ON THE ORIGINS OF LAND USE REGULATIONS:
THEORY AND EVIDENCE FROM US METRO AREAS[*]

Christian A. L. Hilber [a,b] and Frédéric Robert-Nicoud [b,c,d]

This version: October 2011

Abstract. We model residential land use constraints as the outcome of a political economy game between owners of developed and owners of undeveloped land. Land use constraints benefit the former group (via increasing property prices) but hurt the latter (via increasing development costs). More desirable locations are more developed and, as a consequence of political economy forces, more regulated. Using OLS as well as an IV approach that directly follows from our model we find strong and robust support for our predictions at the US metro area level. We conclude from our analysis that land use regulations are suboptimal as a consequence of redistributive motives.

JEL classification: H7, Q15, R52. **Keywords:** Land use regulations, zoning, land ownership, housing supply

1. Introduction

The spatial allocation of economic activities is crucial for productive and allocative efficiency. As evidence of this, eighty percent of the US population now lives in cities. These cities occupy only two percent of the nation's land but its occupants produce a disproportionate share of its output and virtually all of its innovations as a result of static and dynamic agglomeration economies. However, urban life also imposes hefty costs on its inhabitants in the form of more expensive dwellings, longer and more congested commutes or higher crime and poverty rates. Many of these costs and benefits like agglomeration economies are not priced by the market. Hence, the equilibrium outcome of individual location decisions will normally be inefficient.[1]

In principle, land use regulations can correct the positive and negative externalities associated with urbanization by separating (or mixing) land uses, altering population density or limiting city size. They can – through zoning – also ensure the provision of local public goods such as public parks. Land use regulations thus potentially have an important role to play to correct market failure, achieve efficiency and raise real incomes.[2]

Recent evidence for the US and the UK, however, casts some doubt on the proposition that existing forms of land use regulation are efficient. This evidence highlights the enormous *gross* costs of land use regulations in metro areas such as New York, San Francisco or London (Glaeser *et al.* 2005a, Cheshire and Hilber 2008). In these places, tight land use controls severely constrain the supply of space made available for new construction and thereby raise property prices enormously. The implied 'regulatory tax' appears to far exceed the negative externalities generated by new construction. This raises a natural question that this paper seeks to answer: *Do these regulations solve the allocative problem or are they driven by redistribution motives?* We find that redistribution motives are an important determinant of cross-metro area regulations in the US. Our results suggest that these motives do not only steer the voting behavior of local residents but also lobbying by non-residents – absentee landlords and owners of undeveloped land – who have a stake in the local land markets. Land use regulations may be suboptimal for a variety of reasons. Our theoretical analysis and corresponding supporting evidence lead us to conclude that current patterns of land use depart from first best and that redistributive motives are an important determinant of this phenomenon.

---

[1] The stylized facts are borrowed from Fujita and Thisse (2002), Burchfield *et al.* (2006) and Glaeser (2008).
[2] See e.g. Rossi-Hansberg (2004) for a recent theoretical development on this issue.

The paper also sheds light on the spatial patterns of land use regulations. Land use regulations are a recent phenomenon from a historical perspective. In the early 20th century, when only about a quarter of humans lived in urbanized areas, virtually no city had any zoning laws. San Francisco in 1880 and New York City in 1916 were early exceptions. Now that over half of the world population lives in cities, land use regulations are ubiquitous in all developed (and most developing) countries. Another striking phenomenon is that the tightness of land use regulations differs vastly across space. Land use regulations are often quite restrictive in highly attractive metro areas of a country (e.g. New York or San Francisco in the US, or the Greater London Area in the UK) but quite relaxed in many other places of the country (e.g. Houston or Pittsburgh in the US or Newcastle in the UK).

Our theory assembles ingredients from the urban economics, political economics and industrial organization literatures so as to propose an original explanation for why highly desirable metro areas are more tightly regulated than the rest of the country.

Our *influential landowner hypothesis* starts from the observation that one of the most salient economic effects of land use regulations is to increase the cost of future developments by restricting the amount of land zoned for development or increasing construction costs: either shifts the supply of new housing up and to the left, raising prices. This is good news for owners of the existing stock of developed land but the extra conversion cost is bad news for owners of hitherto undeveloped land (usually land developers). As the most obvious winners and losers, these two groups have strong incentives to influence the regulatory environment. As places grow over time, the share of developed land rises and with it the economic power of the owners of developed land relative to the owners of undeveloped land. Assuming that relative economic power is monotonically transformed into relative political power, the outcome is that places become increasingly regulated as they develop.

The cross-sectional equivalent of our main theoretical prediction is that places with a higher share of developed land should be more regulated than places with a lot of undeveloped land. This proposition should hold both within and across metro areas. In our model the degree of regulation pertaining to a metro area can be characterized as the average of the local outcomes. Figure 1 (panel a) plots the share of developed residential land (or SDL) in 1992 against the Wharton Residential Land Use Regulation Index (or WRLURI) in 2005 for our reference sample of the 93 largest US Metropolitan Statistical Areas (MSAs): the correlation $\rho = .31$ is statistically larger

than zero. Figure 1 (panel b) suggests that this pattern was already visible in the data in the late 1970s ($\rho = .34$).[3] Figure 1 (panel c) uses the aggregate property value per square meter of developable land in the MSA as an alternative to SDL; its correlation with WRLURI, $\rho = .25$, is also meaningful.

In the rigorous econometric work that constitutes the bulk of our paper we take the equilibrium predictions of the model to our MSA-sample and test the influential landowner hypothesis, controlling for various other explanations. These include the welfare economics view (regulation internalizes externalities or provides public goods), the majority voting view (the 'homevoter hypothesis' – more on this in section 2 below), political ideology and sorting by income. We also run a battery of robustness checks to ensure that our proxy does indeed capture our influential landowner hypothesis and not another explanation.

The empirical evidence in favor of our mechanism is threefold. First, the point of at least some regulation is to zone areas away from development; thus, if the cross sectional variation of the degree of restrictiveness of residential land use regulations was totally random, then we would expect to find a *negative* correlation between the regulatory variable (our dependent variable) and the SDL (our independent variable).[4] The very fact that, controlling for all other explanations, we find a positive and statistically significant relationship between the two variables in our OLS estimates provides strong evidence in favor of our paradigm. Put differently, if the welfare economics view and the homevoter hypothesis were the only mechanisms at play, the coefficient of SDL should be statistically (weakly) negative. Yet, it is anything but. A Two-Stage-Least-Square (TSLS) instrumental-variable strategy directly inspired by our theory provides the second piece of evidence. It addresses the reverse-causation (i.e., the downward bias of our SDL measure inherent in our OLS estimates) and omitted-variable issues. We confront our theory to the data in *both* stages of the econometric work to identify causal effects. The causal effect of SDL on regulatory restrictiveness finds strong and extremely robust support in our MSA-sample. We use no fewer than four different instrumental variables and we show that our results are robust to the exclusion of any individual instrument or any pair of instruments. Third, out-of-sample evidence

---

[3] We use the 1992 satellite data as in Burchfield *et al.* (2006) to construct our SDL measure. The WRLURI measure is due to Gyourko *et al.* (2008) and pertains to year 2005. We construct the share of developed residential land using 1976 data collected by aerial photographs and use the regulatory index constructed by Saks (2008) for 81 of the largest US MSAs.

[4] This logic finds strong support in the data, as we document in Section 6.

from other studies suggests that our paradigm also finds support *within* metro areas (Fischel 2004, Gyourko *et al.* 2008, Rudel 1989).

Finally, our quantitative exercises reveal that the causal mechanism we uncover is economically meaningful: granting Kansas City (the 9th least regulated city in our sample) with San Francisco's amenities and topography (SF is the 16th most regulated city in the sample) makes it the median city in terms of regulation. The implications of this set of results can hardly be overstated: *the current pattern of land use regulations in the US is neither efficient nor 'democratic'*, in that it goes beyond maximizing social well-being and does not conform to the wishes of a majority of households.

The rest of the paper is organized as follows. Section 2 reviews related work. Section 3 presents the model, with special emphasis on the theoretical predictions that we take to the data; it may be skipped by the empirically-oriented reader (its results are summarized above). Section 4 describes the data and our identification strategy and provides baseline results. Section 5 reports robustness checks. Section 6 revisits the effect of land use regulations on MSA growth and Section 7 concludes.

