Land Use Regulation, Regulatory Spillover and Housing Prices*

Desen Lin† Susan Wachter‡

This version: September 11, 2020

Abstract

We estimate the effect of city land use regulation on housing prices in the presence of regulatory spillover. The total effect of regulation is decomposed into a direct effect in which regulation lowers housing productivity and an indirect effect in which household location choice mitigates the price effects of regulatory restrictions. Using housing sales data from California, we structurally estimate a closed-form housing price equation based on a housing model with spatial arbitrage. We find that the total price effect of a one standard deviation increase in city restrictiveness is 9.3% on average, ranging from 4.1% to 14.4% across cities. The spillover effect is economically significant, with the size of the indirect effect equal to 21% of the direct effect for an average city, ranging from 0 to 47%. We point to the importance of identifying direct and indirect effects by controlling for regulation in surrounding locations. For jurisdictions with the power to impose regulation on a larger number of locations, regulation has a stronger price impact due to limits on regulatory spillover.

Keywords: housing prices, land use regulation, spillover effect, California

JEL: R10, R13, R31, R52, R58

* Acknowledgement: We thank David Albouy, Brent Ambrose (discussant), Daniel Broxterman (discussant), Zhenhua Chen (discussant), Salim Furth, Stuart Gabriel, Joe Gyourko, Christian Hilber, Stephen Malpezzi, David E Rappoport, Chris Severen, Tsur Somerville, Weihua Zhao (discussant) and conference participants at 47th AREUEA National Meeting, ESCP-TAU-UCLA Conference on Low Income Housing Supply and Housing Affordability, 2019 OSU PhD Conference on Real Estate and Housing, 14th Urban Economics Association Meeting, 2020 ASSA-AEA Meeting 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.

† Email: desenlin@sas.upenn.edu. PCPSE 535, 133 South 36th Street, Department of Economics, University of Pennsylvania, Philadelphia, PA 19104.

‡ Email: wachter@wharton.upenn.edu. 422 Vance Hall, 3733 Spruce Street, Department of Real Estate, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104.

1. Introduction

Land use regulation may raise housing prices through local restrictions on housing supply, as shown by a large empirical literature (Quigley and Rosenthal, 2005; Gyourko and Molloy, 2015). While land use restriction occurs within localities, housing demand goes beyond local boundaries through location choice, with the result that adjacent restrictiveness impacts local housing prices. However, few empirical studies (*e.g.* Pollakowski and Wachter, 1990) examine the joint impact of in-zone and adjacent restrictiveness on price outcomes. One challenge is the need to quantify the magnitude of regulatory spillover, which requires using spatially disaggregated data to capture the characteristics of nearby localities. Another challenge is the need to deal with the endogeneity of land use regulation, as the restrictiveness of home and nearby localities, together with other local characteristics, including income, are simultaneously determined.

In this paper, we estimate the effect of regulation on housing prices in the presence of regulatory spillover, based on a housing model with spatial arbitrage. Land use regulation imposes a direct effect on local housing prices by raising the cost of housing production, or, equivalently, lowering the productivity of housing supply (Albouy and Ehrlich, 2018).[1] In addition, it has an indirect effect that influences the relative attractiveness of substitutable localities and the location choice of households. For each locality, we build an imperfectly substitutable community whose characteristics reflect those of the nearby localities weighted by distance. We use this to estimate the surrounding impact of regulation on home housing prices. Most empirical studies of regulation do not consider the indirect effect of spillover but regress housing price against a single regulation index, excluding the surrounding regulatory regimes (Quigley and Rosenthal, 2005; Ihlanfeldt, 2007; Quigley et al., 2008; Jackson, 2018).[2] We introduce a spillover parameter to the structural model to empirically test the spillover effect and we reject the null hypothesis of no spillover. The size of the indirect effect as a share of the direct effect is 21% on average and exhibits large dispersion across space, with the size of the indirect effect ranging from 0 to 47% of the direct effect.

Studies that ignore regulatory surroundings or put no spatial weight on neighbors produce estimates that mix direct and indirect effects. Separating these effects is crucial in estimating the effect of regulation on the price of housing, especially for the case in which regulation covers multiple localities. In such a scenario, the restrictiveness of home and nearby localities is correlated to a large extent, which limits regulatory spillover and increases the regulatory price effect. Tighter regulation across nearby municipalities will have larger effects on housing prices than if only one municipality has tighter regulation, even for that single locality. While this spatial relationship has been hypothesized, empirical tests are sparse. We estimate a housing price

[1] Housing prices could increase in regulation due to an amenity effect (*e.g.* lower density) which we do not identify separately. As the components of the restrictiveness measure in the paper are mostly restrictions on housing production on the supply side, we believe the direct effect on amenities is limited. Moreover, Albouy and Ehrlich (2018) find the metro-level evidence that the supply-side effect of regulation dominates the demand-side effect.

[2] For example, Jackson (2018) tests for the regulatory impact in California with a pooling sample of cities and finds 5% higher housing prices attributable to one standard deviation increase in regulation. An earlier paper by Quigley et al (2008) for the San Francisco Bay Area finds a regulatory effect on housing prices in the range of 1-2% based on OLS and 4-5% when regulation is instrumented by political preference. For a comprehensive summary of the literature, see Table 1.

equation which explores the home and nearby variation of regulation, income, and land cost, with the coefficients disciplined by the equilibrium conditions of a housing model.

To estimate the housing price equation, we use housing sales for the period 2012-2017 from Zillow's ZTRAX database and a measure of city land use restrictiveness built from the California land use survey by Glickfeld and Levine (1992). We focus on California where the degree of restrictiveness varies greatly at the municipality level (Fischel, 1995). Using a tract housing price index built from sales data, we produce tract estimates of the regulatory effect that incorporate the inter-city heterogeneity of substitutable locations. To measure land use restrictiveness, we use a standardized city index that counts the affirmative answers of ten yes-no questions in Glickfeld and Levine (1992) on whether a residential regulatory measure was adopted. A city refers to an incorporated city (also known as municipality) in our context.

The estimation of the housing price equation additionally employs the home and nearby variation of the tract per capita income and the city land cost. We estimate parameters using generalized method of moments (GMM) and endogenize indices of regulation, income, and land cost, instrumented by political preference, local demographics, and natural amenities.

We find that the city housing price impact of one standard deviation increase in city restrictiveness (1.4 more regulatory measures) is 9.3% on average, ranging from 4.1% to 14.4% across cities. Our result shows a positive correlation between the city population and the total effect of regulation on housing prices. We find evidence that land use regulation shows a weaker indirect effect (cities are spatially less substitutable) in the largest cities than the average city in California.

The paper is organized as follows. Section 2 reviews the literature. Section 3 sets up the housing model with regulation and spatial arbitrage. Section 4 describes the data. Section 5 reports the estimation results. We conclude in Section 6.

2. Literature Review

There are a number of theory papers that show how household mobility mitigates the price effects of regulatory housing supply restrictions. Examples include Helsley and Strange (1995) and Van Nieuwerburgh and Weill (2010). Helsley and Strange (1995) introduces a strategic growth control model with multiple cities. The paper highlights the migration channel that links nearby regulation to home housing prices. Van Nieuwerburgh and Weill assume perfect mobility and simulate a Rosen-Roback model showing that there are very limited house price effects of regulation due to location substitutability.[3]

Quigley and Rosenthal (2005) survey the empirical literature on regulation and housing prices through 2005, and we summarize more recent work on this topic in Table 1. While recent empirical studies focus on

[3] Brueckner (1995) presents another strategic growth control model to show the spatial correlation of regulation. Under Leontief preferences, regulation in the home city will become stricter if it becomes stricter in a nearby city, with the prediction empirically tested in Brueckner (1998). Some other examples related to housing, migration, and regulation include Engle et al. (1992), Van Nieuwerburgh and Weill (2010), Hilber and Rober-Nicoud (2013), Hsieh and Moretti (2019), Annenberg and Kung (2020).

different housing markets and differ in the measurement of land use restrictiveness, there is a remarkable convergence on the estimated effects of regulation on the level of housing prices.[4] Across these studies, the impact of one standard deviation increase in regulation is associated with a more or less 5% increase in housing prices. Most studies do not treat regulation as endogenous, except for Ihlanfeldt (2007), Quigley et al. (2008), Zabel and Dalton (2011), and Albouy and Ehrlich (2018). Most studies explore the cross-sectional price variation (except for Zabel and Dalton, 2011, who use a panel data set to identify the price impact of minimum lot size by the difference-in-difference approach). Most studies use a reduced-form approach (except for Albouy and Ehrlich (2018) who adopt a structural approach in the estimation of housing cost and use the metro variation in the prices of land and structural input). In this paper, we structurally estimate a cross-sectional tract housing price equation and endogenize regulation.

One thread of the empirical literature on zoning and restrictions shows that not all price effects that occur through zoning are local, but tests for inter-jurisdictional regulatory price effects are empirically sparse. Pollakowski and Wachter (1990) estimate the housing price effects of home regulation and the relative restrictiveness of nearby jurisdictions in Montgomery County in Maryland to show the presence of the spillover effect across localities. The in-zone and adjacent restrictiveness produce stronger effects collectively than individually on local housing prices.[5]

While not directly testing the regulation externality, the early literature on monopoly zoning (Hamilton, 1978; Fischel, 1980; Rose, 1989; Bates, 1993; Thorson, 1996) empirically considers the impact of the surrounding areas and finds that urban areas with more monopoly zoning power have higher housing prices.[6] More recently, Ihlanfeldt (2007) examines housing prices in Florida using a reduced-form approach and lets the regulation index interact with the number of cities in the county to control the degree of monopoly zoning power. As the interactive term is negative, the author finds evidence that the impact of restrictiveness is increasing in monopoly zoning power.

Glaeser and Ward (2009) who empirically examine the price impact of the minimum lot size in Greater Boston find that the regulatory effect on housing prices disappears when local characteristics are controlled. Glaeser and Ward link this result to the substitutability of towns: "In the case of prices, it is possible that land use restrictions have no impact on prices in one locale, relative to a close substitute town, but still have an impact on area level prices. The basic economics of price restrictions tells us that we should not expect the

---

[4] The housing markets studied in recent work cover the US (Huang and Tang, 2012; Albouy and Ehrlich, 2018), Boston (Glaeser and Ward, 2009; Zabel and Dalton, 2011), Florida (Ihlanfeldt, 2007), and California (Quigley and Raphael, 2005; Quigley et al., 2008; Kok et al, 2014; Jackson, 2018). To measure restrictiveness, these studies use either a standardized sum of regulatory measures (Quigley and Raphael, 2005; Ihlanfeldt, 2007; Kok et al, 2014; Jackson, 2018), the principal factor of sub-indices (Quigley et al, 2008; Huang and Tang, 2012; Albouy and Ehrlich, 2018), or raw restriction measures (Glaeser and Ward, 2009; Zabel and Dalton, 2011).

[5] Also see Cho and Wachter (1991).

[6] Monopoly zoning literature puts forward the hypothesis that local governments with more monopoly power pursue stricter zoning policies to maximize the property value of homeowners. As summarized by Quigley and Rosenthal (2005), the more fragmented the governance structure of an urban area, the less monopoly power any one town will have due to entry price competition from its neighbors. The literature uses the count of towns in a metro area or the suburban land concentration ratio to measure the degree of monopoly zoning power.

price of a good to rise relative to a perfect substitute if that goods supply is restricted." However, Zabel and Dalton (2011) use the same data as Glaeser and Ward use but directly control a measure of amenity difference and identify a significant and positive price effect of regulation.[7]

Several studies in the literature on regulation use sharp spatial discontinuities in regulatory regimes to identify the causal impact of regulation on prices within certain distances from the boundaries of regulated cities. Turner et al. (2014) decomposes the regulatory effect on land value near straight regulatory boundaries in the US, while Severen and Platinga (2018) decomposes the total effect of the California Coast Act on housing price and rent. Both studies separate the cost of regulation on a parcel (local effect) and the amenity value of regulation on one's neighbors (external effect).[8] Properties on boundaries enjoy the same amenities but differ in regulation, leading to an identification of the effect of restrictiveness on the price of housing through the discontinuity of regulatory regimes. These empirical studies test for localized effects of regulatory spillover, but they find that these local effects are strong.

