wharton Survey) are plotted in Figure 6.

Table 11. Comparison with Quigley, Raphael and Rosenthal (2008)

Quigley, Raphael and Rosenthal (2008)				
Location	San Francisco Bay Area			
Number of cities	86			
Source of Price data	2000 US Census			
Regulatory Index	BLURI (from Berkeley Land Use Survey)			
Number of sub-indices	10			
Estimation method	OLS and IV			
Results	OLS		IV	
Marginal effect of regulation	1.2%-2.2%		3.8%-5.3%	
This paper				
Location	San Francisco-Oakland-Hayward, MSA			
Number of cities	25			
Source of Price data	ZTRAX, 1993-2017			
Regulatory Index	CALURI (from Wharton Residential Land Use Survey)			
Number of sub-indices	8			
Estimation method	GMM-IV			
Results (GE separated)	production	amenity	GE	total
Marginal effect of regulation	6.25%	0.32%	-1.73%	4.84%
Results (GE incorporated)	production with GE	Amenity with GE		total
Marginal effect of regulation	4.58%	0.27%		4.84%

Table 12a. Spillover Effect: Los Angeles-Long Beach-Anaheim, MSA

Variable	(1) Benchmark	(2) Inv.dist2	(3) Gravity
CALURI	0.0595***	0.147***	0.182***
	(0.0015)	(0.0093)	(0.0100)
RRI		0.0878***	0.124***
		(0.0091)	(0.0099)
Adjusted R^2	0.563	0.564	0.565
N	52,102	52,102	52,102

Note: robust standard errors in the parentheses. * p<0.10, ** p<0.05, *** p<0.010. CALURI = California Land Use Regulation Index; RRI = Relative Restrictiveness Index. *Inv.dist2* uses the inverse distance squared to weigh neighboring CALURI. Gravity indicates the specification with the city-level income per capita divided by the squared distance as the weight. Omitted control variables in all specifications include log city-level per capita income where a property is located and its squared term, the number of bedrooms, the number of bathrooms, the log distance to the Central Business District (centroid of the nearest core city of an MSA), the log size of a property, the property use (single-family, condominium) and the property age. We use the housing transactions in 2014 from ZTRAX. The data of the city-level per capita income is aggregated from the census tract data from the Summary File of the 5-year American Community Survey 2010-2014.

Table 12b. Spillover Effect: San Francisco-Oakland-Hayward, MSA

Variable	(1) Benchmark	(2) Inv.dist2	(3) gravity
CALURI	0.0158*	0.0600***	0.0715***
	(0.0081)	(0.017)	(0.016)
RRI		0.0410***	0.0512***
		(0.014)	(0.013)
Adjusted R^2	0.510	0.511	0.511
N	19,137	19,137	19,137

Note: robust standard errors in the parentheses. * p<0.10, ** p<0.05, *** p<0.010. CALURI = California Land Use Regulation Index; RRI = Relative Restrictiveness Index. *Inv.dist2* uses the inverse distance squared to weigh neighboring CALURI. Gravity indicates the specification with the city-level income per capita divided by the squared distance as the weight. Omitted control variables in all specifications include log city-level per capita income where a property is located and its squared term, the number of bedrooms, the number of bathrooms, the log distance to the Central Business District (centroid of the nearest core city of an MSA), the log size of a property, the property use (single-family, condominium) and the property age. We use the housing transactions in 2014 from ZTRAX. The data of the city-level per capita income is aggregated from the census tract data from the Summary File of the 5-year American Community Survey 2010-2014.

Table 12c. Spillover Effect: San Diego-Carlsbad, MSA

Variable	(1) Benchmark	(2) Inv.dist2	(3) Gravity
CALURI	0.125***	0.257***	0.262***
	(0.0076)	(0.023)	(0.021)
RRI		0.103***	0.106***
		(0.018)	(0.016)
Adjusted R^2	0.604	0.605	0.605
N	21,985	21,985	21,985

Note: robust standard errors in the parentheses. * p<0.10, ** p<0.05, *** p<0.010. CALURI = California Land Use Regulation Index; RRI = Relative Restrictiveness Index. *Inv.dist2* uses the inverse distance squared to weigh neighboring CALURI. Gravity indicates the specification with the city-level income per capita divided by the squared distance as the weight. Omitted control variables in all specifications include log city-level per capita income where a property is located and its squared term, the number of bedrooms, the number of bathrooms, the log distance to the Central Business District (centroid of the nearest core city of an MSA), the log size of a property, the property use (single-family, condominium) and the property age. We use the housing transactions in 2014 from ZTRAX. The data of the city-level per capita income is aggregated from the census tract data from the Summary File of the 5-year American Community Survey 2010-2014.

Table 12d. Spillover Effect: Oxnard-Thousand Oaks-Ventura, MSA

Variable	(1) Benchmark	(2) Inv.dist2	(3) Gravity
CALURI	-0.0191**	0.0607***	0.0679***
	(0.0089)	(0.022)	(0.022)
RRI		0.0805***	0.0879***
		(0.018)	(0.019)
Adjusted R^2	0.429	0.431	0.431
N	6,272	6,272	6,272

Note: robust standard errors in the parentheses. * p<0.10, ** p<0.05, *** p<0.010. CALURI = California Land Use Regulation Index; RRI = Relative Restrictiveness Index. *Inv.dist2* uses the inverse distance squared to weigh neighboring CALURI. Gravity indicates the specification with the city-level income per capita divided by the squared distance as the weight. Omitted control variables in all specifications include log city-level per capita income where a property is located and its squared term, the number of bedrooms, the number of bathrooms, the log distance to the Central Business District (centroid of the nearest core city of an MSA), the log size of a property, the property use (single-family, condominium) and the property age. We use the housing transactions in 2014 from ZTRAX. The data of the city-level per capita income is aggregated from the census tract data from the Summary File of the 5-year American Community Survey 2010-2014.