2. Related literature

The restrictiveness of land use regulations varies strongly across the United States (Glaeser *et al.* 2005a, b) and Europe (Cheshire and Hilber 2008). The first empirical contribution of this paper is to identify the origins of this cross-sectional variation. Evenson and Wheaton (2003) and Glaeser and Ward (2009) regress measures of various types of land use regulations on historical and other characteristics of Massachusetts towns. For instance, historical population density (1915 in the case of Glaeser and Ward) has a positive effect on current minimum lot size restrictions. They conclude from their exercise that 'the bulk of these rules seem moderately random and unrelated to the most obvious explanatory variables' (Glaeser and Ward 2009: 266). Our analysis shows that looking at aggregated measures of regulation across the major US MSAs reveals systematic patterns. The most closely related study to ours is Saiz (2010) and the papers complement each other in important ways. For each MSA in his sample, Saiz builds a measure of *developable land* and regresses WRLURI on this measure. His findings suggest that cities with a relatively small fraction of developable land are more regulated. By contrast, we create a measure of *developed*

land (SDL) that has developable land at the denominator.[5] Therefore, we take the physical constraints to expanding human settlements in existing MSAs as given and, guided by our theory, we aim to understand how the fraction of land *actually developed* influences regulation, emphasizing political economy mechanisms. Our model also suggests that the most desirable places should indirectly be the most regulated. This accords well with Glaeser *et al.* (2005a), who find that the regulatory tax is highest in Manhattan and in the Bay area (exceeding 50% of house values), while they find no evidence for a regulatory tax in places such as Pittsburgh or Detroit. In addition, our paper enables us to shed a new perspective on some of the results unveiled by Burchfield *et al.* (2006). For instance, they find that cities with better natural amenities sprawl more than others – likely because of minimum lot size restrictions that reduce the capital-to-land ratio. In the model, we attribute this phenomenon to endogenous land use constraints; in our empirical work, we find that locations with more desirable amenities are more developed and more regulated.

In the US, land use regulations are largely determined by local planning boards whose members are elected by local residents. Accordingly, the dominant political economics view suggests that local land use regulations correspond to the wishes of a majority of voters (Fischel 2001, Ortalo-Magné and Prat 2007). Fischel's 'homevoter hypothesis' postulates that homeowners – in contrast to renters – favor regulations because it raises their property value and, in turn, suggests that jurisdictions with a larger share of homeowners should be more regulated. Available evidence is strongly suggestive that 'homevoters' (and conservationists) are influential in regulating land use *locally* (e.g. Dehring *et al.* 2008). However, these influences may have less explanatory power in explaining *across metro area* differences in planning restrictiveness. As Figure 1 (panel d) illustrates, the homeownership rate and land use regulations are significantly *negatively correlated* in our sample of MSAs (the homeownership rate is roughly twice as high in the loosely regulated Pittsburgh and Houston compared to the tightly regulated New York).

Like Fischel (2001) and Ortalo-Magné and Prat (2007), we understand land use regulations as the outcome of political economic forces at work. By contrast, we assume that landowners (who include the group of homeowners) lobby and influence planning boards and that these implement policies that maximize land value as a result. Such policies, by catering to 'land based interests'

---

[5] Saiz (2008) excludes water bodies, wetlands and slopes of 15% or more to construct his measure of developable land. We use a comparable dataset, except that we base our definition of non-developable land on land cover data. See also Burchfield *et al.* (2006) and Hilber and Mayer (2009).

(Molotch 1976), may then hurt consumers and overall welfare, as stressed by Brueckner (1995) and Helsley and Strange (1995).[6]

A dynamic interpretation of our model is also consistent with the findings of Fischel (2004), Rudel (1989) and Gyourko *et al.* (2008). According to Fischel (2004), land use regulations originate within larger cities and then zoning spreads quickly to the suburbs and surrounding towns as the city grows. The most direct evidence that the timing and restrictiveness of zoning is tied to the distance from the central city comes from Rudel (1989) who shows that the Connecticut municipalities located at a greater distance to New York City adopted land-use laws later than those closer to the Big Apple. Gyourko *et al.* (2008) find that municipalities within MSAs tend to be more highly regulated than their counterparts outside of MSAs.

In our empirical analysis we find that regulations slow down new development, confirming the findings of Quigley and Raphael (2005) for California or Glaeser and Ward (2009) for Boston. There is thus a two-way relationship between regulation and urban development. Our estimates imply that a one standard deviation difference in MSA-level SDL in 1976 results in a roughly 2 percentage point-decrease in the growth of housing supply between 1990 and 2000 via differentially affecting regulatory restrictiveness during the late 1970s and 1980s. This accounts for about 15 percent of the growth in new construction during that period.

We also contribute to the theoretical literature in two ways. First, our combination of a discrete choice model and of a standard monocentric city (MCC) model for 'macro' (across-city) and 'micro' (within city) location decisions, respectively, is unique in the urban economics literature. This combination provides a useful generalization of the currently available extreme versions of the MCC model, whereby each MSA is either fully isolated ('closed') – the population supply to each city is inelastic and utility varies across MSAs – or small and fully 'open', that is, the utility level is exogenous and population supply is infinitely elastic (Brueckner 1987). In our model, both MSA sizes and average utility levels vary across MSAs and are determined endogenously. Second, jurisdictions set their policies non-cooperatively and a jurisdiction's policy spills over to other communities because consumers are mobile, as in Brueckner (1995) and Helsley and Strange (1995). In our model, the Nash equilibrium in land use regulations exists under fairly mild conditions; Helsley and Strange (1995) use a model in which at least one jurisdiction is inactive and acknowledge that 'it is not possible to consider population controls when all

---

[6] Epple *et al.* (1988) also develop a model in which owners of developed land ('early arrivals' in their terminology) impose policies that hurt later entrants when they control the political agenda.

communities are active without substantially modifying [their] model' (p. 456). We propose one such modification.

3. The model

The set of players and the timing of the game are as follows. In stage 1, the planning boards of a set of local jurisdictions (which differ in exogenously given characteristics) simultaneously choose a zoning policy, taking the other planning boards' choices as given. Each jurisdiction belongs to exactly one MSA and the set of MSAs is a partition of the set of jurisdictions. In stage 2, households make location decisions of two kinds. They first choose a jurisdiction where to live; a bidding process for land then allocates households within each jurisdiction. Finally, payoffs are realized. The equilibrium concept is a (subgame perfect) Nash equilibrium in zoning policies: all agents are rational and forward-looking.[7] We now formally describe the set of players, their strategy sets and their payoff functions.

3.1. Households' location choice

In stage 1, a continuum of H households indexed by $h \in [0, H]$ allocate themselves to a number J > 1 of jurisdictions indexed by $j \in \Im \equiv \{1,...,J\}$. Households established in jurisdiction j derive utility u_j. Following the *Random Utility Theory*, which finds its origins in psychology (Thurstone 1927), we assume that u_j is a random variable and we model the fraction f_j of households that choose to live in jurisdiction j as

$$f_j = \Pr\left\{u_j = \max_{k \in \Im} u_k\right\}. \tag{1}$$

Specifically, the household-specific realization of u_j, denoted as $u_j(h)$, has a common component V_j and an idiosyncratic, random households-specific component $\varepsilon_j(h)$ with cumulative density G. These components add up as

$$u_j(h) = \ln V_j + \varepsilon_j(h), \qquad \varepsilon_j(h) \sim \text{i.i.d. } G(\cdot). \tag{2}$$

The common component V_j is deterministic and summarizes the costs and benefits from living in jurisdiction j, expressed in monetary units; we refer to it as the deterministic utility or real income. The idiosyncratic component $\varepsilon_j(h)$ is random (Manski 1977, Anderson *et al.* 1992) and summarizes the idiosyncratic utility that h derives from consuming local amenities. Households

---

[7] As in all papers in which interest groups lobby the executive power in order to raise their rent, we assume entry of developers away as entry erodes rents. Baldwin and Robert-Nicoud (2007) show how both ingredients can be simultaneously included in a dynamic stochastic model. This issue is beyond the scope of this paper so we omit it for simplicity.

are heterogeneous in their appreciation of these amenities.[8] In order to get simple, explicit solutions, we assume that the ε_j's are iid distributed according to the double exponential distribution with mean zero (so that the average and median households are a-priory indifferent about where to live) and variance $\sigma^2 \pi^2 / 6$. The resulting location choice corresponds to a standard multinomial logit model (Anderson *et al.* 1992). The degree of household heterogeneity σ, which governs the sensitivity of f_j with respect to the utility differentials, is not observable in the data and we therefore normalize it to unity in order to simplify the equilibrium expressions.[9] As a result of these assumptions and of (2), the location choice probabilities in (1) are equal to

$$f_j = \frac{\exp\left(\ln V_j / \sigma\right)}{\sum_{k \in \Im} \exp\left(\ln V_k / \sigma\right)} = \frac{V_j}{\sum_{k \in \Im} V_k} = \frac{1}{J} \frac{V_j}{\bar{V}} \tag{3}$$

with $f_j = 0$ or $f_j = 1$ in an obvious manner if the right-hand-side (RHS) above falls outside the unit interval and $\bar{V} \equiv \sum_{k \in \Im} V_k / J$ (throughout the paper, we use 'upper bars' to denote averages across jurisdictions). An implication of (3) is that jurisdictions that command a higher-than-average deterministic utility V_j attract more households than the average jurisdiction (i.e. $V_j > \bar{V} \Rightarrow f_j > 1 / J$).