To measure land use restrictiveness, the literature relies on land use surveys taken by local planners. Among the surveys on California land use (Glickfeld and Levine, 1992; Donovan and Neiman, 1992; Gyourko et al., 2008; Quigley et al., 2008; Jackson, 2018; Pendall et al., 2018; Mawhorter and Reid, 2018; Gyourko et al., 2019), Glickfeld and Levine's survey has the highest response rate, covering 410 out of 482 municipalities (incorporated cities) in California.[9] The extensive coverage reduces measurement errors, allowing us to build robust measures of nearby characteristics to examine the spillover effect.

3. Model

3.1 A Graphical Example

Figure 1 shows an example where a home housing market has an identical and substitutable neighboring market, both at the equilibrium (E_0). The home and neighboring markets are two areas (*e.g.* census tracts) either in the same city or in two different cities. The two markets are not perfectly substitutable, and demand curves are downward sloping. The purpose of the example is to show how the price effect of city regulation depends on the status of the neighboring market. We assume tighter regulation lowers housing productivity, which shifts the supply curve leftward. The total effect captures the cross-sectional housing price differential

[7] As the ability to raise prices depends on the absence of close substitutes and locations are imperfectly substitutable, the authors find a positive regulatory price effect. Using a structural break method to infer the change of minimum lot size and examining a panel of housing prices, Zabel and Dalton controls for location fixed effects in the panel data to deal with the endogeneity of regulation.

[8] The total effect of regulation in Turner et al. (2014) is decomposed into an own-lot effect, which reflects the cost of regulation to the owner of a parcel; an external effect, which reflects the value of regulatory constraints on one's neighbors; a supply effect, which reflects the effect of regulated scarcity of developable land. Severen and Platinga (2018) classify the own-lot effect in Turner et al. (2014) as a local effect that combines a true own-lot effect and a neighbor effect.

[9] In the paper, cities refer to incorporated cities or municipalities. The survey by Glickfeld and Levine (1992) covers 410 out of 482 incorporated cities, compared to 147 cities in Donovan and Neiman (1992), 185 cities in Gyourko et al. (2008), 86 cities in Quigley et al. (2008), 366 cities in Jackson (2018), 172 cities in Pendall et al. (2018), 252 cities in Mawhorter and Reid (2018), 172 cities in Gyourko et al. (2019). See Lewis and Marantz (2019) for further comparison of the surveys.

attributed to land use regulation and is decomposed into a direct effect and an indirect effect. The x- and y-axis stands for the log quantity and the log price respectively.

Panel (a) shows the first case where home and neighboring markets are in the same city, so regulation of the two markets is perfectly correlated. To examine the price differential, we show a case in which regulation is tighter in both and shift the supply curve leftward in both. The direct effect captures the price differential ($E_0 \rightarrow E_1$) due to the supply channel. As the relative attractiveness is the same in the two localities, there is no room for spatial arbitrage and the indirect effect ($E_1 \rightarrow E_2$) is zero.

Panel (b) illustrates the second case where home and neighboring markets are in different cities, so regulation in the two markets is uncorrelated. We show again the case of a tighter home city regulation and shift the supply curve leftward in the home market only. Relative attractiveness to households favors the neighboring market, as a higher price of housing in the home market creates disutility. The indirect effect in the home market thus partially offsets the direct effect, with the housing demand in the home market spilling over to the neighboring market. The total effect of regulation is smaller in the second case than in the first case where regulation in the two markets is correlated.

The example points to the importance of regulatory surroundings and the correlated impact of regulation on housing prices. Next, we introduce the model and derive a housing price equation for estimation.

3.2 Household Choices

Households make consumption and location choices, consumption and location. In location j which represents a census tract, a household with the Cobb-Douglas preference $U(n_j, h_j, \tilde{u}_{ij}) = \left(\frac{h_j}{\alpha}\right)^{\alpha} \left(\frac{n_j}{1-\alpha}\right)^{1-\alpha} \tilde{u}_{ij}$ consumes the numeraire n_j and housing h_j, where $\alpha \in (0,1)$ is the housing share and $\tilde{u}_{ij}$ is a stochastic utility shifter that is independently and identically distributed across households. Denote the housing price to be p_{hj}. We assume the local income is equal to Z_j. The budget constraint is $p_{hj}h_j + n_j = Z_j$. Hence, the housing demand in location j is $h_j^D(p_{hj}) = \alpha Z_j / p_{hj}$, leading to the deterministic part of the utility to be $U_j = Z_j p_{hj}^{-\alpha}$.

Given the consumption choice, a household chooses where to locate. We assume heterogeneous location preference, with a household indexed by a home locality j. A household makes a binary location choice $j^*(j) = \arg max_{j' \in \{j,-j\}} \{U_{j'} \tilde{u}_{ij'}\}$ where j and $-j$ in the choice set denote staying in and outside location j respectively. The assumption allows us to derive an equilibrium price equation in a closed form and maintain the necessary element of spatial arbitrage. Empirically, we build a location substitute $-j$ by averaging the characteristics of nearby localities weighted by the distance, meaning that the outside option depends on j and there is moving friction. This deviates from the full mobility condition in the open city model in which the location choice is based on the absolute ranking of localities. Assume the household utility shifter $\tilde{u}_{ij}$ follows a Fréchet distribution (Redding and Rossi-Hansberg, 2017), with the shape parameter $\varphi \geq 0$ controlling the degree of location substitutability. The household propensity of choosing location j is

$$q_j(\vec{p}_{hj}) = \frac{Z_j^{\varphi} p_{hj}^{-\varphi\alpha}}{Z_j^{\varphi} p_{hj}^{-\varphi\alpha} + Z_{-j}^{\varphi} p_{h,-j}^{-\varphi\alpha}}, \text{ where } \vec{p}_{hj} \equiv \{p_{hj}, p_{h,-j}\} \tag{1}$$

In Section 5, a locality refers to a census tract and location substitutes are built from the nearby tracts in close distance. With the weight on nearby tracts decaying by the distance, a location substitute for most cases is in the same county or metro area as a home locality. The indirect effect driven by the spatial arbitrage in this paper is thus more local than the spatial effect in the studies that focus on the mobility arbitrage across metro areas.[10] Equation (1) assumes that the part of amenities not capitalized in the local income is constant in the home and nearby localities, so the terms about amenities in the numerator and the denominator drop out of equation (1).

3.3 Housing Production

Each location j has a single housing developer who builds housing H_j using land l_j and structure s_j as the inputs. The production technology is Cobb-Douglas and decreasing returns to scale. The cost function of the housing developer is thus $C(p_{lj}, p_s, H_j) = \left(\frac{p_{lj}}{\gamma}\right)^{\gamma} \left(\frac{p_{sj}}{1-\gamma}\right)^{1-\gamma} \left(\frac{H_j}{A_j}\right)^{1/v}$. p_{lj} and p_{sj} denote the unit cost of land and structure respectively; $\gamma \in (0,1)$ is the land share; $v \in (0,1)$ controls the production returns to scale; $A_j > 0$ is the housing productivity. The housing supply curve in location j is thus [11]

$$H_j^S(p_{hj}) = A_j^{\frac{1}{1-v}} \left(\frac{v p_{hj}}{c_j}\right)^{\frac{v}{1-v}}, \text{ where } c_j \equiv \left(\frac{p_{lj}}{\gamma}\right)^{\gamma} \left(\frac{p_{sj}}{1-\gamma}\right)^{1-\gamma} \tag{2}$$

We assume the structure is produced one-to-one from the numeraire, so the factor price is a constant, $p_{sj} = 1$. To model the regulatory impact, we assume the regulation intensity $R_j \in (0, \infty)$ affects housing supply by decreasing the housing productivity (Albouy and Ehrlich, 2018), $lnA_j = -\delta lnR_j$. While the sign of δ is expected to be positive to capture counter-productivity of regulation, a negative sign may indicate that regulation is productive by directly improving local amenities (less congestion, more open space, etc.).

3.4 Equilibrium Housing Price

In equilibrium, the housing markets clear, with the demand equal to the supply at home and at nearby localities, $N_j q_j(\vec{p}_{hj}) h_j^D(p_{hj}) = H_j^S(p_{hj}), \forall j$, where $N_j q_j(\vec{p}_{hj})$ is the number of householders who make the choice between $\{j, -j\}$ and prefer location j. As a location represents a census tract with a relatively

[10] For example, see Van Nieuwerburgh and Weill (2010) and Hsieh and Moretti (2019) for Rosen-Roback models in which households spatially arbitrage wage, amenities, and housing cost across metro areas. Hsieh and Moretti (2019) in addition consider the case of imperfect mobility across metro areas.

[11] The parameter v is related to the supply elasticity. While there is empirical support for heterogeneity across locations (Glaeser et al, 2006; Saiz, 2010; Hsieh and Moretti, 2019), the link is questioned by Larson et al. (2018) who numerically support the null relationship between regulation and supply elasticity. Broxterman and Liu (2019) show in a standard urban model that testing for the effects of regulation on housing supply should be in levels rather than elasticities.

stable population size, we normalize N_j to 1.[12] The aggregate housing demand on the left side of the market clearing condition is decreasing in price, because the individual demand is downward sloping and locations are imperfectly substitutable. The equilibrium condition shows that local housing markets are inter-related, with the price in location j depending on the prices in nearby markets. The effect of regulation will thus spill over to other localities through location choice. The market clearing condition implies that $lnp_{hj} = (1 - v)(-lnb_j + lnq_j)$ where $lnb_j \equiv \frac{1}{1-v} lnA_j - \frac{v}{1-v} ln\,(c_j/v) - lnZ_j - ln\alpha$. In appendix, we show lnq_j can be written as

$$\ln q_j = -\ln(1+\exp(-\Delta y)), \text{ where } \Delta y = y_j - y_{-j}$$

$$y_j = \varphi(1-\lambda)\ln Z_j + \lambda \ln b_j \text{ and } \lambda \equiv \frac{\alpha\varphi(1-v)}{\alpha\varphi(1-v)+1}$$

With the linear approximation of $lnq_j(\Delta y) = -\ln(1+\exp(-\Delta y)) \sim -\ln(2) + 0.5\Delta y$ in the neighborhood of $\Delta y = 0$, the housing price in location j can be written as follows.

$$\begin{aligned} \ln p_{hj} &= \delta \ln R_j + (1-v)\ln Z_j + \gamma v \ln p_{lj} \\ &\quad + \tfrac{1}{2}\lambda\left[\delta\Delta \ln R_{-j} - (\alpha^{-1} - 1 + v)\Delta \ln Z_{-j} + \gamma v \Delta \ln p_{l,-j}\right] + constant \end{aligned} \tag{3}$$

with $\Delta lnx_{-j} \equiv lnx_{-j} - lnx_j$ to be the log deviation from x_j.

The first line of equation (3) summarizes the direct effect; housing price in location j is increasing in regulation, income, and land cost at the home location. When $\lambda > 0$, the bracketed term in the second line captures the indirect effect due to spatial arbitrage. The housing price in location j is higher, if relative restrictiveness (ΔlnR_{-j}) is larger, relative income (ΔlnZ_{-j}) is lower and relative land cost ($\Delta lnp_{l,-j}$) is higher. The price equation nests the case of no spillover ($\lambda = 0$) as a special case. With the indirect effect linear in $\lambda \in [0,1]$, the nearby localities have no impact on the housing price of location j if $\lambda = 0$.