Figures

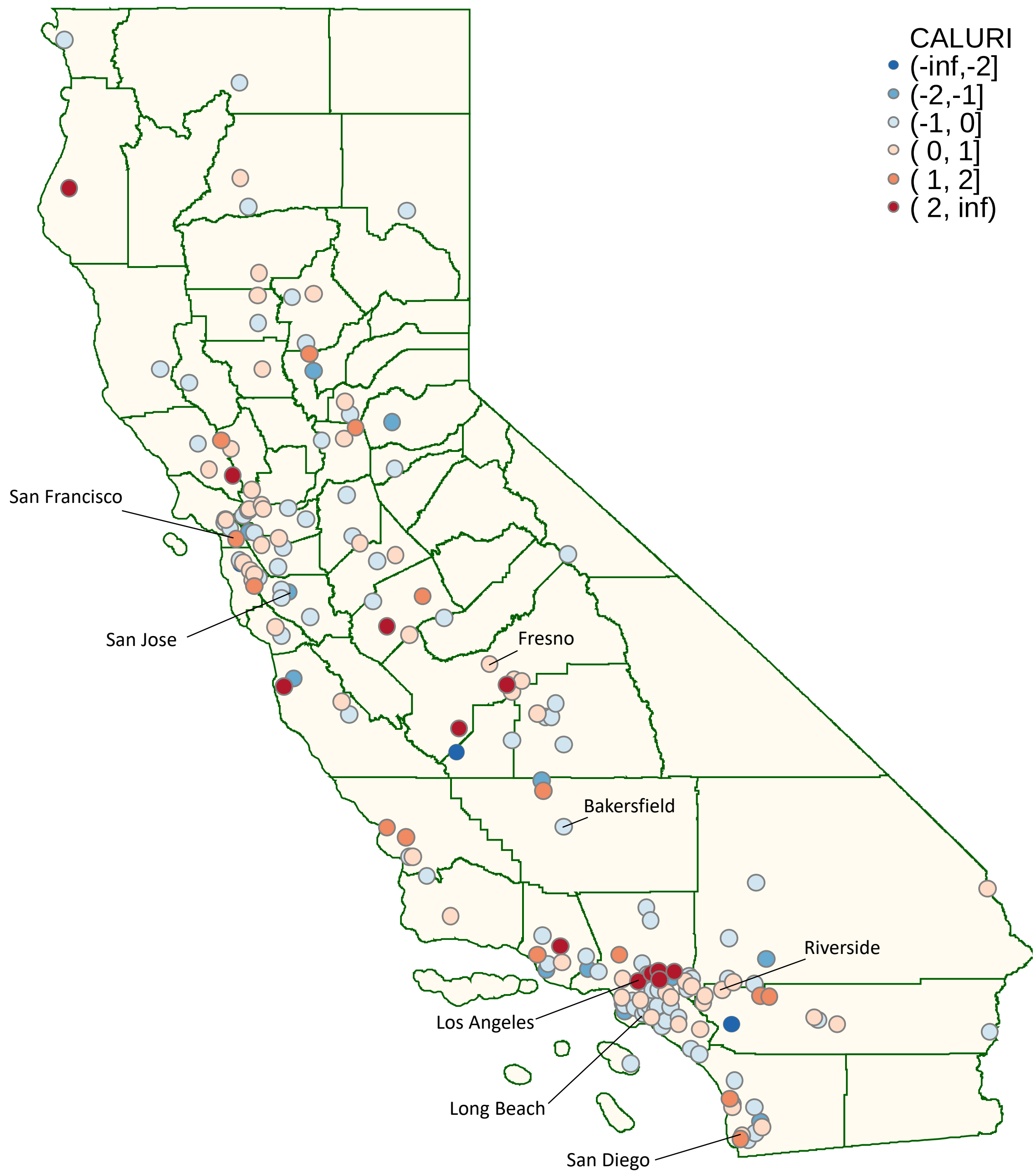


Figure 1: spatial distribution of land use regulation intensity in California. California Land Use Regulation Index (CALURI) is based on the sub-indices from WRLURI. A higher index value indicates higher regulation intensity. There are 185 jurisdictions in total.
Source: Gyourko, Saiz and Summers (2008) and authors' calculation.

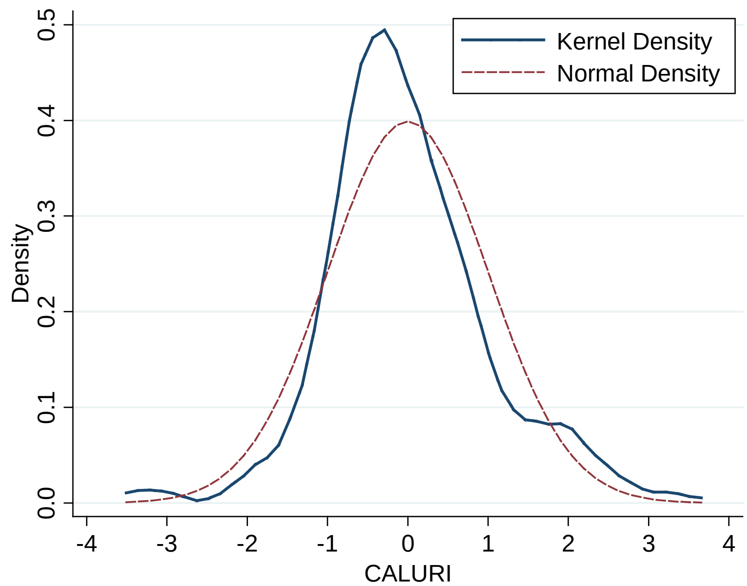
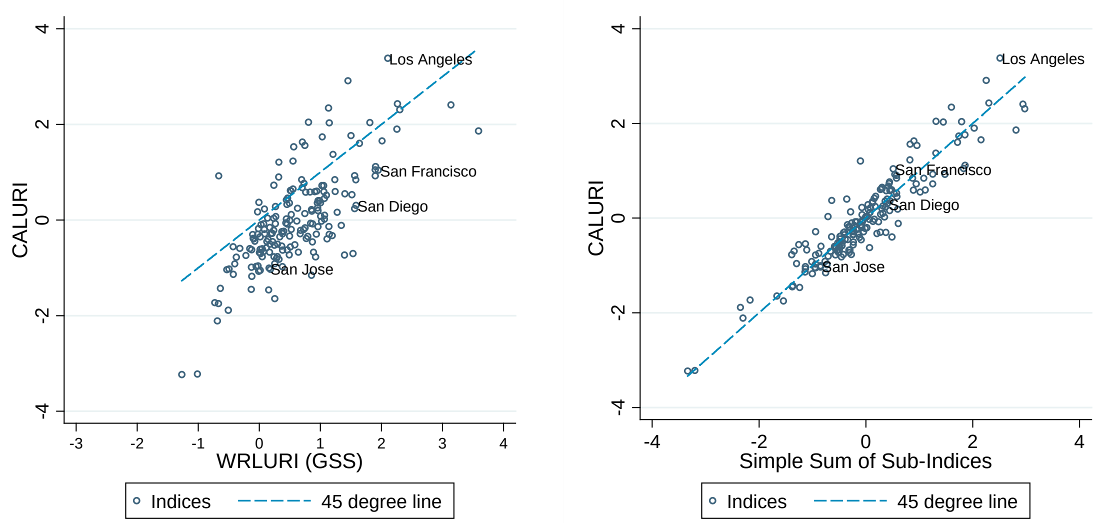


Figure 2: comparison of the kernel density of California Land Use Regulation Index (CALURI) and the normal density. CALURI is based on the sub-indices from WRLURI. A higher index value indicates higher regulation intensity.

Source: Gyourko, Saiz and Summers (2008) and authors' calculation.



(a) CALURI vs WRLURI (b) CALURI vs Simple Sum of Sub-indices

Figure 3: quantile-quantile plots of WRLURI, CALURI and Simple Sum of Sub-indices. We compare the index based on the first factor of the principal factor analysis with the simple sum of the 8 sub-indices underlying CALURI. For comparability, we normalize the sub-indices and their sum, so all indices in comparison have zero mean and unit variance.