We assume that the common, deterministic component V_j is a function of economic and non-economic variables pertaining to jurisdiction j: let

$$V_j = a_j + w - c_j - t_j, \tag{4}$$

where a_j is a measure of the observable *quality* of local amenities (converted in monetary units), w denotes the household's income, c_j captures the monetary costs of living associated with j (henceforth 'urban costs') and t_j is a local 'regulatory tax' levied on residents (more on this below). Household h's global appreciation of jurisdiction j's amenities is thus equal to $a_j + \varepsilon_j(h)$. a_j summarizes the attributes of local amenities that can be ranked across the average population (hence the term 'quality').[10]

---

[8] Consider the following examples to fix ideas. Within the metro area of Los Angeles, people-watchers prefer to live in Venice Beach and recreation and golf lovers in Bel Air, ceteris paribus. Similarly, at the more aggregate level, skiers prefer local jurisdictions in the Boulder MSA, whereas windsurfers prefer locations in the San Francisco MSA In the empirical section below, we use an MSA's access to a major coast and its mild winter temperatures as the natural amenities people care about (Rosen 1979, Roback 1982, Glaeser *et al.* 2001, Rappaport 2007).

[9] See Appendix A for expression (3) with an arbitrary σ.

[10] As an illustration, let us compare a representative jurisdiction in the Boulder MSA to a representative jurisdiction in the San Francisco MSA ($j = B, SF$). Ranking access to mountain slopes versus access to the ocean is clearly a mater of individual taste but most people prefer mild to very cold winter temperatures: the latter implies $a_{SF} > a_B$. Put

In this paper, the land market outcomes play the central role, so we treat a_j and w as parameters but we endogenize the urban cost as follows; we show in Appendix C.2 how the qualitative properties of the model are unaffected by relaxing the parameterization of a_j and w to allow for agglomeration and other external (dis)economies. Assume that jurisdiction j is a linear monocentric city (Alonso 1964), in which the per-unit-distance commuting cost is equal to τ and the unit cost of converting land into housing is constant and equal to m_j. Then, if H_j households live in j, c_j is equal to $H_j\tau + m_j$.[11] Substituting this expression for c_j in (4) yields:

$$V_j = \omega_j - H_j\tau - t_j; \qquad \omega_j \equiv a_j + w - m_j, \tag{5}$$

where ω_j summarizes the 'fundamental' (parametric) determinants of welfare in jurisdiction j. We say that a jurisdiction characterized by a high ω_j is a 'desirable' location *ex ante* (or that it is *fundamentally desirable*). The urban cost τH_j and regulatory cost t_j are endogenous to the model. The former rises with jurisdiction size; the latter is the outcome of the political economy game of section 3.2 below. Plugging (5) into (3) establishes that the fraction f_j of households wishing to live in jurisdiction j is decreasing in the regulatory tax it levies, decreasing in its level of congestion and increasing in the desirability of jurisdiction j, which includes amenities and wages.

Define the vectors $\mathbf{t} \equiv [t_1,...,t_J]'$ and $\mathbf{H} \equiv [H_1,...,H_J]'$. Households treat $\mathbf{t}$ and $\mathbf{H}\tau$ as parameters. We define as a *spatial equilibrium* for $\mathbf{H}$ a situation in which, given the induced equilibrium values of (3) and (5), no household wishes to relocate to another jurisdiction. Formally, the actual fraction of households living in j, H_j/H, must be equal to f_j. Using (5) and (3), we may define the spatial equilibrium as

$$\forall j \in \mathfrak{I}_+: \qquad f_j = \frac{H_j}{H} \implies \frac{H_j}{H} = \frac{1}{J}\frac{\omega_j - t_j - H_j\tau}{\bar{\omega} - \bar{t} - \bar{H}\tau} . \tag{6}$$

---

differently, the distribution of $a_{SF} + \varepsilon_{SF}$ stochastically dominates the distribution of $a_B + \varepsilon_B$ and, keeping the economic attributes of V_B and V_{SF} in (4) equal, a larger fraction of households would then choose to live in a municipality in the San Francisco MSA rather than in one in the Boulder MSA.

[11] To see this, assume that all city dwellers consume one unit of land and that the central business district (CBD) is located at $d = 0$, so that a city of size H_j stretches out from 0 to H_j. Assume further that that the unit cost of converting farm land for housing consumption is equal to m_j. Without loss of generality, we assume that the opportunity cost of land at the urban fringe is zero. Each city dweller commutes to the CBD at a constant per unit distance cost $\tau > 0$. The city residential land market is at an equilibrium when the sum of commuting costs and land rent are identical across city locations (a no arbitrage condition), thus the equilibrium bid rent schedule (net of conversion costs m_j) is $r(d) = (H_j - d)\ \tau$. As a result, the urban cost is $c_j = H_j\tau + m_j$, the aggregate land rent is equal to $H^2\tau/2$ and $\omega_j = a_j + w - m_j$, as in (5).

9

That is, the fraction of people living in j is increasing in the local well-being net of the regulatory tax and urban costs and decreasing in the well-being net of regulatory tax and urban costs of other jurisdictions. Since households directly consume one unit of land for housing purposes in the linear MCC model, the equilibrium H_j is also the equilibrium fraction of developed land in jurisdiction j. We readily obtain the following:[12]

Proposition 1 (existence and uniqueness of the spatial equilibrium). Assume that the fraction of households that wish to live in jurisdiction j is given by (3) and that the observable real income is given by (5). Then the spatial equilibrium defined in (6) exists and is unique.

Proof. See Appendix A.

Figure 2 illustrates the equilibrium concept. The downward-sloping schedule illustrates the fact that as jurisdictions get more populated they get more congested and thus less desirable, ceteris paribus. A higher ω_j and/or a lower t_j shift this schedule upwards. Thus, at equilibrium more desirable locations have more households. To see this formally, solve (6) explicitly for H_j:

$$\frac{H_j}{H} = \frac{1}{J}\frac{\omega_j - t_j}{\overline{\omega} - \overline{t}} = \frac{1}{J}\left[1 + \frac{\left(\omega_j - \overline{\omega}\right) - \left(t_j - \overline{t}\right)}{\overline{\omega} - \overline{t}}\right]. \tag{7}$$

Together, (5) and (7) imply

$$V_j = \left(\omega_j - t_j\right)\left[1 - \frac{H\tau}{J\left(\overline{\omega} - \overline{t}\right)}\right]. \tag{8}$$

For now, we assume $t_j < \omega_j$ for all j (this will be true at equilibrium) and we impose $H\tau < J\left(\overline{\omega} - \overline{t}\right)$ so that $V_j > 0$ and $H_j \in \left(0, H\right)$ for all j.[13] Several aspects of the spatial equilibrium characterized by (7) and (8) are noteworthy. First, jurisdictions that receive more households than the average $\overline{H}$ either are fundamentally desirable, or have a low regulatory tax, or both, relative to the average jurisdiction. Second, precisely because such desirable places end up being more congested in equilibrium, households obtain a real income that is only a fraction of the local fundamental desirability net of regulatory taxes. Third, *all jurisdictions yield about the same welfare ex post*: congestion and labor mobility between jurisdictions together ensure that in

---

[12] The normalization $\sigma = 1$ has no bearing on Proposition 1; see Appendix A for a proof.