4. Data

4.1 Land Use Regulation Data

For land use regulation, we use the California land use survey conducted by Glickfeld and Levine (1992, 1999). We focus on municipalities (incorporated cities) in California because the degree of restrictiveness varies greatly at the city level (Fischel, 1995). Places that are not incorporated usually do not have jurisdictional power to set growth control policies. As addressed by Lewis and Marantz (2019), Glickfeld and Levine's survey has the best city coverage (410 out of 482 California cities), compared to 147 cities in Donovan and Neiman (1992), 185 cities in Gyourko et al. (2008), 86 cities in Quigley et al. (2008), 366 cities in Jackson (2018), 172 cities in Pendall et al. (2018), 252 cities in Mawhorter and Reid (2018), and 172 cities

[12] According to the Census (https://www2.census.gov/geo/pdfs/education/CensusTracts.pdf), a tract is an area with the population between 1,200 and 8,000. When the population goes beyond the upper or the lower threshold, a tract is either split or merged into neighboring tracts. For robustness, we check that the normalization does not affect our main results.

in Gyourko et al. (2019). Another reason for using the historical survey is to alleviate the concern of reverse causality that higher housing prices may induce stricter regulation.

To construct an aggregate regulatory measure, we count the affirmative answers of ten yes-no questions on whether a residential regulatory measure was adopted.[13] Ten regulatory measures include (1) the cap on residential permits, (2) population growth limitation, (3) adequate service requirements for residential development, (4) rezoning requirement of residential land to open space, (5) reduction of permitted residential density, (6) requirement of voter control over density increase, (7) requirement of super-majority vote to increase density, (8) phased development, (9) lot subdivision restriction, (10) floor area restriction. We normalize the count to zero mean and unit variance as the California Land Use Regulation Index ($CALURI$) whose model counterpart is lnR. The normalization makes our estimated regulatory effects comparable to those in the studies using standardized indices (see Table 1). In Figure 2, we map the spatial distribution of $CALURI$ across 410 cities in the survey, while in Figure 3, we compare the kernel density of $CALURI$ to the standard normal distribution. In Table 2, we summarize the share of cities for each regulatory instrument. On average, the number of growth controls adopted by a city is 1.72, with the range going from 0 to 6. Floor area restriction (43.9%) and adequate service requirements for residential development (39.5%) are the most adopted regulation in California.[14] We recognize that the regulation that we measure through this index is incomplete.

4.2 Housing Price, Land Cost and Income Data

Housing prices come from sales data in the Zillow Transaction and Assessment Dataset (ZTRAX). We include a set of housing characteristics to account for housing heterogeneity in the analysis: sale year, property use, number of bedrooms, property age, floor size, and mile distance from the city centroid.[15] We focus on housing sales for the period from 2012-2017, because the period is consistent with the time availability of land cost and the income data. In Table 2, we report the summary statistics for housing characteristics. The housing sales cover 5,141 tracts in 434 cities, out of 8,057 tracts in 482 cities in California. After matching to multiple data sources, the final data set used in the estimation has 4,720 tracts from 376 cities. The mean sale price in 2012 dollars is $505,337, with 80.4% of sales come from a single-family

[13] Among the ten yes-no questions in Glickfeld and Levine (1992), 7 questions come from Questions 5a-5g of the survey questionnaire, and the rest come from Questions 9-11. An aggregate index instead of binary sub-indices is used to provide a smooth measure of regulation intensity (Glaeser and Ward, 2009; Hilber and Robert-Nicoud, 2013).

[14] Our regulatory index measures the overall strictness of the regulatory environment, different from a particular physical or geographical constraint that directly affects factor inputs in the housing production. Our regulatory measure is positive correlated with other regulatory indices. Using the 163 cities that are present in the surveys conducted by Glickfeld and Levine (1992) and Gyourko et al. (2008), the correlation of CALURI and the Wharton Index (WRLURI) is 0.25 (unweighted) and 0.54 (weighted by the city population in 1990).

[15] We code housing sales into 5 categories of property use (single-family, townhouse/row house, cluster home, condominium/cooperative, planned unit development), 8 categories by the number of bedrooms, 9 categories by the property age inferred from the effective built year and the sale year, and the floor size in square foot.

residence. The mean floor size is 1,758 square feet, with 3.1 bedrooms on average. The mean property age is 34 years and the mean distance from the tract to the city centroid is 3.2 miles.

For land cost data, we use the land cost per quarter acre from the cross-section tract table of Davis et al. (2019) whose estimation is based on the cost-approach appraisals for 2012-2018. As a non-trivial fraction of the tract estimate of land cost is missing, we focus on the city estimate instead. Out of 482 California cities, the land cost of 458 cities are estimated using the tract average weighted by the tract population from the 2010 Decennial Census. We approximate the land cost for the rest of cities with the average land cost of the county where a city is located. For income data, we use per capita income by census tract from the 2016 ACS 5-year estimate for the period 2012-2016, which is consistent with the period used for housing sales.

4.3 Instrumental Variable Measures

If regulation, land cost, or income are endogenous in the housing price equation, their estimated effects on housing prices will be biased. In particular, we are concerned that the regulatory index may be a function of income and that a finding of significance will be due to this. Hence, we consider a set of instrumental variables that affect housing prices through regulation, land cost and income, but with no direct impact on the tract housing prices. Because the model with constrained coefficients is over-identified, we test the validity of the over-identifying restriction (Hansen, 1982) after estimation.

First, we follow Quigley et al. (2008) to include political preference as part of the instruments for restrictiveness. Using the Record of American Democracy Data from King et al. (2007), we aggregate the precinct-level voting counts by political parties to the city level and calculate the city voting share for the Republican candidate in the 1992 general election.[16] Second, we follow Ihlanfeldt (2007) to include demographic variables as the instruments in the housing price equation. We include the college graduation share and the mean householder age from the 2016 ACS 5-year tract estimate. These two instruments mainly capture the variation of the tract income. Third, we follow Albouy and Ehrlich (2018) and adopt a broader set of natural amenities to instrument the land cost. We use the winter temperature and the winter sun hours from the USDA natural amenity file (McGranahan, 1999). For the nearby values (lnx_{-j}) in equation (3), we separately construct their instruments and explain how we do this in the next section.[17]

[16] The idea of instrumenting regulation with political preference and local demographics is motivated by the Homevoter Hypothesis by Fischel (1995, 2001). The hypothesis says that local government officials pursue land use and taxation policies that increase the property value of the homeowners who have the greatest political influence by controlling the majority of votes. In general, homeowners tend to be older and better educated (Ruggles et al., 2018).

[17] With multiple endogenous variables, the instruments should provide exogenous variation jointly for each endogenous variable. The validity of the instruments requires that they are excludable. A potential concern is that the instruments (of regulation) are correlated with the unobserved price factors (Davidoff, 2016). The exclusion restriction is that the instruments are orthogonal to the housing prices conditioning on regulation, income, and land cost, which is weaker than the requirement that the instruments are orthogonal to unconditional housing prices (or housing prices conditioning on a subset of the variables).

4.4 Define Nearby Indices

By Census' definition, as the tract area is not necessarily nested in the city area, we use the following way to match a tract if possible, to a city using the spatial packages from *R* software. For a tract whose centroid lies within a city's boundary, we assign the tract with the land cost and regulation of the matched city. For a remaining tract whose polygon intersects with any city area, we match it to the nearest city with a common area. The rest of the tracts are rural and disjoint from any city area, which we exclude in our analysis. The tracts at the city center and on the periphery face the same home city regulation ($lnR_j = CALURI_j$), but they differ in the nearby regulation (lnR_{-j}) that depends on the spatial distribution of the adjacent tracts.

To create a housing price index, we adjust the sale prices for housing heterogeneity by regressing the log real price in sale i in tract j on a vector of housing characteristics X_{ijt}, $lnp_{hijt} = X_{ijt}\beta^h + \chi_{t,co} + \varepsilon^h_{ijt}$ where $\chi_{t,co}$ is the year-county fixed effect. We create tract housing price index $ln\,\tilde{p}_{hj}$ by aggregating the residuals by tract for housing sales in 2012-2017.[18] Similarly, we create the city land cost index $ln\,\tilde{p}_{lj}$ and the tract income index $ln\,\tilde{Z}_j$ by removing the county fixed effect, with the regressions weighted by the population from the 2010 Decennial Census.

To construct the nearby value x_{-j} of a variable x in tract j, we use the geometric mean of 30 nearest tracts using the inverse-distance weight $w_{jj'}$.[19]

$$\ln x_{-j} = \sum_{j' \in S(j)} w_{jj'} \left(\ln x_{j'}\right), \text{ where } w_{jj'} \equiv d_{jj'}^{-2} \left(\sum_{j' \in S(j)} d_{jj'}^{-2}\right)^{-1} \tag{4}$$

where $d_{jj'}$ is the distance between tract j and j'. $S(j)$ is the set of nearby tracts with positive weights. We use spatial weighting to construct the nearby indices of land cost, regulation, income, as well as the instruments we use to endogenize these variables. The nearby indices depend on where a home locality is and how far the nearby localities are. Hence, the location choice is local, not based on the absolute ranking of attractiveness.

Taking the nearby index of regulation lnR_{-j} as an example, we show on the heatmaps in Figure 4 how the nearby regulation varies across two major metro areas in California: San Francisco-Oakland-Berkeley MSA and Los Angeles-Long Beach-Anaheim MSA. While regulation is constant within a city, the nearby regulation index of tracts on the city periphery puts more weight on the restrictiveness of adjacent cities. We show the extensive coverage of housing sales at the tract level and land use regulation at the city level in our sample, which is necessary to build robust measures of nearby characteristics to study regulatory spillover.

[18] The set of hedonic prices β^h is estimated using all housing sales in California for the period of 2000-2017. Allowing β^h to vary over time won't substantially affect our estimation results.

[19] We identify the 30 nearest tracts using all 8,057 tracts in California in the estimation, not the subset of tracts that we have data on. Including a larger number of neighbors than 30 does not change the nearby indices much. The choice on the number of nearby tracts is large enough to cover all contiguous tracts of any given tract in California.

5. Estimation

5.1 Estimation Approach

In the empirical model, we define a location j to represent a census tract. Housing prices and per capita income vary at the tract level, while land cost and regulation vary at the city level. We assume $lnZ_j = ln\tilde{Z}_j + \eta ln\tilde{Z}_{-j}$ and $lnZ_{-j} = ln\tilde{Z}_{-j} + \eta ln\tilde{Z}_j$ to model the income impact of nearby tracts on the housing demand of a home tract, with $\eta \geq 0$ to be the weight on the nearby income index.[20] By replacing lnZ_j with $ln\tilde{Z}_j$, equation (3) is written as follows.

$$\begin{aligned} \ln \tilde{p}_{hj} &= \delta \ln R_j + (1-v)\ln \tilde{Z}_j + (1-v)\eta \ln \tilde{Z}_{-j} + \gamma v \ln \tilde{p}_{lj} \\ &+ \tfrac{1}{2}\lambda\left[\delta \Delta \ln R_{-j} - (\alpha^{-1} - 1 + v)(1-\eta)\Delta \ln \tilde{Z}_{-j} + \gamma v \Delta \ln \tilde{p}_{l,-j}\right] + constant \end{aligned} \tag{5}$$

with the direct and indirect effect on the price level highlighted on the first and the second line respectively. By rearranging the terms, we have a general form of the price equation with the constrained coefficients.

$$\begin{aligned} &\ln \tilde{p}_{hj} = \beta_0 + \beta_{R1}\ln R_j + \beta_{R2}\ln R_{-j} + \beta_{l1}\ln \tilde{p}_{lj} + \beta_{l2}\ln \tilde{p}_{l,-j} + \beta_{z1}\ln \tilde{Z}_j + \beta_{z2}\ln \tilde{Z}_{-j} + \varepsilon_{hjt} \\ &\text{where } \beta_{R1} = \delta(1-\tfrac{1}{2}\lambda),\ \beta_{R2} = \tfrac{1}{2}\delta\lambda,\ \beta_{l1} = \gamma v(1-\tfrac{1}{2}\lambda),\ \beta_{l2} = \tfrac{1}{2}\lambda\gamma v \\ &\beta_{z1} = (1-v) + \tfrac{1}{2}\lambda(\alpha^{-1} - 1 + v)(1-\eta),\ \beta_{z2} = (1-v)\eta - \tfrac{1}{2}\lambda(\alpha^{-1} - 1 + v)(1-\eta) \end{aligned} \tag{6}$$

As the housing share α is determined outside of the model, we assume $\alpha = 0.2$.[21] Besides the constant term β_0, there are 5 parameters to estimate $(\delta, \gamma, v, \eta, \lambda)$ with the constraint $\lambda \in [0,1]$, so we need at least 6 moment conditions to achieve identification. Regulation index (lnR_j), land cost index ($ln\,\tilde{p}_{lj}$) and income index ($ln\tilde{Z}_j$) identify δ, γ and v respectively. As shown by equation (5), the linear combination of home and nearby income indices pins down η. The linear combination of three relative indices jointly determines λ that summarizes the strength of location substitutability. We will report the specifications with unconstrained and constrained coefficients. For the constrained models, we estimate equation (6) using two-stage Generalized Method of Moments (GMM) and weigh each tract with the number of housing sales in the period of 2012-2017.