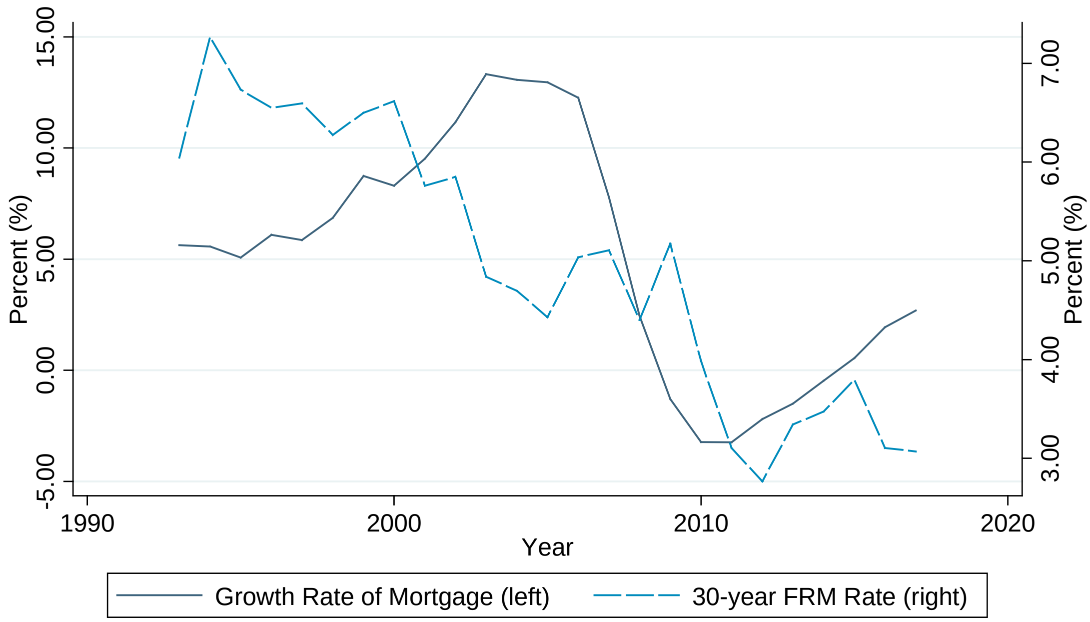
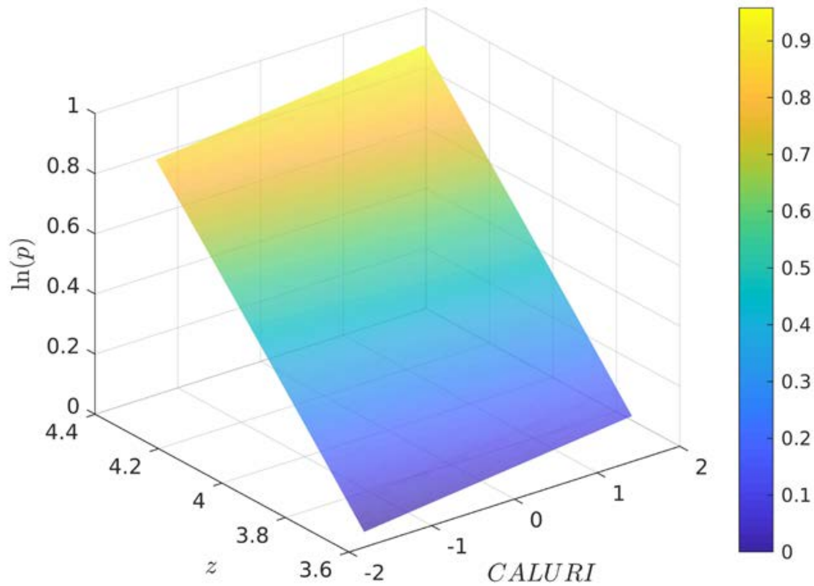
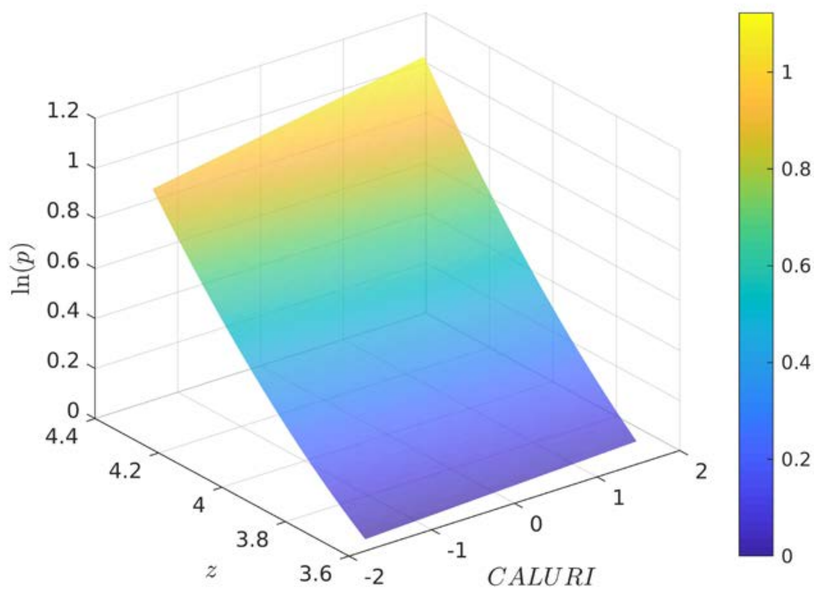


Figure 4: Annual growth rate of the residential mortgage debt of US households and 30-year US average fixed-rate mortgage rate. The mortgage rate has been adjusted for inflation. Source: Z.1 Financial Account Table from the Board of Governors of Federal Reserves and Freddie Mac.



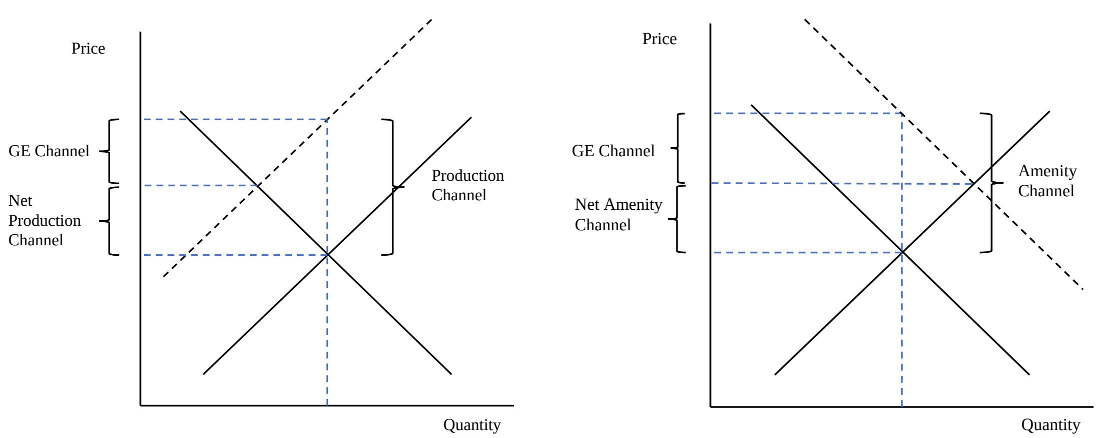
Panel (a): no interactive or quadratic term



Panel (b) with the interactive term and the quadratic term of the log GDP per capita

Figure 5: the log housing price as the function of the log GDP per capita (z) and land use regulation intensity (CALURI). The grid of each dimension is simulated using normal distribution, with the mean and the standard deviation estimated from the data. Grid points within 90% confidence intervals along each dimension are plotted. The parameters are evaluated at the estimated values of Model 4 in panel (a) and Model 7 in panel (b). The min value along the z-axis is normalized to 0.