[13] The former technical condition former assumes away corner solutions for simplicity. The latter require aggregate urban costs and aggregate taxes to be smaller than aggregate well-being so that equilibrium consumption is positive; these are needed for (3) to be well-defined. These conditions are implied by (14) below when the t_j's are endogenous.

each jurisdiction the marginal household is indifferent between staying put and living in its next best alternative. All infra-marginal households are strictly better off in the jurisdiction of their choosing. To get a sense of this, consider (3) at the limit $\sigma \to 0$ (homogeneous population). In this deterministic case, $f_j > 0$ if and only if $V_j = \max_k V_k$. That is to say, all jurisdictions with a strictly positive equilibrium population yield the same (deterministic) utility. By contrast, for $\sigma > 0$ households do not enjoy exactly the same real wage everywhere at the spatial equilibrium because they are willing to forego some economic benefits to live in jurisdictions that offer the non-economic amenities that they value.

3.2. Planning boards choose regulation

We assume that each jurisdiction j has a planning board that regulates the use of land. Land use restrictions can take many forms. To simplify the analysis, we assume that the main effect of such regulations is to increase the individual cost of living in the jurisdiction of each household by t_j. We interpret t_j as a 'regulatory tax' (Glaeser *et al.* 2005a, b) and assume that it is capitalized into the price of developed land (Oates 1969, Palmon and Smith 1998). This capitalization effect captures in a parsimonious way the fact that land use regulations reallocate the local demand for land away from potential new developments to existing ones (keeping total demand for land H_j constant). In addition to this *direct* effect that benefits owners of developed land at the expense of owners of undeveloped land, a higher regulatory tax in j decreases the desirability of j as per (5), which in turn reduces the equilibrium population size and equilibrium amount of developed land in jurisdiction j as per (6); the former effect reduces the average land rent in the jurisdiction.[14] Thus, the overall *indirect* effect tends to hurt all landowners.

Here, we depart from the standard literature by assuming that the planning board caters to the landowners' interests. In the wake of Bernheim and Whinston (1986) and Dixit *et al.* (1997), we assume that the owners of developed land and the owners of undeveloped land (or land developers) form two competing lobbies that influence the planning board by way of lobbying contributions. Specifically, we assume that the planning board maximizes aggregate lobbying contributions $C_j \equiv \sum_\Lambda c_j^\Lambda$, with $\Lambda \in \{$owners of developed land, land developers$\}$. This objective function conveys the idea that the planning board caters only to the interests of land stakeholders.

---

[14] This point is easily made: $TR_j \equiv \int_0^{H_j} r(d)\mathrm{d}t = \tau H_j^2 / 2$, so the average land rent (defined as $AR_j \equiv TR_j / H_j$) is linearly increasing in H_j. It is not difficult to show that AR_j is increasing in H_j also in nonlinear MCC models (Brueckner 1987).

We return to this bold assumption below. Note that land stakeholders include absentee landlords, local landlords, land developers, homeowners and even renters when rent controls are in place (de facto, rent controls act as way to share land rents between the owner and the renter).

We also assume that each group offers a 'menu' of contributions to the planning board, contingent on the degree of regulation t_j actually chosen so that $c_j^\Lambda = c_j^\Lambda(t_j)$. Many contribution schemes are possible (and thus many Nash equilibria exist), but Bernheim and Whinston (1986) show that the set of best responses of each lobby to *any* contribution scheme chosen by the other players includes a linear schedule of the form $c_j^\Lambda(t_j) = R_j^\Lambda(t_j) - c_j^\Lambda$, where R_j^Λ is the aggregate land rent pertaining to lobby Λ and c_j^Λ is a constant determined at equilibrium.[15] Thus, the owners of developed land are offering a contribution schedule that is increasing in the degree of regulation; land developers' contributions are decreasing in t_j. The literal interpretation of this working hypothesis is that stakeholders bribe the planning boards in order to sway its decisions. We may also understand the word 'influence' in a broader and more benign sense, such as pressure groups acting as experts and conveying useful information to the executives. By using legal contributions so as to buy access to executives (Austen-Smith 1995, Lohmann 1995), pressure groups provide credible information to legislators.[16] As a result of these assumptions, the planning board maximizes total land rents plus the regulatory tax revenue:

$$R_j(\mathbf{t}) \equiv \int_0^{H_j} \tau\left[H_j(\mathbf{t}) - x\right] dx + t_j H_j(\mathbf{t})$$
$$= \frac{\tau}{2}\left[H_j(\mathbf{t})\right]^2 + t_j H_j(\mathbf{t}), \tag{9}$$

where $H_j(\mathbf{t})$ is given by (7). Two aspects of the program $\max_{t_j} R_j(\mathbf{t})$ are noteworthy. First, the planning board gives equal weight to the cost and benefit to landowners of raising the local regulatory tax t_j. Second, maximizing only the first component of $R_j(\mathbf{t})$ above and ignoring strategic interactions among jurisdictions would lead planning boards to choose the first best

---

[15] The timing of the contribution game is as follows. The lobbies (the 'principals') move first and simultaneously, the planning board (the 'agent') then chooses to accept the contributions or not and, contingent on accepting some, both, or no contributions, chooses t_j. Since the principals move first, at equilibrium they choose c_j^Λ so as to ensure that the agent accepts the contribution and enforces a regulation that is closer to the interests of lobby Λ. These linear contribution schedules also have the desirable property to produce the unique 'coalition proof Nash equilibrium' of the game. Goldberg and Maggi (1999) show how a model in which t_j is set by cooperative Nash bargaining produces a similar policy outcome.

[16] The non-partisan research group Centre for Responsive Politics (CPR) reports that the National Association of Realtors topped the CPR's top-20 list of Political Action Committees contributing to federal candidates both in 2005/6 and 2007/8. The National Association of Home Builders ranked 4th and 12th, respectively.

policy by the Henry George theorem; this would be $t_j = 0$ for all j since there is no market failure in the model.

3.3. Subgame perfect equilibrium

We solve for a Nash (subgame perfect) equilibrium (SPE henceforth) in regulatory taxes. Thus, j's planning board chooses $t_j \in \mathfrak{R}_+$ so as to maximize (9) subject to (7) taking the vector $\mathbf{t^0_{-j}} \equiv \{t^0_k\}_{k \in \mathfrak{I}\backslash\{j\}}$ as given (the superscript '0' pertains to equilibrium values). Then the first order condition for this program may be written as

$$\frac{\mathrm{d}}{\mathrm{d}t_j} R_j(t_j, \mathbf{t^0_{-j}})\bigg|_{t_j = t^0_j} = H^0_j + \left(H^0_j \tau + t^0_j\right)\frac{\mathrm{d}}{\mathrm{d}t_j} H_j(t_j, \mathbf{t^0_{-j}})\bigg|_{t_j = t^0_j} \leq 0, \qquad t^0_j \geq 0, \qquad (10)$$

with complementary slackness; H^0_j is the equilibrium jurisdiction size, namely (7) evaluated at the Nash tax vector $\mathbf{t^0} \equiv [t^0_1, ..., t^0_J]'$ (also $\mathbf{H^0} \equiv [H^0_1, ..., H^0_J]'$). The RHS consists of the direct effect (keeping the population constant) and the indirect effect (allowing it to vary in response) of an increase in the specific regulatory tax. At an interior equilibrium, the cost and benefits of raising t_j are equal at the margin.

Let

$$f^0_j \equiv \frac{1}{J}\frac{\omega_j - t^0_j}{\bar{\omega} - \bar{t}^0} \in [0,1] \qquad (11)$$

be the equilibrium fraction of people settling in jurisdiction j. Then, using (7) and (11) and assuming that the parameters of the model are such that the resulting H^0_j's in (10) are all interior (precise conditions to follow in (14)), we may develop the first order condition (10) to get an equilibrium relationship between population size $f^0_j H$ and the regulatory tax t^0_j:[17]

$$t^0_j = \kappa^0_j \omega_j - \left(1 - \kappa^0_j\right)f^0_j H\tau, \qquad 1 > \kappa^0_j \equiv \frac{1}{2 - f^0_j} > \frac{1}{2}. \qquad (12)$$

That is, the equilibrium regulatory tax absorbs part, but only part, of the jurisdiction's desirability ω_j and of its urban cost $f^0_j H\tau$. Plugging (11) into (12) yields a system of J third-order polynomials in the components of $\mathbf{t^0}$. Plugging (12) into (5) yields an expression for equilibrium deterministic utility:

---

[17] See Guide to calculations in Appendix C.1 (not intended for publication) for details.

$$V_j^0 = \left(1 - \kappa_j^0\right)\omega_j + \kappa_j^0\left(-f_j^0 H\tau\right). \tag{13}$$

That is to say, V_j^0 is a weighted sum of the fundamental desirability and the urban costs pertaining to j.