5.1.1 Direct and Indirect Effects of Regulation: A Decomposition

By equation (6), we can express the total effect of regulation on housing prices as follows.

$$\frac{\partial \ln \tilde{p}_{hj}}{\partial \ln R_j} = \delta\left(1 - \frac{1}{2}\lambda\right) + \frac{1}{2}\delta\lambda \cdot \frac{\partial \ln R_{-j}}{\partial \ln R_j} = \underbrace{\delta}_{\text{Direct Effect}} + \underbrace{\frac{1}{2}\delta\lambda \cdot \frac{\partial \Delta \ln R_{-j}}{\partial \ln R_j}}_{\text{Indirect Effect}} \tag{7}$$

The total effect is decomposed into two parts. The first term δ is the direct effect of tightening regulation in tract j. One standard deviation increase in regulation leads to a δ percent decrease in housing

[20] The assumption nests two extreme cases. If $\eta = 0$, households live and work in the same tract. If $\eta = 1$, income does not affect spatial arbitrage.

[21] The housing share α is based on the US real housing service and utility expenditure in the real personal consumption expenditures from BEA.

productivity, thus shifting the supply curve leftward and increasing the price of housing by δ percent. The second term is newly introduced in this paper and measures the indirect effect due to spatial arbitrage.[22]

As regulatory policies in California are determined by cities, regulation in the home and nearby census tracts are correlated, with the correlation depending on the spatial distribution of the nearby locations. If the nearby regulation lnR_{-j} puts more weight on the tracts in different cities nearby, the derivative $\partial lnR_{-j}/\partial lnR_j$ is closer to 0 and the response of relative restrictiveness $\partial \Delta lnR_{-j}/\partial lnR_j$ is closer to -1. The indirect effect is estimated to be large. The total effect of regulation is closer to the lower bound $\beta_{z1} = \delta(1 - \frac{1}{2}\lambda)$. If the nearby regulation puts more weight on the tracts in the same city where tract j is located, the response of relative restrictiveness is closer to 0. In the absence of spillover, the total effect of regulation is closer to the upper bound $\beta_{z1} + \beta_{z2} = \delta$. Hence, the indirect effect is non-positive. It is strictly negative, if (1) there is positive spillover $\lambda > 0$, and (2) relative restrictiveness ΔlnR_{-j} responds to the change in the home city restrictiveness. As the marginal effect of regulation lies in the interval determined by the coefficients of home and nearby regulation $[\beta_{z1}, \beta_{z1} + \beta_{z2}]$, equation (7) is written as

$$\frac{\partial \ln \tilde{p}_{hj}}{\partial \ln R_j} = \beta_{z1} + \rho_j \beta_{z2} \tag{8}$$

where $\rho_j \equiv \partial lnR_{-j}/lnR_j \in [0,1]$ captures the share of nearby tracts located in the same city as the home tract, which is exogenous. To calculate the city price impact of regulation, we aggregate the tract impact by taking a weighted average of equation (8).

5.2 Estimation Results

In Table 3, we estimate the housing price equation (6) and report the marginal impact of home and nearby indices on regulation, income, and land cost on the home housing price index. Columns 1 is the unconstrained coefficient models estimated by OLS, while Columns 2-3 are the constrained coefficient models estimated by GMM. Columns 1-2 assume exogenous regulation, income, and land cost, while Column 3 endogenizes home and nearby indices with instrumental variables.

5.2.1 Impact of Nearby Localities: OLS and GMM Estimates

Column 1 estimates an OLS model with unconstrained coefficients and controls the nearby tract indices in addition to the home tract indices. The specification includes the home and nearby tract indices of regulation, income, and land cost. The home and nearby indices of regulation, income, and land cost show positive price effects. For the impact of home indices, the marginal effect of a one standard deviation increase in home restrictiveness on housing prices is 1.6%; the marginal effect of 1% increase in home tract land cost is 0.20%; the marginal effect of 1% increase in home tract income is 0.28%. For the impact of nearby indices, the

[22] Note that the regulatory effect could underestimate the actual effect and should be interpreted as a lower bound, because regulation can directly affect the land cost in a complex way in the price equation. Our model can pick up the regulatory effect through the land cost if regulation affects the factor prices (*e.g.* land cost) in a multiplicative way.

marginal effect of a one standard deviation increase in nearby restrictiveness is 2.7%; the marginal effect of 1% increase in home tract land cost is 0.14%; the marginal effect of 1% increase in home tract income is 0.11%. The nearby indices are jointly and individually significant in Column 1, indicating that nearby localities empirically affect home housing prices.

Column 2 builds on Column 1 by estimating a GMM model and imposes constraints on the coefficients as indicated in equation (6). We first estimate 5 parameters $(\delta, \gamma, v, \eta, \lambda)$ through two-stage GMM and then recover the marginal effect of the home and nearby indices by the Delta method. Note λ, the spillover estimate, captures the indirect effects resulting from spatial arbitrage.

The constraints on the coefficients lead to a decrease in the adjusted R-squared from 74% to 69%. The main distinction between Columns 1 and 2 in terms of the coefficients comes from the regulation indices of home and nearby tracts. Column 2 shows that one standard deviation increase in home and nearby regulation increases housing prices by 2.4% and 1.8% respectively, compared to 1.6% and 2.7% respectively in Column 2. With 5 parameters to estimate besides the constant term in equation (6) and 6 moment conditions built from the home and nearby indices of regulation, income, and land cost, we check the validity of the over-identifying restriction with the J test (Hansen, 1982) whose null hypothesis is instrument consistency. Column 3 fails to reject the null hypothesis with a p-value of 0.22. We find an estimate of the spillover measure of $\lambda = 0.87$, rejecting the hypothesis of no spatial arbitrage ($\lambda = 0$) with 95% confidence. Since in theory λ can vary from 0 to 1, this is consistent with a large spillover effect.

5.2.2 Instrumental Variable Estimates

In Column 2 of Table 3, the indices of regulation, income, and land cost are treated as exogenous. If any of the indices are correlated with an unobserved price factor, the estimated marginal effects of all indices are biased. To deal with this concern, we endogenize the home and nearby indices of regulation, land cost and income in Column 3. Following the literature on endogenous regulation and housing prices (Quigley et al., 2008; Ihlanfeldt, 2007; Albouy and Ehrlich, 2018), we adopt political preference, local demographics, and natural amenities as the instruments. While instrument excludability is hard to test, we believe the instruments adopted in the literature are plausible.

The identification assumption is that the proposed instruments are orthogonal to the housing prices conditioning on the home and nearby indices of regulation, income, and land cost. We use the z-scores of the home and nearby values of the voting share for Republican candidates in the 1992 general election (King et al., 2007), the college graduation share and the mean householder age from 2016 ACS 5-year estimates, the mean temperature and the mean sun hours in January from USDA (McGranahan, 1999). These 10 instruments (5 sets of home and nearby values) are used to estimate 5 parameters besides the constant term.

We find a larger spillover estimate of $\lambda = 0.94$ in Column 3 than in Column 2. The marginal effect of home (nearby) regulation on tract housing prices increases from 2.4% (1.8%) in Column 2 to 8.1% (7.2%) in Column 3. The upward adjustment of the home regulatory impact when regulation is endogenized echoes

the findings by Ihlanfeldt (2007) and Quigley et al. (2008). Column 3 also shows larger marginal effects of home and nearby indices of land cost and smaller marginal effects of home and nearby indices of income.

To confirm the validity of the endogenous specification, we test the instrument relevance and the over-identifying restriction. In Table 4, we report the first-stage regressions of the home and nearby indices of regulation, income, and land cost on these 10 instruments. In the regressions of home and nearby regulation, the voting share for the Republican candidate has a negative sign while the college graduation share has a positive sign, consistent with the results in Brueckner (1998) who tests a strategic growth control model (Helsley and Strange, 1995; Brueckner, 1995) using a California data set. The F statistics of the instrument exclusion for all regressions stay above the rule-of-thumb threshold of 10, suggesting there is a limited issue of weak instruments. Besides, we test the weakness of identification and conduct the rank test of Kleibergen and Paap (2006), with the null hypothesis of under-identification using the proposed instruments. Column 3 rejects the null hypothesis with a p-value of 0.001. To test the over-identifying restriction, the J test of Column 3 shows instrument consistency and fails to reject the null hypothesis with a p-value of 0.31.

5.2.3 Spatial Heterogeneity

We relax the assumption that the parameters estimated in Table 3 are constant across space. We divide 4,720 California tracts we use in the estimation into three Combined Statistical Areas (CSAs): Los Angeles-Long Beach CSA (2,231 tracts in 155 cities), San Jose-San Francisco-Oakland CSA (1,265 tracts in 103 cities), and the rest of California (1,224 tracts in 118 cities).[23]

In Table 5, we re-estimate the constrained coefficient models of Columns 2-3 in Table 3 for the three CSA groups and report the estimates for the exogenous and endogenous specifications. We allow the parameters $(\delta_g, \gamma_g, v_g, \eta_g)$ to vary by group g.[24] When we compare exogenous and endogenous specifications within a group, Table 5 shows that the estimates of the marginal effects of home and nearby indices are qualitatively consistent with the findings in Table 3. Across the CSA groups, we find that San Jose-San Francisco-Oakland CSA (SF CSA) has a stronger price effect of home and nearby regulation than Los Angeles-Long Beach CSA (LA CSA). Both CSAs show stronger regulatory effects than the rest of California. With the endogenous specifications, one standard deviation increase in the home (nearby) restrictiveness is associated with a 6.9% (6.1%) tract housing price increase in the Los Angeles Area and a 7.7% (6.8%) housing price increase in the San Francisco Area, compared to a 4.1% (3.7%) price increase in the rest of California.

[23] As we rely on the cross-city variation of regulation to identify δ, the sample division allows us to have sufficient statistical power to identify δ within a group and to balance the number of cities across the groups.

[24] We keep the spillover estimate λ from Table 3 constant. We consider the spillover measure λ as a global parameter, meaning that location substitutability does not result in the difference in the estimated marginal effects of home or nearby indices across groups.

5.3 City Impact of Regulation with Spatial Arbitrage

We estimate the city-level price impact of regulation and show direct and indirect effects, accounting for spatial heterogeneity across CSAs based on the estimates in Section 5.2.3. We define the city estimate of ρ to be the mean of the share ρ_j across tract j, weighted by the number of housing sales for 2012-2017. The city means of ρ are 0.54 for the Los Angeles Area, 0.55 for the San Francisco Area, and 0.67 for the rest of California. While the means of the CSA groups seem close to the California mean of 0.58, there is a large dispersion of ρ within each group and the distributional pattern differs substantially across groups. The standard deviations of ρ are 0.23, 0.30, and 0.29 for the Los Angeles Area, the San Francisco Area, and the rest of California respectively.

We combine the city estimate of ρ and the estimated coefficients of home and nearby indices of regulation (Table 5) in equation (8) to derive the average marginal effect of regulation on city housing prices. In Table 6, we report the total effects as well as the decomposed effects (direct and indirect) for the average city as well as for the 20 largest cities that cover 38.2% of the overall city population in California.