Figure 6: housing price dynamics of 6 MSAs in California: actual price vs estimated price. The estimation is based on Model 7. The prices are aggregated by year and MSA.



(a) Production Channel only (b) Amenity Channel only

Figure 7: graphical illustration of the (net) production and amenity channels. The shift of the housing supply or the demand curve is triggered by an increase in land use regulation.

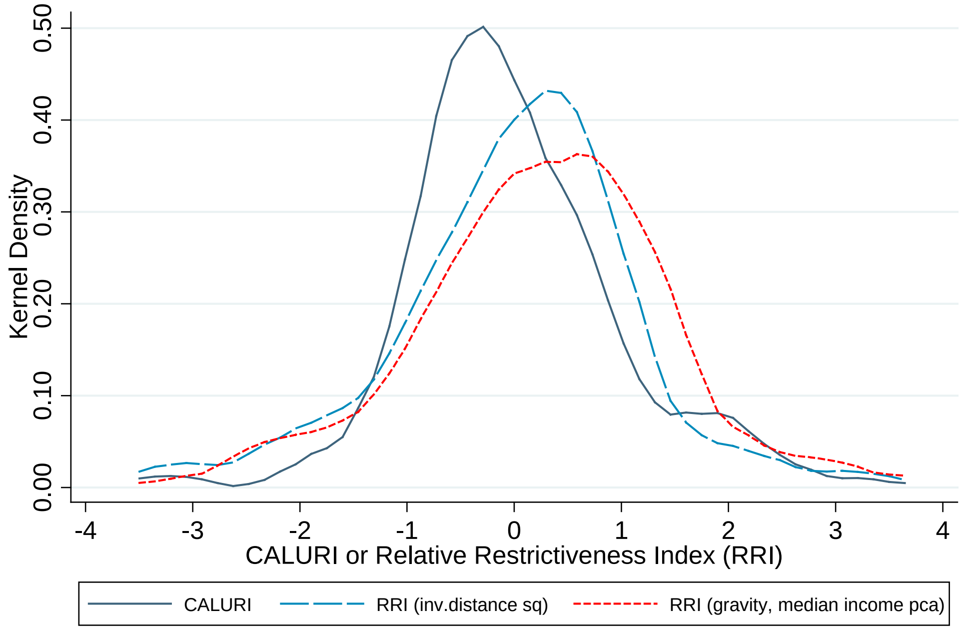
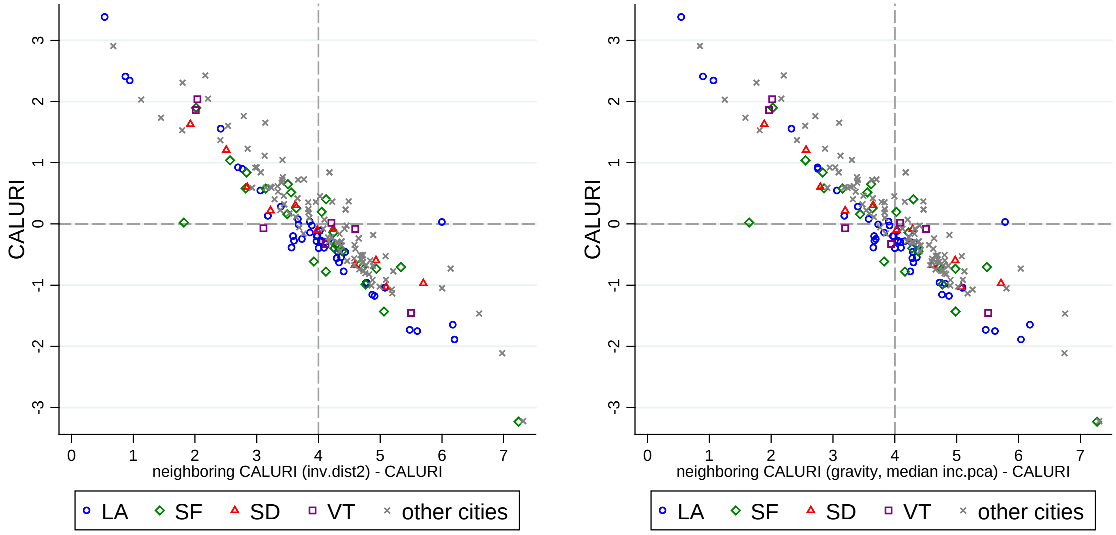


Figure 8: kernel density of CALURI and relative restrictiveness indices (RRI). RRI is defined as the difference between the neighboring regulatory index and CALURI of the city. We report three ways of constructing the neighboring regulatory index, with the weight indicated in the parentheses.



(a) weight: inverse distance sq.

(b) weight: gravity (per capita income*inv. distance sq.)

Figure 9: CALURI vs relative restrictiveness index. Panels (a) and (b) show the scatter plots using different weights in the construction of the neighboring regulatory index. The relative restrictiveness index (RRI) of a city is defined as the difference between the neighboring regulatory index and CALURI of the city. We rescale RRI to the positive real line with the same mean (4) for comparability. We separately mark 4 MSAs (LA = Los Angeles-Long Beach-Anaheim MSA; SF = San Francisco-Oakland-Hayward MSA; SD = San Diego-Carlsbad MSA; VT = Oxnard-Thousand Oaks-Ventura MSA) that are large in terms of and population and the number of cities, and that have high survey response rates in the Wharton Residential Land Use survey (Gyourko, Saiz and Summers, 2008).

Appendix

A.1 Proof of Uniqueness of the Equilibrium

First, rewrite the market clearing condition of city *j* as follows.

$$q_j = b_j r_j (q_j)^{\frac{1}{1-\theta}}, \text{ where } b_j = \frac{A_0^{\frac{1}{1-\theta}}}{\alpha Y_0 Z_j^{\phi} \tau_j^{\eta}} \left(\frac{\theta}{c_j} \right)^{\frac{\theta}{1-\theta}} \tag{30}$$

We express r_j as a function of q_j. The equilibrium condition of location choices (5) can be written as

$$q_j x = Z_j^{\phi} \tau_j^{\eta} r_j (q_j)^{-\alpha}, \text{ where } x = \sum_{k \in S} Z_k^{\phi} \tau_k^{\eta} r_k^{-\alpha} \tag{31}$$

Combine two equations and eliminate r_j.

$$q_j(x) = Z_j^{\frac{\phi}{\alpha(1-\theta)+1}} \tau_j^{\frac{\eta}{\alpha(1-\theta)+1}} b_j^{\frac{\alpha(1-\theta)}{\alpha(1-\theta)+1}} x^{-\frac{1}{\alpha(1-\theta)+1}} \tag{32}$$

For an arbitrary *n*, we can prove that there is a unique set of moving probabilities that solve the system of equations. We can solve *x* from the following equation.

$$\sum_{k \in S} q_j(x) = 1 \tag{33}$$

LHS of (33) is a strictly decreasing function of *x*, while RHS is a weakly decreasing function of *x*. There is a unique solution to the equation. Given *x*, we can use (33) to fully solve the set of moving probabilities.