All the properties of the spatial equilibrium continue to hold at the SPE that (11) and (12) characterize. Three additional properties resulting from strategic interactions are noteworthy (formal proofs follow). First, the equilibrium regulatory tax increases in own desirability and decreases in the desirability of other jurisdictions. Second, this effect is stronger, the lower the urban costs τH. This cross-effect arises because, when urban costs are large, cross-jurisdiction differences along other dimensions matter relatively less for households' location choices. Finally, places that are more desirable are more developed at equilibrium, *despite being more regulated*. That is, endogenous regulation does not change the ranking of jurisdictions according to their ω_j.

The second derivative of R_j with respect to t_j is negative for *any* **t**; so we may write:

Proposition 2 (existence of a SPE in the tax setting game). The subgame perfect equilibrium characterized by (11) and (12) exists.

Proof. See Appendix A.

We can also formally establish the subgame perfect equilibrium properties of our model:

Proposition 3 (properties of the SPE). Assume:

$$\min_{k\in\mathfrak{J}} \omega_k > \frac{H\tau}{J-1}. \tag{14}$$

Then (11) and (12) imply the following properties for any SPE:

    (i) Places that are fundamentally more desirable are more *developed*: $\omega_j > \omega_k \Rightarrow H_j^0 > H_k^0$;

    (ii) Places that are more developed are more *regulated*: $H_j^0 > H_k^0 \Rightarrow t_j^0 > t_k^0$.

Proof. See Appendix A.

These are the properties that we test in section 4: (i) is the first stage of our TSLS instrumental variable (IV) approach, whereas (ii) is the second stage. The equilibrium properties of the model that we do not directly test include:

Corollary 3.1 (further properties of the SPE). Assume that (14) holds. Then:

    (iii) Regulatory taxes are strictly positive for all j;

(iv) The fundamental amenities of a jurisdiction are not fully capitalized into the regulatory tax: $\omega_j > \omega_k \Rightarrow \omega_j - t_j^0 > \omega_k - t_k^0$;

(v) Despite being more developed and more regulated, fundamentally more desirable places command a larger deterministic utility: $\omega_j > \omega_k \Rightarrow V_j^0 > V_k^0$.

Proof. See Appendix A.

The parameter restriction (14) ensures that all jurisdictions have positive population and regulation in equilibrium. We make this assumption for analytical convenience only. Relaxing it would require us to replace the strict inequalities in Proposition 3 by weak inequalities.

3.4. Lobbying, voting and benevolent planning

So far we have been assuming that planning boards only cater to 'land-based interests'. A utilitarian urban planner would choose t_j so as to maximise total welfare $H_j(t_j)V_j(t_j) + R_j(t_j)$; one that seeks to please voters may maximise the welfare of the current residents $H_j V_j(t_j) + \alpha R_j(t_j)$, where α is the share of land rents earned by local residents. A parsimonious way to model the behaviour of a planning board that responds simultaneously to social welfare, electoral considerations and lobbying pressure is to assume that it maximises the weighted sum

$$\Omega_j(t_j) \equiv b\left[H_j(t_j)V(t_j) + R_j(t_j)\right] + v\left[H_j V(t_j) + \alpha R_j(t_j)\right] + (1-b-v)C_j(t_j),$$ where parameters b and v respectively capture the 'benevolence' of the planning board and its responsiveness to the average voter's well-being. Imposing $b = v = 0$ as in e.g. Krishna (1998) gives us $\Omega_j(t_j) = C_j(t_j)$ and yields clear cut results.

Removing this assumption has no bearing on Proposition 2 (existence of the SPE) but it has the following implications. First, insofar as the regulatory tax benefits landowners (at least when they are low to start with) at the expense of other residents and voters, the equilibrium regulatory taxes tend to be lower than in (10) if b, $v > 0$. To see this, rewrite $\Omega_j(t_j)$ as $\Omega_j(t_j) = [1-v(1-\alpha)]R_j(t_j) + \{vH_jV_j(t_j) + bH_j(t_j)V_j(t_j)\}$ and note that the term in the curly bracket, which is new with respect to (9), is decreasing in t_j by (5) and (7). Then $\partial t_j^0 / \partial b < 0$ and $\partial t_j^0 / \partial v < 0$ follow by the second order condition and the envelope theorem. Second, if b and v are small enough (i.e. if $b+v(2-\alpha) < 1$ holds) then the qualitative results summarized in Proposition 3 carry through unaltered. Otherwise, if the planning board cares about social welfare or the voters' well being enough, then it can be shown that jurisdictions that are more developed are less

15

regulated at equilibrium. The implications of our theory and those of two important alternative hypotheses are thus mutually exclusive.

3.5.   Cross-Metro Area theoretical predictions

As we shall see in the immediate sequel, our data is a cross-section of MSAs. Yet, in the US regulatory decisions are taken at the local level. Also, the theory so far has cross-sectional implications for jurisdictions. These implications tend to hold across MSAs, too. To fix ideas, assume that there is a number $M < J$ of MSAs in the economy indexed by $m \in \{1, ..., M\}$; the set of MSAs is a partition of $\Im$. In other words, each MSA is comprised of at least one jurisdiction and each jurisdiction belongs to exactly one MSA; we use $\Im_m$ to denote the subset of jurisdictions that belong to MSA m. Consider an arbitrary MSA m; then we can define any average variable pertaining to MSA m as $x_m \equiv |\Im_m|^{-1} \sum_{j \in \Im_m} x_j$, $x_j \in \{\omega_j, H_j, t_j\}$.

The relationships we want to test below – between amenities and land development, on the one hand, and land development and local regulation, on the other – are both monotonic; so they also tend to hold across MSAs. Yet, the relationships between H_j^0 and ω_j, on the one hand, and between t_j^0 and H_j^0, on the other, are non-linear, so aggregating over MSAs and using ω_m, H_m^0 and t_m^0 instead yields some measurement error. There are two ways to deal with this issue. Since we use linear regressions in the empirical section, the first solution is to do nothing more about it: such regressions can be seen as estimating a linear approximation of the model. The second way is to instrument for t_j^0 and H_j^0, which we also do.

4. Empirical Analysis

The main purpose of our empirical analysis is to explore the *causal* effect of physical residential development on regulatory restrictiveness at the MSA-level. It follows from the spatial equilibrium of our theory (7) and from the very purpose of regulation that residential development is endogenous to the regulatory environment. Therefore, we need an exogenous source of variation of urban development in order to identify its effect on regulation. Our theory readily suggests two sets of instruments: natural amenities and topography: desirable locations (those with a high a_m) and locations that contain a lot of plains (and hence have a low average conversion cost m_m) are more developed at equilibrium by (7). In our main empirical analysis we use land use data from 1992 to explore the causal effect of the share developed residential land on regulatory restrictiveness around 2005.

16

4.1. Description of data

Our data is derived from various sources and geographical levels of aggregation. We match all data to the MSA level using GIS. Table 1 provides summary statistics for all variables. The top panel pertains to our main sample period (turn of 21st century); the bottom panel reports the variables that belong to the earlier sample (around 1980).

The more recent of our two land use datasets, the National Land Cover Data 1992 (NLCD 92), is derived from satellite images. The earlier dataset, the Land Use and Land Cover GIRAS Spatial Data, comes from aerial photos taken around 1976. Both datasets cover the surface areas of all MSAs in our sample.[18] There is a considerable difference in map resolution and in land use definitions between 1976 and 1992, making direct comparisons of the data difficult. The two datasets are described in more detail in Burchfield *et al*. (2006).

We define the share developed land in an MSA (henceforth SDL) as

$$SDL = \frac{developed\ residential\ land\ area}{developable\ residential\ land\ area}, \tag{15}$$

where the 'developable residential land area' is the total land area minus the surface area that is covered by industrial land or 'non-developable' land uses (i.e., soil that does not sufficiently support permanent structures and/or is extremely costly to develop).[19] *SDL* is our proxy for H_j in the model and captures the political influence of owners of developed land relative to the influence of owners of undeveloped land.