When the home and nearby indices are endogenized, the total effect of regulation on the housing price of an average city in California is 9.3%, which is decomposed into an 11.8% direct effect and a -2.5% indirect effect. The estimated city average indirect effect in absolute value (2.5%) is far larger than that for the 20 largest cities, as expected. The total effect in the endogenous specification is three times stronger than that in the exogenous specification, suggesting that most studies surveyed in Table 1 are likely to underestimate the regulatory price effect. The ratio of the indirect to the direct effect in absolute value, $\frac{1}{2}\lambda(1-\rho)$, is governed by the marginal effect of home regulation on nearby regulation ρ. The ratio has a mean value of 0.21. As ρ can take values from 0 to 1, the size of the indirect effect can vary from 0 to 47% of the direct effect.

San Francisco (14.4%) San Jose (13.9%), Oakland (13.8%), Los Angeles (12.3%), and Long Beach (12.2%) show the strongest total effect of regulation among the top 10 largest cities in California. By decomposing the total effect, we show the regulatory effect in the largest cities is attributed to both a strong direct effect and the limited indirect effect. The marginal effects of home regulation on nearby regulation ρ for the five cities are greater than 0.87, compared to the California mean of 0.58. The evidence suggests that land use regulation shows a stronger direct effect (regulation is locally more impactful) and a weaker indirect effect (cities are spatially less substitutable) in the largest cities than the average city in California.[25]

In Table 7, we show how the total effect and the decomposed effects of regulation are correlated with various factors. We examine cities by CSA, showing that the total effects in the San Francisco Area (11.4%) and in the Los Angeles Area (10.2%) are much stronger than for the rest of California (6.3%), which is attributed to stronger direct effects in the two areas. The total effect of regulation is increasing in per capita income, restrictiveness, and population. While the correlation with the total effect is mostly driven by the correlation with the direct effect, the indirect effect plays different roles. The indirect effect in absolute value

[25] That the same level of regulatory restrictiveness has a greater impact on larger cities could be attributed to the unobserved regulatory measures in larger cities that adopt a more complex set of regulation than in smaller cities, or reflect the impact of higher density or a sorting effect in larger cities.

is decreasing in regulation and population, thus strengthening the direct effect of regulation. The indirect effect in absolute value is increasing in per capita income, thus mitigating the direct effect of regulation.

6. Discussion and Conclusion

We estimate and decompose the total effect of land use regulation on housing prices in the presence of regulatory spillover. In this paper, regulation has a direct positive effect on housing prices by lowering housing productivity through the supply channel. In addition, it has an indirect effect that influences the relative attractiveness of localities and the location choice of households through the demand channel. Based on a housing model with spatial arbitrage and housing sales data from California, we estimate a structural housing price equation with constrained coefficients on the characteristics of the home and nearby localities.

The city housing price impact of one standard deviation increase in city land use restrictiveness (1.4 more regulatory measures) is 9.3% on average and ranges from 4.1% to 14.4% across California cities. We produce a city estimate of the regulatory effect that considers the heterogeneity of substitutable localities. We find an economically and statistically significant spillover effect of regulation across localities, with the size of the indirect effect equal to 21% of the direct effect for an average city. The evidence suggests that larger cities (by population) have a stronger total effect of regulation on housing prices, which we attribute to a stronger direct effect (regulation is locally more impactful) and a weaker indirect effect (cities are spatially less substitutable).

We point out the importance of decomposing the total effect of regulation into direct and indirect effects. When a jurisdiction, whether it is a city, a county or a state, has greater power to impose regulation on a larger number of localities, regulation shows a stronger impact on housing prices by limiting regulatory spillover across these localities. Without controlling for the regulation of surrounding localities, the estimated effect of regulation understates (overstates) the total impact of regulation in the jurisdictions with greater (less) zoning power.

While our estimation of the regulatory price effect adds new insight to the literature, we would like to point out the limitations of our estimates. First, to identify the indirect effect more precisely, we circumscribe the effect and only look at nearby locations. We are likely to underestimate the indirect effect by leaving out the migration to farther out locations and locations outside of California. Second, we are underestimating the total effect of regulations through land cost. Importantly, we do not capture the impact of regulation on sorting effects through the interaction of labor, housing and land markets. Future work will focus on modeling zoning and restrictiveness in a unified framework to identify the regulatory effect through multiple channels.

We use a survey-based dataset to measure regulatory restrictiveness that provides survey responses for many jurisdictions in California. However, the extent to which these data describe all relevant regulation is limited. There is a great need to expand information on local land use regulations. While we recognize that regulations are determined endogenously along with local income and have attempted to take account of this.

Much additional work needs to be done to understand the indirect and direct effects of regulation on local and regional house prices.

References

Albouy, D., & Ehrlich, G. (2015). 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.

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

Anenberg, E., & Kung, E. (2020). Can more housing supply solve the affordability crisis? Evidence from a neighborhood choice model. *Regional Science and Urban Economics*.

Bates, L. J. (1993). Municipal monopoly power and the supply of residential development rights. *Eastern Economic Journal*, 19(2), 173-184.

Broxterman, D. and Liu, Y., What is the Relation between Land Use Controls and Housing Prices? (September 21, 2019). Working paper.

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

Brueckner, J. K. (1998). Testing for strategic interaction among local governments: The case of growth controls. *Journal of urban economics*, 44(3), 438-467.

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

Davidoff, T. (2016). Supply Constraints Are Not Valid Instrumental Variables for Home Prices Because They Are Correlated With Many Demand Factors. *Critical Finance Review*, 5(2), 177-206.

Davis, M. A., Larson, W. D., Oliner, S. D., & Shui, J. (2019). The Price of Residential Land for Counties, ZIP Codes, and Census Tracts in the United States. *AEI Paper & Studies*.

Donovan, T., & Neiman, M. (1992). Citizen mobilization and the adoption of local growth control. *Western Political Quarterly*, 45(3), 651-675.

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

Federal Housing Finance Agency (2012). *U.S. House Price Index Report - 1Q 2012 / March 2012*

Fischel, W. A. (1980). Zoning and the exercise of monopoly power: A reevaluation. *Journal of Urban Economics*, 8(3), 283-293.

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. (2006). 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*.

Gyourko, J., Hartley, J., & Krimmel, J. (2019). The Local Residential Land Use Regulatory Environment Across US Housing Markets: Evidence from a New Wharton Index (No. w26573). National Bureau of Economic Research.

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. (1978). Zoning and the exercise of monopoly power. *Journal of Urban Economics*, 5(1), 116-130.

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

Helsley, R. W., & Strange, W. C. (1995). Strategic growth controls. *Regional Science and Urban Economics*, 25(4), 435-460.

Hilber, C. A., & Robert-Nicoud, F. (2013). On the origins of land use regulations: Theory and evidence from US metro areas. *Journal of Urban Economics*, 75, 29-43.

Hsieh, C. T., & Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2), 1-39.

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.

King, G., Palmquist, B., Adams, G., Altman, M., Benoit, K., Gay, C.; Lewis, J. B.; Mayer, R., &Reinhardt, Eric, (2007), *Record of American Democracy, All Key Data Files*

Kneebone, E., & Holmes, N. (2015). The growing distance between people and jobs in metropolitan America. Washington, DC: *Brookings Institution: Metropolitan Policy Program*

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.

Larson, W., Yezer, A., & Zhao, W. (2018). An Examination of the Link between Urban Planning Policies and the High Cost of Housing and Labor. Working paper.

Levine, N. (1999). The effects of local growth controls on regional housing production and population redistribution in California. Urban Studies, 36(12), 2047-2068.

Lewis, P. G., & Marantz, N. J. (2019). What planners know: Using surveys about local land use regulation to understand housing development. *Journal of the American Planning Association*, 85(4), 445-462.

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.

Mayer, C. J., & Somerville, C. T. (2000). Residential construction: Using the urban growth model to estimate housing supply. *Journal of urban economics*, 48(1), 85-109.

McGranahan, D. A. (1999). Natural amenities drive rural population change (No. 1473-2016-120765).

Van Nieuwerburgh, S., & Weill, P. O. (2010). Why has house price dispersion gone up?. *Review of Economic Studies*, 77(4), 1567-1606.

Pendall, R., Wegmann, J., Martin, J., & Wei, D. (2018). The growth of control? Changes in local land-use regulation in major US Metropolitan Areas from 1994 to 2003. *Housing Policy Debate*, 28(6), 901-919.

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. Working Papers. Berkeley Program on Housing and Urban Policy, Institute of Business and Economic Research. UC Berkeley.

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.

Redding, S. J., & Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9, 21-58.

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

Rose, L. A. (1989). Urban land supply: Natural and contrived restrictions. *Journal of Urban Economics*, 25(3), 325-345.

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

Rosenthal, S. S. (2014). Are private markets and filtering a viable source of low-income housing? Estimates from a" repeat income" model. *American Economic Review*, 104(2), 687-706.

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.

Thorson, J. A. (1996). An examination of the monopoly zoning hypothesis. *Land Economics*, 43-55.

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

Zabel, J., & Dalton, M. (2011). The impact of minimum lot size regulations on house prices in Eastern Massachusetts. *Regional Science and Urban Economics*, 41(6), 571-583.

Tables

Table 1. Summary of Recent Studies on Land Use Regulation and Housing Prices

Author (Year)	Market	Method	Regulatory Measure	Housing Price Measure	Data Period	Regulatory Effect on Prices
Quigley and Raphael (2005)	407 cities in California	OLS	Number of growth control measures (out of 15) adopted by each city	Housing price index constructed from Census Public Use Micro Sample (constant-quality index)	1990, 2000	1 more measure leads to 3.1% increase in prices in 1990 and 4.5% in 2000
Ihlanfeldt (2007)	105 cities in Florida	OLS, 2SLS	Number of restrictive land use management techniques (out of 13)	Sales data from property tax rolls with housing characteristics	2000-2002	1 more regulation leads to 7.7% price increase with endogenous regulation
Quigley, Raphael and Rosenthal (2008)	86 cities in the San Francisco Bay Area	OLS, IV	Sum/principal factor of 10 sub-indices from Berkeley Land Use Survey	Housing prices with housing characteristics from 2000 Census	2000	1 SD increase leads to 1.1%-2.2% with OLS, or 3.8%-5.3% with IV
Glaeser and Ward (2009)	187 cities and towns in Great Boston Area	OLS	Min lot size, sum of three dummies of non-lot restriction (septic systems, wetlands, sub-divisions)	Banker and Tradesman data on housing prices with housing characteristics	2000-2005	Effect of regulation is insignificant when town characteristics are included
Zabel and Dalton (2011)	187 cities and towns in Great Boston Area	DID with structural break	Min lot size	Single-family housing sales with housing characteristics from Warren Group	1987-2006	1 acre (1.5 SD) increase in min lot size increases prices by 9%.
Huang and Tang (2012)	327 cities in US	OLS	WRLURI	Zillow hedonic price index	2000-2009	1 SD increase leads to 5.6% price increase for 2000-2006; 4.5% price decrease for 2006-2009
Kok, Monkkonen, Quigley (2014)	110 cities in San Francisco Bay Area	OLS	Number of independent reviews/approvals required by a locality before of permit issuance; number of separate reviews by local authorities required for zoning change approval	Mean sales price from DataQuick	1990-2000	1 SD decrease in the number of reviews for permit approval (zone change) related to price decrease of 4–8% (1-2%)
Jackson (2018)	366 cities in California	OLS	Standardized sum of the 9 sub-indices from California Land Use Survey in 2018	Zillow hedonic price index	Jan 2000, April 2006, Jan 2012	1 SD increase leads to 5% increase in price
Albouy and Ehrlich (2018)	230 metros in US	Structural, OLS, IV	WRLURI	Housing price from 1% ACS sample with housing characteristics	2005-2010	1 SD increase leads to 6.5%-8.8% increase in prices.

Note: for studies on land use regulation and housing prices before 2005, see the summary by Quigley and Rosenthal (2005).