For the special case of $n = 2$, we can solve the model. With $q_j + q_k = 1$ and $S = \{j, k\}$,

$$q_j = \frac{(Z_j^{\phi} \tau_j^{\eta})^{1-\lambda} b_j^{\lambda}}{(Z_j^{\phi} \tau_j^{\eta})^{1-\lambda} b_j^{\lambda} + (Z_k^{\phi} \tau_k^{\eta})^{1-\lambda} b_k^{\lambda}}, \text{ where } \lambda = \frac{\alpha(1-\theta)}{\alpha(1-\theta)+1} \tag{34}$$

Combined with (30), the log housing price can be expressed in the linear form (13).

A.2 CALURI by MSA and City

Table A1. City and CALURI

MSA and City	CALURI	MSA and City	CALURI
Bakersfield	0.291	Signal Hill city	-0.203
McFarland city	1.735	Redondo Beach city	-0.245
Bakersfield city	-0.308	Pico Rivera city	-0.279
Delano city	-1.052	Lakewood city	-0.279
Chico	0.190	Tustin city	-0.284
Orland city	0.721	La Palma city	-0.289
Paradise town	0.527	Palmdale city	-0.297
Willows city	-0.163	Claremont city	-0.302
Gridley city	-0.288	Los Alamitos city	-0.351
Chico city	-0.343	Commerce city	-0.385
Fresno	1.032	Whittier city	-0.389
Huron city	2.908	South Pasadena city	-0.396
Selma city	2.429	Lancaster city	-0.455
Kingsburg city	0.841	La Canada Flintridge city	-0.459
Fresno city	0.452	Avalon city	-0.544
Parlier city	0.369	Hermosa Beach city	-0.561
Reedley city	0.236	Alhambra city	-0.631
Hanford-Corcoran	-1.280	Calabasas city	-0.775
Corcoran city	-0.508	Carson city	-0.962
Avenal city	-2.112	Huntington Beach city	-0.975
Los Angeles-Long Beach-Anaheim	-0.195	La Habra city	-1.042
Los Angeles city	3.382	Agoura Hills city	-1.157
Glendora city	2.408	Palos Verdes Estates city	-1.178
El Monte city	2.342	Covina city	-1.648
San Fernando city	1.558	Montebello city	-1.730
Irvine city	0.924	Santa Ana city	-1.751
Seal Beach city	0.897	Baldwin Park city	-1.889
Brea city	0.546	Arcadia city	NA
Pomona city	0.322	San Marino city	NA
Compton city	0.280	**Madera**	-0.772
La Habra Heights city	0.131	Mammoth Lakes town	-0.623
El Segundo city	0.077	Chowchilla city	-0.772
Rancho Santa Margarita city	0.037	**Merced**	0.830
Beverly Hills city	0.032	Los Banos city	2.046
Anaheim city	-0.008	Merced city	1.231
Dana Point city	-0.025	Dos Palos city	0.728
San Clemente city	-0.115	Gustine city	-0.081
Gardena city	-0.142	**Modesto**	-0.036
Fountain Valley city	-0.198	Waterford city	0.458
Long Beach city	-0.198	Ceres city	-0.684

Table A1. City and CALURI (continued)

MSA and City	CALURI	MSA and City	CALURI
Napa	0.414	Rancho Cordova city	0.070
Calistoga city	1.114	West Sacramento city	-0.353
St. Helena city	0.363	Rocklin city	-0.510
American Canyon city	0.242	Placerville city	-1.072
Oxnard-Thousand Oaks-Ventura	0.254	**Salinas**	-0.294
Santa Paula city	2.037	Carmel-by-the-Sea city	2.031
San Buenaventura (Ventura) city	1.861	Soledad city	0.226
Camarillo city	0.020	Greenfield city	-0.914
Oxnard city	-0.071	Seaside city	-1.466
Ojai city	-0.081	**San Diego-Carlsbad**	-0.253
Simi Valley city	-0.327	Encinitas city	1.630
Port Hueneme city	-1.453	Coronado city	1.207
Redding	-0.307	Del Mar city	0.599
Shasta Lake city	0.173	San Diego city	0.303
Anderson city	-0.584	El Cajon city	0.217
Weed city	-0.768	Vista city	-0.086
Riverside-San Bernardino-Ontario	-0.081	Lemon Grove city	-0.102
Beaumont city	1.761	National city	-0.596
Banning city	1.654	Poway city	-0.676
Rancho Mirage city	0.921	Solana Beach city	-0.972
Riverside city	0.842	Santee city	-1.035
Coachella city	0.675	**San Francisco-Oakland-Hayward**	-0.219
Needles city	0.617	Portola Valley town	1.899
Chino city	0.590	San Francisco city	1.040
Corona city	0.419	Belmont city	0.839
Loma Linda city	0.402	Redwood city	0.648
Norco city	0.353	Hercules city	0.582
Palm Desert city	-0.180	San Leandro city	0.578
Yucaipa city	-0.236	Larkspur city	0.515
Chino Hills city	-0.287	Woodside town	0.402
Blythe city	-0.299	Martinez city	0.256
Colton city	-0.599	Corte Madera town	0.196
Montclair city	-0.625	San Ramon city	0.159
Barstow city	-0.674	Burlingame city	0.022
Hesperia city	-0.745	Mill Valley city	-0.139
Big Bear Lake city	-1.136	Fremont city	-0.338
Canyon Lake city	-3.222	Brentwood city	-0.397
Sacramento-Roseville-Arden-Arcade	-0.001	Pittsburg city	-0.450
Folsom city	1.370	Millbrae city	-0.614
Lincoln city	0.112	Dublin city	-0.664

Table A1. City and CALURI (continued)

MSA and City	CALURI	MSA and City	CALURI
Sausalito city	-0.700	Santa Maria city	-0.519
Menlo Park city	-0.703	**Santa Rosa**	0.653
Pinole city	-0.732	Sonoma city	2.309
Piedmont city	-0.778	Rohnert Park city	0.719
San Pablo city	-0.987	Windsor town	-0.027
Emeryville city	-1.430	**Stockton-Lodi**	-0.110
Hillsborough town	-3.232	Ripon city	0.592
San Jose-Sunnyvale-Santa Clara	-0.657	Jackson city	-0.219
Campbell city	-0.158	Manteca city	-0.407
Santa Clara city	-0.605	Lodi city	-0.769
Morgan Hill city	-0.824	**Vallejo-Fairfield**	0.187
San Jose city	-1.007	Benicia city	0.187
San Luis Obispo-Paso Robles-Arroyo Grande	0.531	**Visalia-Porterville**	-0.292
San Luis Obispo city	1.603	Visalia city	0.606
Morro Bay city	1.046	Exeter city	-0.060
Arroyo Grande city	0.590	Woodlake city	-0.079
Grover Beach city	-0.526	Farmersville city	-0.674
Santa Cruz-Watsonville	-0.036	Porterville city	-0.806
Scotts Valley city	0.358	**Yuba City**	0.849
Capitola city	-0.731	Live Oak city	1.532
Santa Maria-Santa Barbara	-0.158	Williams city	0.922
Buellton city	0.098	Yuba city	-1.026

Note: MSAs are sorted in alphabetic order. Within each MSA, cities are sorted by CALURI in descending order. CALURI is defined as the first factor using the principal factor analysis. 8 sub-indices that have city-level variations from the Wharton Residential Land Use Survey are used: local political pressure index (LPPI), local zoning approval index (LZAI), local project approval index (LPAI), density restriction index (DRI), open space index (OSI), exactions index (EI), supply restriction index (SRI), approval delay index (ADI). Source: Gyourko, Saiz and Summer (2008) and authors' calculation.