The homeownership rates (*HOR* henceforth) for 1980 and 1990 are extracted from the Neighborhood Community Database (NCDB). We compute population densities in the developed residential area (*POPD*) using NCDB and NLCD data. (*POPD* and *HOR* respectively control for the alternative welfare economics and homevoter hypotheses.)

We use the two regulatory indices as the counterpart in the data to the regulatory tax t_m in the model. Each of these indices is derived from a different source and pertains to a different time

---

[18] A special case is the MSA of Washington, DC. While we have data for the surface area of the MSA outside the District of Columbia, we do not have any information for the District itself. Hence we imputed SDL by assuming that land uses within the District are similar to that at the boundaries. Since the District covers only about 1 percent of the MSA's surface area, this adjustment increases the SDL measure for the MSA by only about half a percentage point. None of our results changes notably if we assume that the District is either not at all or fully developed nor if we drop the observation altogether.

[19] These land uses include barren, water, ice, wetlands, and shrubland (1992 classifications) and 'undefined', barren, water, ice, wetlands, shrub/brush land, dry salt flats, beaches, sandy areas, bare exposed rock, strip mines, all categories of tundra except herbaceous tundra (1976 classifications). We experimented classifying *all* tundra as 'non-developable' or 'developable'. Results are virtually unchanged in all the specifications to follow.

period. *WRLURI* is a measure of differences in the local land use regulatory climate across more than 2600 communities across the US based on a 2005 survey and a separate study of state executive, legislative, and court activities. It is arguably the most comprehensive survey to date. See Gyourko *et al.* (2008) for details on the compilation. Saiz (2010) reports *WRLURI* values for 95 MSAs (our sample consists of 93 MSAs; we loose two observations for lack of data on 1880 population density). A *WRLURI* value of 1 implies that the measure is one standard deviation above the national mean. The *SAKS* measure was created by Saks (2008) as a 'comprehensive index of housing supply regulation' by using the simple average of six independent surveys conducted during the second half of the 1970s and the 1980s (see Saks 2008 for details). Saks reports regulatory index values for 83 MSAs. We loose two observations for lack of land use data and for lack of information on historic density.[20] Similar to the *WRLURI*, the *SAKS* index is scaled to have a mean of 0 and a standard deviation of 1.

There is considerable variation in the degree of land use regulation across US MSAs. Gyourko *et al.* (2008) suggest that there is more variation across than within MSAs. Other empirically motivated reasons also lead us to choose to run our regressions at the MSA level. In many MSAs only few municipalities responded to the surveys that are the foundation of the *WRLURI* and *SAKS* measures and many potentially important controls are available at the MSA- or state-level only. See also Gyourko *et al.* (2008) and Saiz (2010) on the merits of using MSA aggregates in this context. Our decision to use *aggregate* indices – rather than various measures of different types of land use regulation – allows us to capture the overall regulatory environment, while avoiding the loss of statistical clarity associated with trying to look at the effects of various types of regulations simultaneously (Glaeser and Ward 2009).

Our amenity measures and the region dummies are derived from the Environmental Systems Research Institute's (ESRI) Census 2000 MSA-level shape file. The sources for our other (excluded) instruments and controls are listed in the note to Table 1.

4.2.   Baseline empirical specification and results using OLS

Our objective in this section is to test the predictions of our model as directly as possible. The key prediction, stated in Proposition 3 (ii), follows from (12): places that are more developed are more regulated, i.e. $\partial t_m^0 / \partial H_m^0 > 0$. The homevoter hypothesis argues that places with a higher

---

[20] Glaeser *et al.* (2005a) estimate a regulatory tax for 21 MSAs using 1998 data. The interested reader may find this regulatory tax, *WRLURI*, *SAKS* and *SDL* values for these cities in Table U1 (not for publication). We also report pair-wise correlations and rank correlations of these variables as well as of the homeownership rates.

homeownership rate should be more regulated. The welfare economics view suggests that regulation corrects for market failures in the urban economy (e.g. 'externality zoning'). We use population density as a proxy for the intensity of these market failures. The motivation for this is that all urban economic theories predict that externalities that are conductive to agglomeration economies and urban costs are sensitive to distance: denser places generate more non-market interactions and pecuniary externalities, both conductive to urban growth (e.g. knowledge spillovers, labor market matching) and to urban costs (e.g. noise). These complementary explanations can be nested in the following model:

$$ WRLURI_m = \beta_0 + \beta_1 \left(\boldsymbol{SDL}_m \right) + \beta_2 \left(\boldsymbol{HOR}_m \right) + \beta_3 \left(POPD_m \right) + \beta_4 \left(controls_m \right) + \varepsilon_m, \quad (16) $$

where *WRLURI* is our measure for the restrictiveness of regulation and ε_m is the error term with the standard assumed properties. The priors are $\beta_1 > 0$ by the 'influential landowner' hypothesis, $\beta_2 > 0$ by Fischel's 'homevoter' hypothesis and $\beta_3 > 0$ by the welfare economics hypothesis. The variables in bold are potentially endogenously determined. Putting this issue aside, we start by running (16) by OLS. The controls include *share democratic votes*, namely, the state share of votes that went for the Democratic candidate in the 1988 and the 1992 presidential elections (allowing for the fact that regulatory restrictiveness may be driven by political ideology), *average household wage* (to control for the possibility that the findings are driven by income sorting), and *regional dummies* (to capture all other region-specific unobservable characteristics). The estimation results are reported in column (1) of Table 2.[21] The adjusted R^2 of 0.377 is reasonably high. Among the coefficients of interest, only β_1 has the expected sign and is statistically significant. This preliminary finding is encouraging for our influential landowner hypothesis. Turning attention to the controls, we see that MSAs in Democrat-leaning states are more regulated. Our interpretation is that liberal voters (in North American parlance) are ideologically more sympathetic to regulation than conservative voters.[22] This result is robust to adding an interaction term between *share democratic votes* and *average income*; the coefficient is insignificant, suggesting that blue collar and white collar Democrats do not hold significantly different views on regulations.[23] *Region dummies* reveal that broad geographic patterns emerge,

---

[21] Throughout the paper the standard errors are clustered by state because the *share of democratic votes* is state-specific.

[22] To ensure that our results may not be spurious, we include one explanatory variable at a time. We find that the OLS coefficient of SDL varies from 1.34 when only *SDL* and *HOR* are included to around 2 when we include the *share Democrat votes* variable. Thus, in addition to being an important explanation on its own, ideology is helpful in identifying the role of *SDL*. See Table U2 (not intended for publication) for details.

[23] See Table U7 (not intended for publication), Columns (1) to (3) for details.

with the West being the most regulated region and the Midwest (the omitted category) the least regulated.

4.3. Identification strategy and results for IV-specifications

One important caveat with the OLS estimates of (16) reported in Table 2 (column 1) is that our key explanatory variable is likely endogenously determined, causing the estimate to be biased. Among possible sources of endogeneity, it directly follows from our theory that regulation works as an impediment to development by (7). This implies that the estimation of β_1 in (16) is biased downwards. We address this issue by instrumenting for *SDL*.

We rely on the model to find credible sources of exogenous variations in *SDL* that are not directly correlated with our regulatory measure *WRLURI*. Our identifying assumption for the *SDL* variable is that places endowed with desirable amenities and located on plains are developed earlier, attract more residents over time and, as a result, are more developed in our cross-section of MSAs, but that these characteristics are not directly related to regulatory restrictiveness. These predictions directly follow from Proposition 3 (i). Its first component is a demand factor: ceteris paribus, people prefer to live in nice places. We thus use a dummy variable that equals one if the MSA *has a major border with a coastline* and *average temperatures in January* as instruments for *SDL*. January temperatures should not have a direct and systematic influence on a broad index of residential land use regulations. However, the reader may worry that valuable ocean coasts require protection in the form of regulation. In practice, *three* properties of the *WRLURI* measure suggest that *border with coast* is a valid instrument. First, the *WRLURI* measure does not include attitudes towards regulation of coastal areas. Second, the majority of municipalities responding to the survey do not have access to the coast; this is true even for municipalities that belong to MSAs with access to the coast. Finally, federal regulations that may protect the coast are excluded from *WRLURI* by construction. We conduct additional robustness checks in Sections 5.3 and 5.4.