Table 2. Summary Statistics of Housing Sales and Land Use Regulation in California

Housing Characteristics	Mean	SD	Pct.25	Pct.50	Pct.75
Real sale price (2012M1 $)	505,337	765,694	218,708	364,108	593,496
Floor size (sq.ft.)	1,758	12,451	1,198	1,547	2,064
Number of bedrooms	3.087	0.98	3	3	4
Property age	40.126	24.822	21	38	58
Is single-family (binary)	0.804	0.397	1	1	1
Is townhouse/row house (binary)	0.006	0.078	0	0	0
Is cluster home (binary)	0.003	0.059	0	0	0
Is condo/coop (binary)	0.168	0.374	0	0	0
Is planned unit development (binary)	0.018	0.132	0	0	0
Miles from tract to city centroid	3.19	4.24	1.07	1.89	3.49
Spatial Coverage (#/Total)	CBSA	Counties	Cities	Tracts	Sales
	33/34	52/58	434/482	5,141/8,057	2,073,813
Growth Control Measures (Yes = 1, No = 0)	Mean	SD	Min	Max	
Cap on residential permits	0.124	0.330	0	1	
Population growth limitation	0.093	0.290	0	1	
Adequate service of residential development	0.395	0.489	0	1	
Rezoning of residential land to open space	0.088	0.283	0	1	
Reduction of permitted residential density	0.351	0.478	0	1	
Voter control over density increases	0.051	0.221	0	1	
Super-majority vote to increase densities	0.024	0.154	0	1	
Phased development	0.122	0.328	0	1	
Lot subdivision restriction	0.041	0.200	0	1	
Floor area restriction	0.439	0.497	0	1	
Sum of growth control measures	1.729	1.417	0	6	
Regulation intensity: CALURI	0.000	1.000	-1.221	3.014	
Spatial Coverage (#/Total)	Counties	Cities			
	53/58	410/482			

Note: Housing sales in California come from Zillow's ZTRAX database for 2012-2017. The growth control measures come from Questions 5a-5g, and Questions 9-11 in the Glickfeld and Levine (1992)'s California land use survey. California Land Use Regulation Index (CALURI) is defined as the total number of growth controls standardized to zero mean and unit variance.

Table 3. Marginal Effect of Regulation, Income and Land Cost on Tract Housing Price

	Dependent Variable: $ln\tilde{p}_{h,j}$		
	(1)	(2)	(3)
Home tract: Regulation Index lnR_j	0.016***	0.024***	0.081***
	(0.005)	(0.004)	(0.027)
Nearby tract: Regulation Index lnR_{-j}	0.027***	0.018***	0.072***
	(0.006)	(0.004)	(0.028)
Home tract: Land Cost Index $ln\tilde{p}_{l,j}$	0.199***	0.194***	0.318***
	(0.015)	(0.028)	(0.101)
Nearby tract: Land Cost Index $ln\tilde{p}_{l,-j}$	0.142***	0.149***	0.282**
	(0.017)	(0.030)	(0.113)
Home tract: Income Index $ln\tilde{Z}_j$	0.276***	0.277***	0.242***
	(0.009)	(0.013)	(0.023)
Nearby tract: Income Index $ln\tilde{Z}_{-j}$	0.105***	0.104***	-0.086**
	(0.012)	(0.017)	(0.042)
Constant	-0.032***	-0.031***	-0.059***
	(0.003)	(0.003)	(0.006)
Number of cities	376	376	376
Number of tracts	4,720	4,720	4,720
Spillover $\lambda \in [0,1]$		0.870	0.940
Hansen Test, p-value		0.216	0.312
Rank Test, p-value			0.001
Constrained coefficients	No	Yes	Yes
Variables are endogenous	No	No	Yes
Estimation method	OLS	GMM	GMM
Adjusted R-squared	0.741	0.687	0.469

Note: robust standard errors in the parentheses, clustered by county. * p<0.10, ** p<0.05, *** p<0.010. The dependent variables in all models are the California census tract housing price indices ($ln\tilde{p}_{h,j}$) for 2012-2017 which are the log real sale prices adjusted for the housing characteristics and the county-year fixed effect. lnx_j and lnx_{-j} stand for the home and nearby values of tract j respectively. The nearby values of a tract are constructed using the mean of the 30 nearest neighbors, with the inverse squared distance as the weight. Column 1 estimated using OLS without constraints on coefficients, while Columns 2-3 are estimated using GMM with the constraints on coefficients. Tract housing price indices are weighted by the number of sales for 2012-2017. Columns 1-2 treat the home and nearby values as exogenous, while Column 3 treats the variables as endogenous. The instrument variables include the z-scores of the home and nearby values of voting share of the Republican candidate in the 1992 general election (King et al., 2007), the college graduation share and the mean householder age from 2016 ACS 5-year estimates, the mean temperature and the mean sun hours in January from USDA (McGranahan, 1999). Columns 2-3 estimate 5 parameters ($\delta, \gamma, v, \eta, \lambda$) in addition to the constant term, so they are over-identified with the degree of freedom equal to 1 and 5 respectively. The null hypothesis of the Hansen over-identification test is instrument consistency. The null hypothesis of Kleibergen-Paap rank test is the instruments under-identify the model.

Table 4. First-Stage Relevance Test of Instrumental Variables

	(1) lnR_j	(2) $ln\tilde{Z}_j$	(3) $ln\tilde{p}_{l,j}$	(4) lnR_{-j}	(5) $ln\tilde{Z}_{-j}$	(6) $ln\tilde{p}_{l,-j}$
home tract: Republican voting share	-0.021***	0.043***	0.213***	0.113**	0.027***	0.033*
	(0.004)	(0.014)	(0.022)	(0.053)	(0.010)	(0.019)
home tract: college graduation share	0.027	0.400***	0.024*	-0.002	0.015**	-0.009
	(0.037)	(0.008)	(0.013)	(0.032)	(0.006)	(0.011)
home tract: mean householder age	0.013	0.177***	-0.002	0.010	0.013***	-0.005
	(0.026)	(0.006)	(0.009)	(0.022)	(0.004)	(0.008)
home tract: mean temperature in January	0.153	0.053	-0.619**	-0.758	0.172	-0.206
	(0.758)	(0.173)	(0.267)	(0.662)	(0.130)	(0.234)
home tract: mean sun hours in January	0.194	-0.054	0.025	0.689**	-0.033	0.255**
	(0.371)	(0.085)	(0.131)	(0.324)	(0.064)	(0.114)
nearby tract: Republican voting share	-0.190***	0.014	-0.229***	-0.339***	0.030***	-0.042**
	(0.064)	(0.015)	(0.023)	(0.056)	(0.011)	(0.020)
nearby tract: college graduation share	0.128***	-0.063***	0.285***	0.108***	0.325***	0.300***
	(0.046)	(0.010)	(0.016)	(0.040)	(0.008)	(0.014)
nearby tract: mean householder age	-0.126***	-0.067***	-0.038***	-0.082**	0.088***	-0.037***
	(0.042)	(0.010)	(0.015)	(0.037)	(0.007)	(0.013)
nearby tract: mean temperature in January	0.644	-0.508***	0.218	1.545**	-0.631***	-0.186
	(0.768)	(0.175)	(0.271)	(0.670)	(0.131)	(0.237)
nearby tract: mean sun hours in January	-0.124	0.231***	0.133	-0.608*	0.211***	-0.109
	(0.373)	(0.085)	(0.131)	(0.325)	(0.064)	(0.115)
Constant	-1.075***	0.423***	0.362***	-1.044***	0.412***	0.358***
	(0.098)	(0.022)	(0.035)	(0.086)	(0.017)	(0.030)
F Test of Instrument Exclusion	118	1,401	189	141	1,400	199
Adjusted R-squared	0.214	0.766	0.305	0.246	0.765	0.316
Number of tracts	4,720	4,720	4,720	4,720	4,720	4,720

Note: robust standard errors in the parentheses, clustered by county. * p<0.10, ** p<0.05, *** p<0.010. The dependent variables are the home and nearby values of the regulation index (Columns 1, 4), the income index (Columns 2, 5) or the land cost index (Columns 3, 6). Models are estimated using OLS weighted by the number of housing sales in a tract for 2012-2017. lnx_j and lnx_{-j} stand for the home and nearby values of tract j respectively. The nearby values of a tract are constructed using the mean of the 30 nearest neighbors, with the inverse squared distance as the weight. The instrument variables include the z-scores of the home and nearby values of voting share of the Republican candidate in the 1992 general election (King et al., 2007), the college graduation share and the mean householder age from 2016 ACS 5-year estimates, the mean temperature and the mean sun hours in January from USDA (McGranahan, 1999). The null hypothesis of the F test is that all instruments are irrelevant.

Table 5. Estimated Impact of Regulation, Income and Land Cost on Tract Housing Price by CSA Group

	Dependent Variable: $ln\tilde{p}_{h,j}$					
	Los Angeles-Long Beach		San Jose-San Francisco-Oakland		Rest of California	
	Exog. Indices (1)	Endo. Indices (2)	Exog. Indices (3)	Endo. Indices (4)	Exog. Indices (5)	Endo. Indices (6)
Home tract: Regulation	0.023***	0.069***	0.030***	0.077***	0.013***	0.041***
Index lnR_j	(0.003)	(0.015)	(0.002)	(0.007)	(0.003)	(0.006)
Nearby tract: Regulation	0.018***	0.061***	0.023***	0.068***	0.010***	0.037***
Index lnR_{-j}	(0.002)	(0.013)	(0.002)	(0.006)	(0.002)	(0.005)
Home tract: Land Cost	0.210***	0.317***	0.188***	0.338***	0.175***	0.239***
Index $ln\tilde{p}_{l,j}$	(0.006)	(0.043)	(0.011)	(0.054)	(0.011)	(0.025)
Nearby tract: Land Cost	0.161***	0.282***	0.145***	0.300***	0.135***	0.212***
Index $ln\tilde{p}_{l,-j}$	(0.004)	(0.038)	(0.009)	(0.048)	(0.008)	(0.022)
Home tract: Income	0.244***	0.158***	0.293***	0.359***	0.306***	0.288***
Index $ln\tilde{Z}_j$	(0.020)	(0.030)	(0.023)	(0.028)	(0.023)	(0.026)
Nearby tract: Income	0.045*	-0.118	0.234***	-0.110	0.103***	0.042
Index $ln\tilde{Z}_{-j}$	(0.026)	(0.083)	(0.042)	(0.104)	(0.028)	(0.046)
Constant	0.008	-0.006	-0.020***	-0.071***	-0.076***	-0.083***
	(0.005)	(0.023)	(0.005)	(0.008)	(0.005)	(0.006)
Number of cities/Total	155/185	155/185	103/129	91/129	118/168	118/168
Number of tracts	2,231	2,231	1,265	1,265	1,224	1,224
Spillover $\lambda \in [0,1]$	0.870	0.940	0.870	0.940	0.870	0.940
Hansen Test, p-value	0.899	0.119	0.118	0.081	0.308	0.120
Rank Test, p-value		0.001		0.042		0.046
Constrained coefficients	Yes	Yes	Yes	Yes	Yes	Yes
Variables are endogenous	No	Yes	No	Yes	No	Yes
Estimation method	GMM	GMM	GMM	GMM	GMM	GMM
Adjusted R-squared	0.721	0.476	0.779	0.642	0.613	0.560

Note: robust standard errors in the parentheses, clustered by county. * p<0.10, ** p<0.05, *** p<0.010. The dependent variables are the California census tract housing price indices ($ln\tilde{p}_{h,j}$) for 2012-2017 which are the log real sale prices adjusted for the housing characteristics and the county-year fixed effect. lnx_j and lnx_{-j} stand for the home and nearby values of tract j respectively. The nearby values of a tract are constructed using the mean of the 30 nearest neighbors, with the inverse squared distance as the weight. All models are estimated by GMM and weighted by the number of sales for 2012-2017. The regional models assume constant spillover measure λ from Table 3 as given and estimate 4 parameters $(\delta, \gamma, v, \eta)$ in addition to the constant term. Columns 1-2, 3-4, and 5-6 are estimated using the subsample of tracts in Los Angeles-Long Beach CSA, San Jose-San Francisco-Oakland CSA (including Modesto and Merced MSAs, added in 2018), and the rest of California respectively. Columns 1, 3 and 5 treat the home and nearby values as exogenous, while Columns 2, 4 and 6 treats the variables as endogenous. The instrument variables include the z-scores of the home and nearby values of voting share of the Republican candidate in the 1992 general election (King et al., 2007), the college graduation share and the mean householder age from 2016 ACS 5-year estimates, the mean temperature and the mean sun hours in January from USDA (McGranahan, 1999). The null hypothesis of the Hansen over-identification test is instrument consistency. The null hypothesis of Kleibergen-Paap rank test is the instruments under-identify the model.