Table A2. Survey Response Rates by CBSA in California

	City and Town			Principal City		
CBSA (MSA/μMSA)	CA	GSS	%	CA	GSS	%
Bakersfield	11	3	27	1	1	100
Chico	5	3	60	1	1	100
Clearlake	2	1	50	1	0	0
Crescent City	1	1	100	1	1	100
El Centro	7	0	0	1	0	0
Eureka-Arcata-Fortuna	7	1	14	3	1	33
Fresno	15	6	40	1	1	100
Hanford-Corcoran	4	2	50	2	1	50
Los Angeles-Long Beach-Anaheim	122	48	39	25	13	52
Madera	2	1	50	1	0	0
Merced	6	4	67	1	1	100
Modesto	9	2	22	1	0	0
Napa	5	3	60	1	0	0
Oxnard-Thousand Oaks-Ventura	10	7	70	4	3	75
Red Bluff	3	1	33	1	0	0
Redding	3	2	67	1	0	0
Riverside-San Bernardino-Ontario	52	20	38	9	3	33
Sacramento--Roseville--Arden-Arcade	19	6	32	5	2	40
Salinas	12	4	33	1	0	0
San Diego-Carlsbad	18	11	61	4	2	50
San Francisco-Oakland-Hayward	65	25	38	12	4	33
San Jose-Sunnyvale-Santa Clara	17	4	24	7	2	29
San Luis Obispo-Paso Robles-Arroyo Grande	7	4	57	2	1	50
Santa Cruz-Watsonville	4	2	50	2	0	0
Santa Maria-Santa Barbara	8	2	25	3	1	33
Santa Rosa	9	3	33	2	0	0
Sonora	1	0	0	0	0	0
Stockton-Lodi	7	3	43	1	0	0
Susanville	1	1	100	1	1	100
Truckee-Grass Valley	3	0	0	2	0	0
Ukiah	4	1	25	1	1	100
Vallejo-Fairfield	7	1	14	2	0	0
Visalia-Porterville	8	5	63	2	2	100
Yuba City	4	2	50	1	1	100
Total	458	179	39	103	43	42

Note: the list of Core Based Statistical Areas (CBSA) includes both MSAs and μMSAs. There are 482 jurisdictions in California, with 458 tied to the CBSA codes in California. “CA” and “GSS” counts the total number of cities and towns in California (CA) and in the sample of Gyourko, Saiz and Summers (2008) (GSS) respectively. The columns with “%” calculate the city share of GSS sample in California. The definition of the principal cities is based on the historical delineation files of the Principal cities of metropolitan and micropolitan statistical areas (2006) from the Census Bureau. The definition of CBSA is based on 2010 Geographic Terms and Concepts from the Census Bureau.

A.3 Data Filtering and Construction of ZTRAX Variables

The Whole ZTRAX database consists of two parts: ZTrans (transaction data) and ZAsmt (assessment data) that can be linked by a unique parcel ID. For most states, the sample prior to 2005 are scarce; for California, the database can trace back to transactions as early as 1993. I first restrict the sample to the transaction with the sales prices more than 5,000 US dollars in California. California data before 1993 (inclusive) is extremely sparse, so our ZTRAX data starts from 1993:M1 and ends in 2017:M6. For the other US states, the quality of data before 2005 is generally worse than that after the 2005. California data allows us to examine the housing prices and property characteristics in a much longer horizon.

We keep residential properties only and drop any commercials, manufactural, and foreclosure sales. Based on the Property Use Standard Code and Assessment Land Use Standard Code, we identify and focus on the residential types including single family residentials, townhouses, cluster homes, condominiums, cooperatives, planned unit developments and those inferred as single family residentials by Zillow. A transaction can involve multiple parcels, we focus on transactions with a single parcel only. We only keep the transactions that can be linked to the housing properties in the assessment data. About 89% of the transactions are matched to the assessment files.

The data fields we use from the housing data include: transaction date, geographic location (county, city, CBSA, address longitude and latitude), the sales prices, the number of bedrooms, the number of bathrooms, the year a property was built, the square foot of a property and the miles to the nearest core cities. There are other housing characteristics available in the database, but they are in general not commonly populated.

There is no separate field to directly observe the size of a property, so we construct the field as follows. We are able to observe the following fields relevant to the size of a property: building area living, building area finished, effective building area, gross building area, building area adjusted, building area total, building area finished living, base building area, heated building area. To take the maximum of the fields above and define it as the square footage of a property.

The miles of a property to the nearest core cities is constructed as follows. We first identify the CBSA where a property is located. We use the leading principal cities listed in the name of an MSA and geocode the city centers using the application program interface (API) of Google Map. We calculate the great-circle distance in miles from each property to the center of each leading principal city in the CBSA and define the minimum as the distance to the principal city. A small number of cities are not assigned to any CBSA. We thus geocode the distance from the properties in each of the cities to the nearest leading principal cities in all CBSAs in California using the API of Google Map.[53] We assign these cities to the nearest MSAs, so they don't fall out of sample in the analysis.

[53] 6 cities whose fips county codes don't fall in any MSA in California are assigned to the nearest metropolitan statistical area. They are Jackson City, Williams City, Orland City, Willows City, Mammoth Lakes Town, and Weed City.

The number of annual transactions in California ranges from 100,000 to 600,000, depending on the year. There are about 13 million transactions in total from about 1,400 cities available to be matched to the Wharton Land Use Survey data.

A.4 Auxiliary Regression

We log-linearize the definition identity of amenity demand as follows.

$$A_j = Z_j^{\phi-1}\tau_j^{\eta} \Leftrightarrow z_j = -\frac{\eta}{\phi-1}\ln\tau_j + \ln A_j \tag{35}$$

We exogenously estimate the elasticity of per capita income Z_j with respect to τ_j for an additional moment condition in the estimation. Amenity is unobservable, so it is treated as the error plus a constant term. Our data points are 25 MSAs in year 2006 and the regression analysis is cross-sectional. In Table A2, we report three specifications. Model 1 include CALURI as the only independent variable. Model 2 add three more variables: the share of high-tech jobs from the regional dataset of Moody's Analytics collected from BLS and BEA, the mean household age from American Community Survey (ACS) Public Use microdata, and the share of high education (college + graduate study) from ACS microdata. These three factors are highly correlated with per capita income. We show their correlation in Table 6.