The second component of Proposition 3(ii) is a supply factor: it is simpler and cheaper to convert open land into developed land in plains. Hence, we use *share of plains* as an instrument for *SDL*. The way we define *SDL* in (15) is crucial for this to be part of a valid identification strategy. Indeed, some regulations may be designed purposely to protect some local amenities; Saiz (2010) shows how land use regulations correlate with the fraction of undeveloped land in an MSA. These plots, on which it is not practically feasible to build, are excluded from both the numerator and the denominator of (15).

We also use *historical population density* from 1880 as an additional instrument for *SDL*. This is consistent with a dynamic interpretation of our model: desirable locations and plains attracted people early and were developed first, *before land use regulations became part of the urban political life*. This variable captures all the unobserved and time-invariant amenity and cost factors not already included in our set of instruments that lead people to settle in a specific place. It also captures historic amenity and cost factors that were important a long time ago and which started a dynamic development process of cities. They may no longer be important today and yet remain relevant because of inertia, durable housing, or the generation of agglomeration forces.

These considerations lead us to run the following first stage regression by OLS:

$$SDL_j = \alpha_0 + \alpha_1 \left(coast_j \right) + \alpha_2 \left(temperature_j \right) + \alpha_3 \left(share\ plains_j \right)$$
$$+\alpha_4 \left(historical\ density_j \right) + \alpha_5 \left(controls_j \right) + \xi_j, \tag{17}$$

where ξ_j is the error term. Our priors are $\alpha_1, \alpha_2, \alpha_3, \alpha_4 > 0$. The results are reported in column (2) of Table 2. All coefficients have the expected sign and are significant beyond the five percent level with the exception of α_3: *share of plains* is only weakly correlated with *SDL*. Nevertheless, the four instruments are jointly significant. Households obviously value access to the seafront and mild winter temperatures. The quantitative effects are strong:[24] granting a *border with coast* to a hitherto landlocked MSA increases its share of developed land by 65.2% (+8.1 percentage points); an extra standard deviation in *average January temperature* and *historical density* are respectively associated with a 52.7% (+6.5 percentage points) and a 44.3% (+5.5 percentage points) increase in *SDL*. As to our set of controls, MSAs in the South have the lowest share of developed land and MSAs that command a high household average wage are more developed, consistent with the logic of our model of household location. These effects are statistically significant at the one percent level. The adjusted R^2 is high with 0.59. The second stage regression results of (16) with *SDL* being treated as the unique endogenous variable are reported in column (5) of Table 2. The TSLS coefficient of *SDL* is positive and significant at the five percent level. The point estimate is larger than the OLS coefficient of *SDL* and suggests the presence of a downward bias (we provide more direct evidence about this feedback/reverse-causation mechanism in Section 6). These findings provide both direct and indirect evidence consistent with our theory. The coefficients on the homeownership rate (*HOR*) and contemporaneous population density (*POPD*) remain insignificant. The findings that Democrat-

---

[24] We report quantitative effects for all our main specifications in Table U8 (not intended for publication).

leaning MSAs and those in the West prefer above average levels of regulation are very similar for our OLS and TSLS specifications.

We conclude from these findings that the effect of *SDL* on regulation is well identified. The effect is also quantitatively meaningful. To fix ideas, compare Kansas City to San Francisco. The former has no access to coast, average January temperatures are 28.5°F, its share of developable land classified as plains is 63% and its historical population density is 52.3 people per km^2. San Francisco has a border with the Pacific Ocean, January temperatures average 48.2°F, it has no county that is classified as a plain and its population density in 1880 was 239.6 people per km^2. The implied difference in *SDL* is 16.3 percentage points (a full 1.6 standard deviations). This, in turn, implies a 1.05 standard deviation difference of *WRLURI* between the two MSAs. Kansas City is the 9th *least* regulated MSA in our sample (i.e. the 74th most regulated MSA). Granting it with San Francisco's amenities and topography alone hypothetically makes it the 41st most regulated MSA (SF is the 16th most regulated MSA).

The estimation of β_2 in (16) may also be biased if there are omitted variables that are correlated with *HOR* or if land use regulations systematically influence the incentive to own one's home. We use the MSA's share of households that consist of married couples without children as a source of exogenous variation of *HOR* in order to improve the identification of its effect on *WRLURI*. Married couples without children tend to have higher and more stable household incomes and are able to accumulate greater wealth over time compared to married couples with children. This makes them more likely to overcome liquidity and down-payment constraints and thus eases attaining homeownership. Moreover, married couples tend to be in more stable relationships compared to their unmarried counterparts, implying a longer expected duration in their property and, consequently, greater incentives to own rather than rent. By contrast, we do not expect the share of households that consists of married couples without children to help us identify the *SDL*. Our empirical results reported in column (3) of Table 2 are consistent with this prior; by contrast, the patterns of column (2) are unchanged or even reinforced. Column (4) reports OLS estimates of the effect of the *share households with married couples and no children* and the various controls on the homeownership rate. As predicted, the former is positive and highly statistically significant at the 1 percent level. *Historical population density* also helps us to identify the *HOR*; this finding makes sense because denser places have taller buildings and renting (rather than owning) is more efficient in multi-unit buildings. The adjusted R^2 of 0.658 is high. The results contained in columns (2) to (4) thus establish that our proposed instruments for *SDL* and *HOR* fulfill the necessary condition for being valid instruments.

The estimation of (16) with both *SDL* and *HOR* instrumented for are reported in Table 2, columns (6) to (8) (we additionally instrument for *POPD* in Section 5.2). Independent of whether we use the TSLS estimator (column 6), the Limited Information Maximum Likelihood (LIML) estimator (column 7) or the Jackknife (JIVE) estimator (column 8) we find that the coefficient of *SDL* is positive, statistically significant and larger than the OLS coefficient of column (1). These results confirm the presence of a downward bias in the OLS specification and reinforce our influential landowner hypothesis. We run the model using LIML because it is approximately median unbiased for over-identified models (we have five instruments and two endogenous explanatory variables) and produces a smaller bias than TSLS in finite samples. Since its asymptotic properties are the same as those of the TSLS estimator, we hope to find similar coefficients in the TSLS and LIML regressions as a rule of thumb (Angrist and Pischke 2009). We report the results of the regression using the JIVE estimator for further reference only, so we postpone further discussion of this estimation technique. The stability of the magnitude of the estimated β_1 across columns (6) to (8) increases further our confidence in the robustness of our findings and strengthens our IV strategy. We also carry out the usual battery of tests that assess the validity of the instrumental variables, including over-identification tests as well as Hansen-J statistics and Kleibergen-Paap rk LM statistics, and none of these tests indicates a problem at the usual confidence levels. Therefore, we do not report these results in order to save space. The last line of Table 2 reports Kleibergen-Paap rk Wald F-statistics, a test for weak instruments in the presence of robust (clustered) standard errors. The test statistic in column (5) indicates with 95 percent confidence that the maximum TSLS size is just about 15%, implying that our instruments taken together are reasonably strong (Stock and Yogo 2005; Kleibergen and Paap 2006). The statistic in column (6) is much lower, raising concerns that our instruments might be weak in this case; this provides one additional motivation for replicating the analysis using LIML (the Kleibergen-Paap rk Wald F-statistic is the same but the critical values are lower than for TSLS). The result in this case is extremely strong: the reported statistic in column (7) is well above the critical value for a maximum LIML size of 10%. This vindicates our identification strategy.

To summarize, the robust results so far are strongly supportive of various aspects of the influential landowner hypothesis. The theoretical predictions of Proposition 3(i) and (ii) are vindicated; the effect of *regulation* on *SDL* introduces a downward bias.

5. Further specifications and alternative dataset

We explore the sensitivity of our results to the set of instruments we include (Table 3), the set of endogenous variables we instrument for and the estimator we use (Table 4), the proxy we use for

the relative influence of owners of developed land (Table 5), the inclusion of controls that capture preferences for protection of open space (Table 6) and the dataset and time period we use (Table 7).

5.1. Instruments

In our *baseline specification* we instrument for *SDL* and *HOR* using five instruments. We replicate our baseline results in column (1) of Table 3 for convenience. The remaining columns in Table 3 report results for reduced sets of instruments. Panel A reports the first stage estimated coefficients of our instruments (the dependant variable is *SDL*) and Panels B and C respectively report the second stage TSLS and LIML coefficients of the share developed land, the homeownership rate and the population density in 1990. In columns (2) to (5), we replicate these estimations dropping one instrument at a time. In columns (6) to (11), we drop two instruments at a time. Inspection of the coefficients in Panel A reveals that *border with coast*, *average temperatures in January* and *historical density* are particularly helpful in our quest to identify the effect of *SDL* on *WRLURI*: they consistently have the expected sign, their magnitude is stable, and they are statistically significant at the one percent level (except *average temperatures in January* that sometimes ventures in the five percent zone). The effect of *share plains* on *SDL* is also positive and statistically significant throughout, though 'only' at the five or ten percent level. Finally, the *share of married couples without children* (our excluded instrument for *HOR*) is uncorrelated to *SDL* throughout, as in our baseline specification.