Table 6. Average Marginal Effect of Regulation on City Housing Prices for 2012-2017: Largest Cities in California

			Exogenous Specifications			Endogenous Specifications				
City	Population (1k)	Income p.c. (1k)	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect	Total Effect	No. Reg. Policies	ρ
Los Angeles	3,919	29.9	0.040	-0.002	0.038	0.130	-0.007	0.123	5	0.88
San Diego	1,375	35.2	0.023	-0.001	0.022	0.078	-0.004	0.074	3	0.88
San Jose	1,009	37.8	0.054	-0.002	0.052	0.145	-0.006	0.139	5	0.91
San Francisco	850	55.6	0.054	0.000	0.054	0.145	-0.001	0.144	4	0.99
Fresno	514	20.1	0.023	0.000	0.022	0.078	-0.002	0.076	5	0.95
Sacramento	485	27.1	0.023	-0.001	0.021	0.078	-0.005	0.073	1	0.87
Long Beach	470	27.8	0.040	-0.002	0.038	0.130	-0.008	0.122	2	0.87
Oakland	412	35.0	0.054	-0.002	0.052	0.145	-0.006	0.138	2	0.91
Bakersfield	368	24.4	0.023	-0.000	0.023	0.078	-0.000	0.078	0	0.99
Anaheim	347	25.0	0.040	-0.006	0.035	0.130	-0.020	0.110	3	0.67
Santa Ana	334	17.0	0.040	-0.005	0.035	0.130	-0.019	0.111	4	0.69
Riverside	319	23.1	0.040	-0.004	0.037	0.130	-0.013	0.117	6	0.79
Stockton	301	20.7	0.054	-0.001	0.053	0.145	-0.002	0.142	0	0.97
Chula Vista	261	25.8	0.023	-0.002	0.021	0.078	-0.006	0.072	2	0.83
Irvine	247	45.1	0.040	-0.006	0.034	0.130	-0.022	0.108	1	0.64
Fremont	228	43.9	0.054	-0.006	0.048	0.145	-0.017	0.128	1	0.76
San Bernardino	215	14.9	0.040	-0.002	0.038	0.130	-0.007	0.123	3	0.88
Modesto	209	24.0	0.054	-0.003	0.051	0.145	-0.007	0.137	0	0.89
Fontana	205	20.3	0.040	-0.004	0.037	0.130	-0.013	0.117	2	0.79
Oxnard	205	21.1	0.040	-0.002	0.039	0.130	-0.006	0.124	3	0.90
CA City Mean	30.1	31.8	0.039	-0.008	0.031	0.118	-0.025	0.093	1.72	0.58

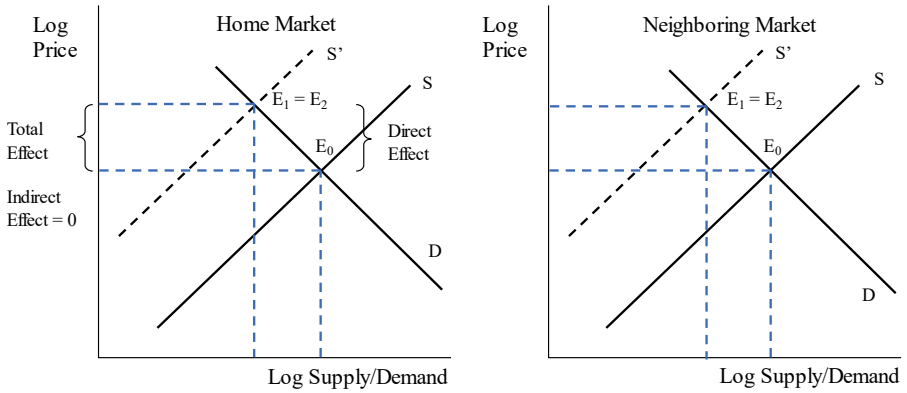
Note: Top 20 largest California cities (municipalities) by the 2016 population in the data sample are reported. Exogenous (endogenous) specifications refer to the models in Table 5 that treat the home and nearby indices of regulation, income and land cost as exogenous (endogenous). The direct, indirect and total effects are the marginal effects of one standard deviation increase in home regulation. The city-level effects are aggregated from the tract-level effects weighted by the number of housing sales for the period 2012-2017. Per capita income and population (in thousand dollars) are based on 2016 ACS 5-year estimates. Number of regulation policies counts the residential regulation out of 10 measures in the California Land Use Survey by Glickfeld and Levine (1992). $\rho \in [0,1]$ is the city mean of the tract share of neighbors located in the same city. The last row reports the mean statistics of all 482 California cities; population is the median city size; per capita income to be the city mean weighted by city population.

Table 7. Average Marginal Effect of Regulation on City Housing Prices for 2012-2017

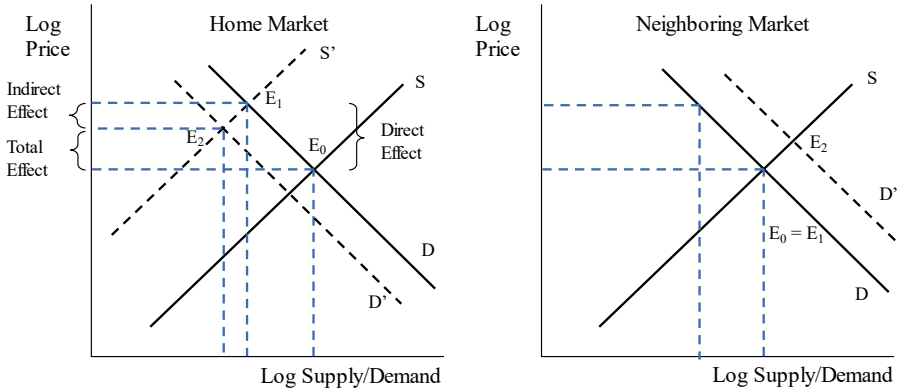
	Exogenous Specifications			Endogenous Specifications		
Group by	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect	Total Effect
CSA Group						
LA CSA	0.040	-0.008	0.032	0.130	-0.028	0.102
SF CSA	0.054	-0.011	0.043	0.145	-0.031	0.114
Rest CA	0.023	-0.004	0.019	0.078	-0.015	0.063
Per Capita Income						
< Q1	0.033	-0.006	0.028	0.107	-0.020	0.087
Q1-Q2	0.036	-0.006	0.030	0.112	-0.020	0.092
Q2-Q3	0.038	-0.007	0.031	0.118	-0.024	0.095
> Q3	0.046	-0.011	0.035	0.133	-0.035	0.098
No. of Reg. Policies						
0	0.034	-0.007	0.027	0.108	-0.026	0.082
1	0.039	-0.008	0.031	0.118	-0.025	0.092
2 or 3	0.038	-0.007	0.032	0.118	-0.022	0.096
4+	0.040	-0.006	0.034	0.120	-0.020	0.100
Population						
< Q1	0.035	-0.011	0.023	0.106	-0.037	0.069
Q1-Q2	0.038	-0.008	0.030	0.117	-0.027	0.090
Q2-Q3	0.041	-0.006	0.035	0.124	-0.021	0.103
> Q3	0.040	-0.004	0.036	0.123	-0.014	0.110
California Cities	0.039	-0.008	0.031	0.118	-0.025	0.093

Note: Exogenous (endogenous) specifications refer to the models in Table 5 that treat the home and nearby indices of regulation, income and land cost as exogenous (endogenous). The direct, indirect and total effects are the marginal effects of one standard deviation increase in home regulation, averaged across cities in a given group. The city-level effects are aggregated from the tract-level effects weighted by the number of housing sales for the period 2012-2017. *LA CSA* = Los Angeles-Long Beach CSA (LA CSA), *SF CSA* = San Jose-San Francisco-Oakland CSA (including Modesto and Merced MSAs, added in 2018); *Rest CA* = rest of California cities. Q1, Q2 and Q3 refer to the lower, mid and upper quantiles. Per capita income and population are based on 2016 ACS 5-year estimates. Number of regulation policies counts the number of growth controls out of 10 measures in the survey by Glickfeld and Levine (1992).

Figures

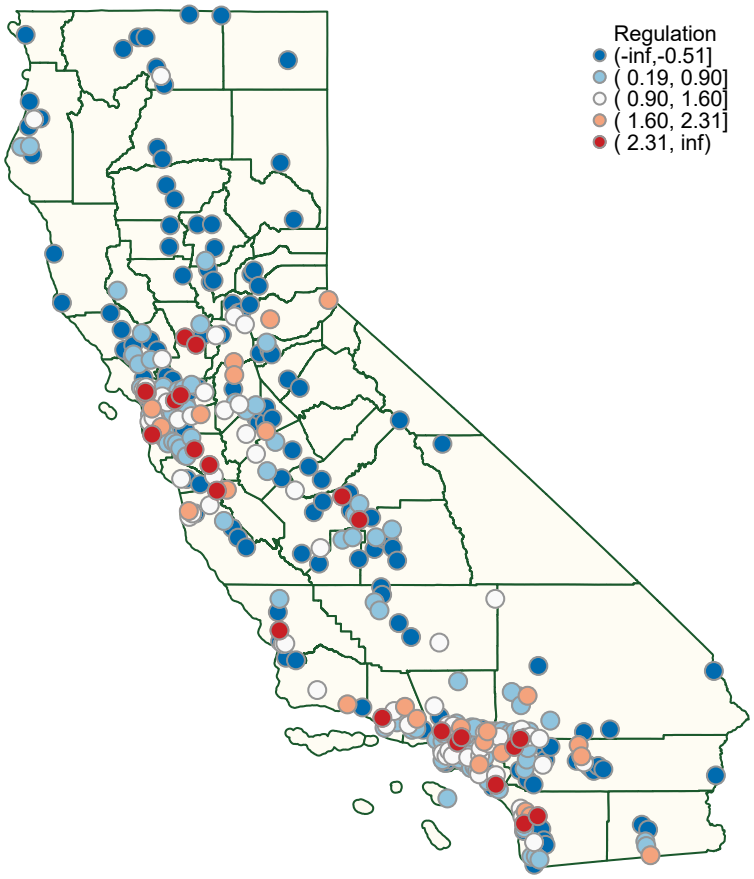


(a) home and neighboring markets in the same city



(b) home and neighboring markets in different cities

Figure 1: the effect of regulation on housing prices.



Note: residential regulation is based on Glickfeld and Levine (1992). Number of cities = 410

Figure 2: Spatial distribution of land use regulation in California. California Land Use Regulation Index (CALURI) is defined as the standardized sum of ten yes-no questions on whether a residential growth control measure was adopted by a city from the California land use survey by Glickfeld and Levine (1992). The cutoffs represent 25th, 50th, 75th, 90th and 95th percentiles of the CALURI distribution. A higher index value indicates more growth control measures thus higher regulation intensity. There are 410 jurisdictions in total.