Model 3 include more controls based on Model 2. Data on the net migrants (in thousand) and total population (in thousand) come from the regional data set of Moody's Analytics collected from the Census Bureau. Data on employment (in thousand) comes from Moody's Analytics collected from BLS (CES and QCEW). The minority share is the fraction of non-white individuals surveyed in ACS microdata. The cost-of-doing-business index is provided by Moody's Analytics. The index is the weighted average of unit labor costs, energy costs, tax burden and office rents. It is an index that standardizes the US average to 100.

We find that the coefficients of CALURI is close to zero. The insignificance of the coefficient is probably due to the small size of the MSAs. We use the estimate from Model 3 to construct the following condition for estimation.

$$\eta = 0.00331\cdot(\phi - 1) \tag{36}$$

Table A3. Auxiliary Regression of log GDP per capita

	Model 1	Model 2	Model 3
CALURI	0.0176	0.0131	-0.00331
	(0.097)	(0.055)	(0.028)
share of high-tech job		0.00747	-0.00167
		(0.008)	(0.005)
log household age		-0.190	0.313
		(0.949)	(0.576)
share of high education		2.275**	-0.0173
		(0.808)	(0.470)
net migrant			0.000664
			(0.002)
log population			-1.001***
			(0.189)
log employment			1.012***
			(0.177)
business cost index			-0.00128
			(0.004)
minority share			0.166
			(0.202)
Constant	3.855***	3.728	3.837
	(0.050)	(3.148)	(2.350)
Observations	25	25	25

Note: robust standard errors in the parentheses. * p<0.10, ** p<0.05, *** p<0.010.

A.5 Structural Parameter Estimates

In Table A4, we report the estimation of the structural parameters. Without housing characteristics properly controlled in Model 1, we tend to underestimate θ by 33%, but to overestimate α and ϕ by 49% and 1.5% respectively, compared to the estimated values in Model 2. GMM-IV estimations produce comparable estimated parameters in Models 3 and 4. Compared with Model 2, Model 4 which treats contemporaneous per capita income as endogenous yield bigger estimated values of ϕ and θ. We find that $\theta = 0.045$ and $\phi = 1.803$ in Model 4, while $\theta = 0.043$ and $\phi = 1.753$ in Model 2. The estimation in Model 4 indicates that the income elasticity of amenity demand is 0.803 (or 1.803 - 1). That is, 1% increase in the per capita income increases the amenity demand by 0.803% on average.[54]

Table A4. Benchmark Estimation: Structural Parameters

	Model 1 GMM	Model 2 GMM	Model 3 GMM-IV	Model 4 GMM-IV
θ	0.029***	0.043***	0.042***	0.045***
	(0.000)	(0.000)	(0.000)	(0.000)
λ	0.787***	0.710***	0.720***	0.751***
	(0.002)	(0.002)	(0.002)	(0.002)
ϕ	1.779***	1.753***	1.754***	1.803***
	(0.005)	(0.004)	(0.004)	(0.004)
α	3.811***	2.554***	2.681***	3.152***
	(0.048)	(0.040)	(0.042)	(0.042)
η	0.003***	0.002***	0.002***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)

Note: robust standard errors in the parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.010$. The lag terms of log real GDP per capita and log mean GDP per capita in California are used as IVs of their contemporaneous terms in Models 3-4; the share of high education, the population age and the share of high-jobs are additional IVs of Model 4.

In Table A5, we replicate the GMM estimations in Table 5, but instead use WALURI as the regulatory index instead. Compared to CALURI constructed only from the subsample of California cities, WALURI is estimated nationally from more than 2,000 jurisdictions.

In Table A6, we report the estimates of the structural parameters under four model specifications. The average income elasticity of amenity demand is adjusted upward from 0.803 in the benchmark to the 1.030 in the fully extended model. The coefficient of the quadratic term ϕ_2 is positive, indicating that the income elasticity of amenity demand increases with income.

[54] In the model, the parameters are not free to take any value on the real line. In the GMM or GMM-IV estimations, we solve the minimization problems without parameter constraints, so our estimations of α may fall out of the unit interval.

Table A5. Estimation with WRLURI

	Model 1 GMM	Model 2 GMM	Model 3 GMM-IV	Model 4 GMM-IV
WRLURI	0.109***	0.117***	0.117***	0.124***
	(0.000)	(0.000)	(0.000)	(0.000)
log GDP per capita	1.185***	1.295***	1.280***	1.265***
	(0.001)	(0.001)	(0.001)	(0.001)
log Avg. GDP per cap	0.556***	0.405***	0.427***	0.482***
	(0.005)	(0.004)	(0.004)	(0.004)
Bedroom: 1		-0.0670***	-0.0677***	-0.0639***
		(0.003)	(0.003)	(0.003)
Bedroom: 2		-0.230***	-0.231***	-0.223***
		(0.003)	(0.003)	(0.003)
Bedroom: 3		-0.318***	-0.320***	-0.315***
		(0.003)	(0.003)	(0.003)
Bedroom: 4+		-0.381***	-0.383***	-0.380***
		(0.003)	(0.003)	(0.003)
Bathroom: 1		0.107***	0.107***	0.0712***
		(0.006)	(0.006)	(0.006)
Bathroom: 2		0.185***	0.187***	0.138***
		(0.006)	(0.006)	(0.006)
Bathroom: 3		0.136***	0.139***	0.0843***
		(0.006)	(0.006)	(0.006)
Bathroom: 4+		0.272***	0.277***	0.229***
		(0.006)	(0.006)	(0.006)
log sq.feet		1.066***	1.065***	1.082***
		(0.001)	(0.001)	(0.001)
log miles to core cities		-0.0236***	-0.0235***	-0.0292***
		(0.000)	(0.000)	(0.000)
SFR		-0.0402***	-0.0420***	-0.0555***
		(0.001)	(0.001)	(0.001)
condominium		0.0118***	0.0118***	0.00944***
		(0.001)	(0.001)	(0.001)
Age: 1-5		0.134***	0.134***	0.120***
		(0.001)	(0.001)	(0.001)
Age: 6-10		0.0858***	0.0860***	0.0740***
		(0.001)	(0.001)	(0.001)
Age: 11-20		0.0645***	0.0649***	0.0556***
		(0.001)	(0.001)	(0.001)
Age: 21-30		0.0544***	0.0553***	0.0408***
		(0.001)	(0.001)	(0.001)
Age: 31-40		0.103***	0.104***	0.0898***
		(0.001)	(0.001)	(0.001)
Age: 41-50		0.122***	0.124***	0.114***
		(0.001)	(0.001)	(0.001)
Age: > 50		0.124***	0.127***	0.118***
		(0.001)	(0.001)	(0.001)
growth rate of mortgage debt	3.057***	3.006***	3.000***	2.865***
	(0.008)	(0.006)	(0.006)	(0.006)
30-year FRM rate	-3.894***	-3.150***	-3.064***	-2.553***
	(0.049)	(0.040)	(0.041)	(0.041)
Constant	5.658***	-1.916***	-1.945***	-2.178***
	(0.021)	(0.021)	(0.021)	(0.021)
Observations	5,259,215	5,259,215	5,259,215	5,259,215

Note: robust standard errors in the parentheses. * p<0.10, ** p<0.05, *** p<0.010. The base levels of the factor variables are: no bedroom, no bathroom, property use other than single-family and condominium, new property (age is zero). The lag terms of log real GDP per capita and log mean GDP per capita in California are used as IVs of their contemporaneous terms in Models 3-4; the share of high education, the population age and the share of high-jobs are additional IVs of Model 4.

Table A6. Estimation with Non-Linear Effects: Structural Parameters

	Model 4 GMM-IV	Model 5 GMM-IV	Model 6 GMM-IV	Model 7 GMM-IV
θ	0.045***	0.045***	0.048***	0.049***
	(0.000)	(0.028)	(0.000)	(0.000)
λ	0.751***	0.748***	0.684***	0.678***
	(0.002)	(0.001)	(0.002)	(0.002)
α	3.152***	3.106***	2.273***	2.212***
	(0.033)	(0.092)	(0.033)	(0.033)
δ_0	1	-0.944***	1	-5.124***
		(0.385)		(0.097)
δ_1	0	0.489***		1.540***
		(0.097)	(0.000)	(0.024)
ϕ_0	0	0	16.748***	17.458***
			(0.066)	(0.067)
ϕ_1	1.803***	1.800***	-6.389***	-6.748***
	(0.005)	(0.053)	(0.032)	(0.033)
ϕ_2	0	0	1.059***	1.104***
			(0.004)	(0.004)
ϕ_{avg}	1.803***	1.800***	2.032***	2.030***
	(0.005)	(0.053)	(0.005)	(0.005)
η	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)

Note: robust standard errors in the parentheses. * p<0.10, ** p<0.05, *** p<0.010.

A.6 Quantifying the Contribution of Production and Amenity Channel to Housing Prices

We evaluate the contribution of the production, the amenity, and the general equilibrium channels at the MSA level, because our measure of the per capita income only varies by MSA. We aggregate the city-level regulatory measure using the probability weight provided by Gyourko, Saiz and Summers (2008). For any MSA-year combination, we calculate the levels of the estimated housing prices P_{mt} and the counterfactual prices that exclude the production, the amenity or the GE channels $P_{mt,-prod}$, $P_{mt,-amen}$ and $P_{mt,-ge}$ as follows.

$$\begin{aligned}
P_{mt} &= \exp[E_{ij}(\ln p_{ijmt})] \\
P_{mt,-prod} &= \exp[E_{ij}(\ln p_{ijmt}) - E_j(prod_{jmt})] \\
P_{mt,-amen} &= \exp[E_{ij}(\ln p_{ijmt}) - E_j(amen_{jmt})] \\
P_{mt,-ge} &= \exp[E_{ij}(\ln p_{ijmt}) - E_j(ge_{jmt})]
\end{aligned} \tag{37}$$

E_{ij} denotes the empirical mean aggregating households and cities. For the counterfactual price excluding the production channel (hereafter, counterfactual production price), we interpret it as the price that normalizes the production effect to the mean but keeps everything else constant. For the counterfactual price excluding the amenity channel (hereafter, counterfactual amenity price), we interpret it as the price that normalizes the amenity effect to the mean but keeps everything else constant. For the counterfactual price excluding the GE channel (hereafter, counterfactual GE price), we can interpret it as the price that normalizes the GE effect to the mean but keeps everything else constant.

To evaluate the production effect of regulation, we conduct the experiments that exclude the production channel. The result by MSA is reported in Column 1 of Table A7. The percentage deviation of an MSA measures the size of the production effect. For an MSA with the average per capita income, a positive (negative) deviation indicates how much the price will increase (decrease) due to the production effect if the cost of housing supply counterfactually increases (decreases) from a below-mean (above-mean) level to the mean.

To evaluate the amenity effect of regulation, we conduct the experiments that exclude the amenity channel. The result by MSA is reported in Column 2 of Table A7. The percentage deviation of an MSA measures the size of amenity effect. For an MSA with the average per capita income, a positive (negative) deviation indicates how much the price will increase (decrease) due to the amenity effect if the amenity level increases (decreases) from a below-mean (above-mean) level to the mean.

To evaluate the GE effect of regulation, we conduct the experiments that exclude the GE channel. The result by MSA is reported in Column 3 of Table A7. The percentage deviation of an MSA measures the size of GE effect. For an MSA with average per capita income, a positive (negative) deviation indicates how much the price will increase (decrease) due to the GE effect if households are counterfactually not moving out (in) for higher utility.

Similarly, we also calculate the levels of the counterfactual prices that exclude the production with GE effect $P_{mt,-prod,ge}$, and the amenity channel with GE effect $P_{mt,-amen,ge}$ as follows.

$$\begin{aligned} P_{mt,-prod,ge} &= \exp[E_{ij}(\ln p_{ijmt}) - E_j(prod_{jmt,ge})] \\ P_{mt,-amen,ge} &= \exp[E_{ij}(\ln p_{ijmt}) - E_j(amen_{jmt,ge})] \end{aligned} \tag{38}$$

The production and the amenity channels with GE effects are reported in Columns 4-5 in Table A7.

Table A7. Counterfactual Experiments: Size of the Channels

	Prices excluding			Prices excluding	
MSA	production	amenity	GE	production with GE	amenity with GE
Bakersfield	1.02	0.10	-0.55	0.49	0.08
Chico	0.33	0.03	-0.18	0.16	0.03
Fresno	-1.61	-0.17	0.90	-0.75	-0.14
Hanford-Corcoran	-0.01	0.33	-1.73	-1.69	0.27
Los Angeles-Long Beach-Anaheim	-3.78	-0.24	1.27	-2.60	-0.19
Madera	1.07	0.23	-1.23	-0.13	0.19
Merced	-0.50	-0.46	2.51	1.91	-0.38
Modesto	0.86	0.14	-0.73	0.14	0.11
Napa	-1.72	-0.12	0.66	-1.10	-0.10
Oxnard-Thousand Oaks-Ventura	-0.74	-0.07	0.39	-0.36	-0.06
Redding	1.50	0.11	-0.60	0.90	0.09
Riverside-San Bernardino-Ontario	-0.43	-0.07	0.39	-0.06	-0.06
Sacramento-Roseville-Arden-Arcade	-0.76	-0.05	0.28	-0.49	-0.04
Salinas	-0.30	-0.04	0.20	-0.11	-0.03
San Diego-Carlsbad	-1.10	-0.07	0.39	-0.73	-0.06
San Francisco-Oakland-Hayward	-0.95	-0.05	0.26	-0.69	-0.04
San Jose-Sunnyvale-Santa Clara	6.02	0.30	-1.60	4.37	0.25
San Luis Obispo-Paso Robles-Arroyo Grande	-3.76	-0.29	1.57	-2.30	-0.24
Santa Cruz-Watsonville	0.35	0.03	-0.16	0.20	0.02
Santa Maria-Santa Barbara	2.25	0.16	-0.85	1.41	0.13
Santa Rosa	-3.70	-0.31	1.66	-2.16	-0.25
Stockton-Lodi	0.87	0.14	-0.74	0.15	0.11
Vallejo-Fairfield	-0.40	-0.06	0.32	-0.09	-0.05
Visalia-Porterville	-0.19	-0.06	0.33	0.12	-0.05
Yuba City	0.78	0.22	-1.19	-0.38	0.18