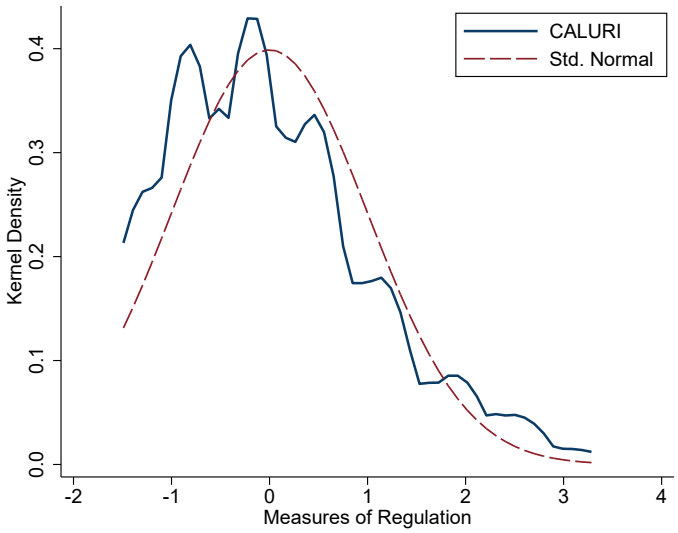
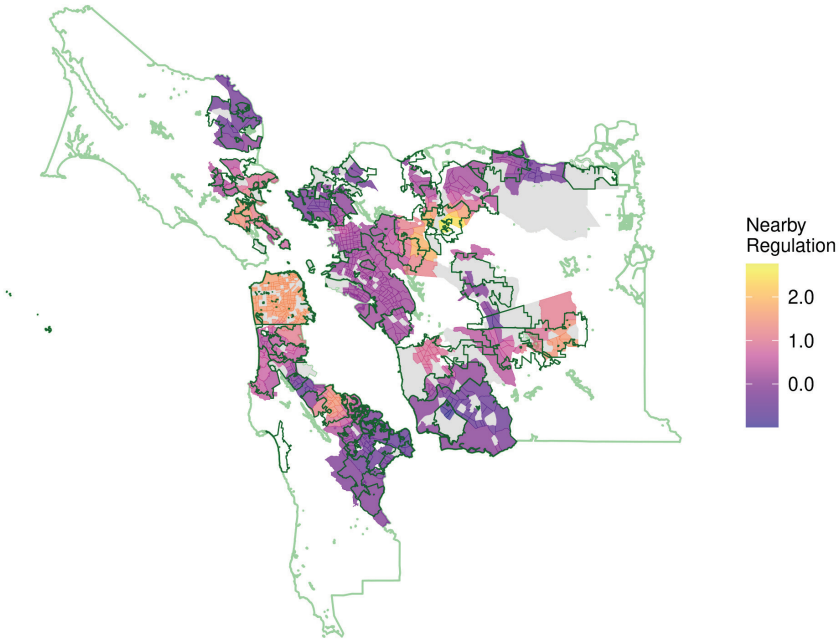
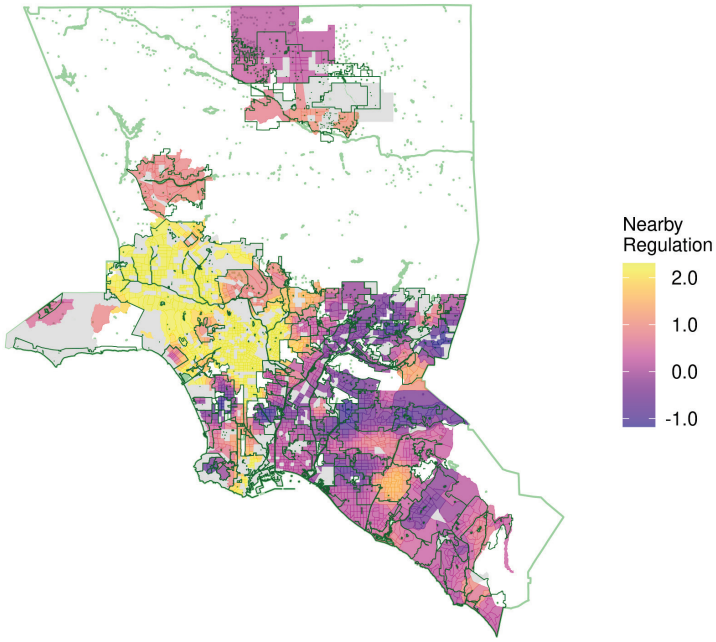


Figure 3: comparison of the kernel density of California Land Use Regulation Index (CALURI) and the standard normal distribution. is defined as the standardized sum of 10 yes-no questions on whether a residential growth control measure was adopted by a city from the California land use survey by Glickfeld and Levine (1992).



(a) San Francisco-Oakland-Berkeley, MSA



(b) Los Angeles-Long Beach-Anaheim, MSA

Figure 4: Nearby Regulation Index of two MSAs. The sample is census tracts matched to the incorporated cities. Colored area: the nearby regulation index (lnR_{-j}) of a tract j, defined as the mean of the home tract regulation index of 30 nearest tracts, using inverse-squared distance as the weight. The larger the index, the tighter the nearby regulation. Green line: city polygon net of water area. Grey area: tracts matched to cities but with missing values on regulation or housing sales. White area: tracts disjoint from (not matched to) cities.

Appendix

A.1 Derivation of the Housing Price Equation

First, rewrite the market clearing condition at location j, $q_j(p_h)h_j^D(p_{hj}) = H_j^S(p_{hj})$, as follows.

$$q_j = b_j p_j^{\frac{1}{1-v}}, \text{ where } b_j \equiv \frac{A_j^{\frac{1}{1-v}}}{\alpha Z_j}\left(\frac{v}{c_j}\right)^{\frac{v}{1-v}} \tag{9}$$

On the other hand, the total number of households are equal 1. We express p_j as a function of q_j. The equilibrium condition of location choices can be written as

$$q_j x = Z_j^{\varphi} p_{hj}^{-\varphi\alpha}, \text{ where } x = \sum_{j'} Z_{j'}^{\varphi} p_{hj'}^{-\varphi\alpha} \tag{10}$$

Combine two equations and eliminate p_{hj}.

$$q_j = (Z_j^{\varphi})^{1-\lambda} b_j^{\lambda} x^{-(1-\lambda)}, \text{ where } \lambda = \frac{\alpha\varphi(1-v)}{\alpha\varphi(1-v)+1} \tag{11}$$

We can show that there is a unique set of q_j that solve the system of equations. We first solve x from the equation. $\sum_j q_j(x) = 1$. The left side is a strictly decreasing function of x, while right side is constant. There is a unique x solving the equation. Given x, we can solve q_j using (11). With a binary location choice set $\{j, -j\}$, we can express q_j in a closed form that does not depend on p_{hj}.

$$q_j = \frac{(Z_j^{\varphi})^{1-\lambda} b_j^{\lambda}}{(Z_j^{\varphi})^{1-\lambda} b_j^{\lambda} + (Z_{-j}^{\varphi})^{1-\lambda} b_{-j}^{\lambda}} \tag{12}$$

Combined with (9), we derive an equation of the housing price, $lnp_{hj} = (1-v)(lnq_j - lnb_j)$. To derive the equation in the main text, we rewrite (12) as follows and make a linear approximation to the function $lnq_j(\Delta y_j) = -ln\big(1 + exp(-\Delta y_j)\big) \sim -ln(2) + 0.5\Delta y_j$ in the neighborhood of $\Delta y_j = 0$, where $\Delta y_j = \varphi(1-\lambda)\Delta lnZ_j + \lambda\Delta lnb_j$. The notation $\Delta lnx_j \equiv lnx_j - lnx_{-j}$ means the log deviation of x_j from x_{-j}. Hence,

$$\begin{aligned} \ln p_{hj} &= \delta \ln R_j + (1-v)\ln Z_j + \gamma v \ln p_{lj} \\ &\quad + \tfrac{1}{2}\lambda\left[\delta\Delta \ln R_{-j} - (\alpha^{-1} - 1 + v)\Delta \ln Z_{-j} + \gamma v \Delta \ln p_{l,-j}\right] + constant \end{aligned} \tag{13}$$

where $constant \equiv (1-v)\ln(\alpha/2) - v[\gamma ln\gamma + (1-\gamma)ln(1-\gamma) + lnv]$ is a constant term.

A.2 Housing Data and Filtering

For the housing data, we rely on the Zillow Transaction and Assessment Dataset (ZTRAX). The entire ZTRAX dataset contains more than 400 million public records from across the US and includes information on deed transfers, mortgages, property characteristics, and geographic information for residential and commercial properties. We are interested in the sale prices in the deed transfers and the housing characteristics in property assessment data in California.

Particularly, we restrict the data to the residential sales that have detailed documentation of housing characteristics. We include the following housing characteristics: the sale date, property use, the number of bedrooms, property age, the floor size and the mile distance from the city centroid. We exclude short sale transactions that will greatly affect the housing price metrics (FHFA, 2012).[26] We compute the property age, the property size and the distance from the city centroid that are not directly observable in ZTRAX. The property age is calculated as the difference of the transaction year and the effective built year. There are multiple fields measuring different aspects of the floor size, so we define the maximum value in those fields as the floor size. Other housing characteristics are available in ZTRAX, but they are either optionally reported or sparsely populated.[27]

Data Filtering and Variable Construction

The 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. We restrict the sample to the transaction with the sales prices more than 5,000 US dollars in California. While we focus on the housing sales for the period 2012-2017, we use the sales in a longer period 2000-2017 in the estimation of the hedonic pricing model.

We keep residential properties only and drop any commercial, manufactural, and foreclosure sales. Based on the Property Use Standard Code and Assessment Land Use Standard Code, we identify and focus on the housing types: single-family, townhouse/row house, cluster homes, condo/coop, and planned unit developments. A transaction can involve multiple parcels, we focus on transactions with a single parcel only. We only keep the sales that can be linked to the housing properties in the assessment data. About 90% of the transactions are matched to the assessment files. To estimate hedonic pricing model, we recode the number of bedrooms into 8 levels (0, 1, 2, 3, 4, 5, 6, 7+) and the property age into 10 levels (0-5, 6-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, > 70). There is no separate field to directly observe the floor size, so we construct the field as follows. We are able to observe the following fields relevant to the property size: 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. We calculate the maximum of the fields above and define it as the square footage of a property. The number of annual transactions in California ranges from 300,000 to 550,000, depending on the year. There are about 7.1 million transactions in total from about 1,400 places (either incorporated or not).

[26] We can identify those distress sales that occur at significant discounts compared with other transactions. FHFA HPI report in 2012Q1 (p.12) indicates that FHFA HPI includes short sales but distress sales can substantially affect housing prices metrics. FHFA plans on releasing a set of distress-free indexes but is constrained by the available data to identify distress sales especially for earlier transactions. One option suggested by FHFA is to construct the index using only transactions that are known definitively to be non-distressed, which we follow in the paper.

[27] 17% and 7% of the sales sample have missing value in lot size and the number of bathrooms respectively. Including these two variables does not significantly increase R-squared and change the hedonic prices of the existing variables but forces us to leave out many sales in small cities that are already thin in housing transactions.

Table A1. Hedonic pricing model: California Housing Sales

	Dependent Variable: log real sales price	
	β	*se*
Log Floor Size	1.035***	(0.001)
Bedroom: 1	0.154***	(0.004)
Bedroom: 2	0.051***	(0.003)
Bedroom: 3	0.005	(0.003)
Bedroom: 4	-0.066***	(0.003)
Bedroom: 5	-0.129***	(0.004)
Bedroom: 6	-0.204***	(0.004)
Bedroom: 7+	-0.256***	(0.009)
Townhouse/Row House	-0.097***	(0.002)
Cluster Home	0.080***	(0.003)
Condo/Coop	-0.046***	(0.001)
Planned Unit Development	-0.016***	(0.001)
Age: 6-10	0.026***	(0.001)
Age: 11-20	0.014***	(0.001)
Age: 21-30	-0.039***	(0.001)
Age: 31-40	-0.050***	(0.001)
Age: 41-50	-0.017***	(0.001)
Age: 51-60	0.009***	(0.001)
Age: 61-70	0.034***	(0.001)
Age: > 70	0.011***	(0.001)
Log Miles from tract to city centroid	-0.001***	(0.000)
Constant	4.888***	(0.008)
County-Year Fixed Effect	Yes	
Adjusted R-squared	0.605	
N	7,100,412	

Note: robust standard errors in the parentheses. * p<0.10, ** p<0.05, *** p<0.010. The base levels of factor or indicator variables (omitted to prevent co-linearity): no bedroom, property use is single-family, property age less than 5 years. The estimation is based on housing sales from 2000 to 2017 in California.

In Table A1, we report the hedonic pricing model we use to derive the housing price index. The model includes the county-year fixed effect that are fully interactive to absorb the macro impact at the county level. We aggregate the residual by tract for sales in the period 2012-2017 as the tract housing price index. Note that the indicator is negative for sales with more than 4 bedrooms. It is because we have separately controlled the floor size. Conditional on floor size, it indicates that dividing a given square footage into more bedrooms lower the housing prices.