Turning to Panels B and C of Table 3, the striking result is that the effect of *SDL* on *WRLURI* is positive and statistically significant in nine cases out of ten: dropping both *border with coast* and *share plains* creates the only combination of instruments that yields a positive coefficient on *SDL* that marginally misses significance at the 10 percent level. Dropping either alone, however, or in combination of any other instrument, *does* identify the effect of *SDL* on *WRLURI*. The coefficients on *HOR* are stable but not statistically larger than zero (with one borderline exception). The coefficients on *POPD* remain statistically insignificant throughout. Ideology and regional dummies (not reported) remain stable and statistically significant.

The Kleibergen-Paap rk Wald F-statistics are in line with those of Table 2: the TSLS statistics fluctuate around the critical values of maximum TSLS sizes of 15% to 20%; the LIML statistics are all well above the critical value for a maximum LIML size of 10%.

5.2.   Endogenous Population Density

In Table 4, we endogenize the *POPD* variable in addition to *SDL* and *HOR*. Various types of land use controls – including minimum lot size restrictions – differentially affect the population density, suggesting reversed causation and biased estimates. We expect two of our excluded instruments to be useful for identifying *POPD*. The first of these instruments is the *share of plains* in an MSA. The identifying assumption is that sprawl is easier in particularly flat areas, where it is particularly easy to build, leading us to expect a negative coefficient for *share of plains* when the dependant variable is *POPD* (in contrast to the *SDL* variable). The second instrument is historical MSA-level population density from 1880. We expect the MSAs that were densely populated in the 19th century (prior to the evolution of land use regulation in the United States) to have a densely populated developable area today. Column (3) shows that *historical population density* and *share of plains* both have the expected sign and are statistically significant. Columns (1) and (2) of Table 4 report results for the first stages of the variables *SDL* and *HOR*, respectively. The results are similar to the corresponding ones in Table 2 (columns 3 and 4).

The results of the second-stage using alternative estimators (TSLS, LIML, and JIVE), reported in columns (4) to (6), are equally supportive. Instrumenting simultaneously for *SDL*, *HOR*, and *POPD* systematically increases the point estimate of the second stage coefficient of *SDL* relative to the specifications with one or two instrumented variables in Table 2, columns (6) to (8). These coefficients also remain statistically significant at the same confidence levels as their Table 2 counterparts. The quantitative effect of *SDL* on *WRLURI* is also enlarged: a one standard deviation increase in *SDL* raises *WRLURI* by more than one third of a standard deviation; this is equivalent to a boost in the regulatory rate league table from the median (rank 47) to the top third. By contrast, the second stage coefficients of *HOR* and *POPD* remain statistically insignificant in all three specifications reported in columns (4) to (6).

While we can calculate and report Kleibergen-Paap rk Wald F-statistics for our TSLS and LIML specifications to assess whether our instruments are jointly 'weak', critical values are not available from Stock and Yogo (2005) for specifications with more than two endogenous variables. To correct for the *possible* presence of weak instruments, we therefore also re-estimate (16) using a JIVE estimator (Angrist, Imbens and Krueger 1999). JIVE gets round the correlation between stage-one and stage-two errors by predicting the value of *SDL* of MSA *j* by running the first stage for *j* on all MSAs but *j* and repeating the procedure for all 93 cities in the sample (thus, the procedure has in a sense as many first stages as observations). As a result, this IV estimator

bias is smaller than the TSLS bias but the standard errors are larger. The regression results of column (6) are consistent with these priors. The estimated coefficient for *SDL* using JIVE is of comparable magnitude to those of TSLS and LIML. This suggests that *even if* the instruments used to identify the endogenous variables were jointly weak the resulting bias would be small.

The estimated coefficients on political ideology (*share democratic votes*) and the region dummies are stable and remain statistically significant.

5.3.   MSA boundaries and representative places

To test the robustness of our results to the MSA definition, we redefine *SDL* so as to include only the land cover within a 20km radius from the centre of each MSA. It turns out that 'more developed' MSAs are more developed at any radius from the center than 'less developed' MSAs (see also Burchfield *et al.* 2006 on this), which leads us to expect our main results to be robust to this change. We also redefine *SDL* in various ways that include industrial land or exclude parks, or both. Finally, to immunize our results to the role of outlier places, we attribute to the MSA the *SDL* of its average or median place. We report the results of various combinations of these robustness checks in Table 5, columns (1) to (9). Column (10) replicates the whole analysis using the *aggregate property value per m^2 of developable land* as the proxy variable for the relative influence of owners of developed land relative to that of owners of undeveloped land. This is the alternative, indirect measure that we use in Figure 1 (Panel c). Panels A, B and C respectively report the first stage, second stage TSLS and second stage LIML results. We instrument for both *SDL* and *HOR* as in our baseline specification throughout; the results ought thus to be gauged against those of Table 2, columns (6) and (7).

The results are again strongly in line with our baseline specification. The *major access to coast* and *historical population density* variables are positive and highly statistically significant in all specifications of Panel A. Aggregate property values are strongly and positively correlated with *border with coast* and *average January temperatures*, which is consistent with the finding that desirable amenities are at least partly capitalized into land prices (Gyourko *et al.* 2006). The second-stage results reported in Table 5 are equally bold. *In all 18 reported specifications* – using quite different definitions for *SDL* and two different estimation techniques – *we find that SDL has a positive and statistically highly significant causal impact on WRLURI*. The coefficient on *aggregate property value* in column (10) is also statistically positive at the ten percent level in both panels. Reassuringly, the coefficients on *all* variables are quite stable across specifications. We also replicate all the regressions in Table 5 endogenizing simultaneously for *SDL*, *HOR* and *POPD*. Our results come out even stronger: the point estimate of the coefficient of *SDL* rises in

all twenty cases and it is statistically significant at the one percent level throughout Panel B and never below the five percent level throughout Panel C. By contrast, *HOR* is statistically significant in only three specifications out of ten and *POPD* is significant in none of them.[25]

5.4. Protection of open space and regulation

Local residents may turn against future development and opt for more restrictive land use regulations if open space becomes scarce in absolute terms (i.e. considering open space as a local public good) or in relative terms (i.e. allowing open space to be subject to crowding). To test that the effect of *SDL* on regulation is not driven by preferences for open space or conservationist motives, we re-estimate our baseline specifications with the total amount of open land (independent of whether the land is developable or not) in an MSA or the amount of open land in an MSA per capita as additional controls. Table 6 provides again strong support for our influential landowner hypothesis; adding the open space controls slightly increases the statistical significance and estimated coefficients on *SDL*. The evidence with respect to the open space hypothesis is mixed. Whereas in some of the TSLS-estimates the controls have the expected sign and are statistically significant, LIML and JIVE-estimates yield statistically insignificant results. These findings are in line with those reported in Table 5, where we find that adding or excluding public parkland from the denominator of the *SDL*-measure does not alter our key findings in a meaningful way.

5.5. The 1970/1980 sample

As an 'out-of-sample' robustness check we apply our two-stage methodology to a different dataset and time period. The dependent variable of interest is the *SAKS* index of residential land use regulations, pertaining to the late 1970s/early 1980s. Our measure for *SDL* is derived from aerial photos taken in the mid 1970s.

We regress (16) with *SAKS* replacing *WRLURI*. Table 7 (Panel A) reports the results. Turn to columns (1) to (3) for the first stage regressions. Interestingly, at that time land development seemed to be well explained by *average January temperature* and *historical population density*, whereas *border with coast* and *share plains* play a lesser role than in the 1992 data. The second stage regressions, reported in columns (4) to (7), provide again strong support for our influential landowner hypothesis. The estimated coefficient of *SDL* is stable across specifications (one or two endogenous variables) and estimators (TSLS and LIML). It is also statistically significant at

---

[25] For details, see Table U3 (not intended for publication).

the one percent level and quantitatively strong: one extra standard deviation in *SDL* raises *SAKS* by over two thirds of a standard deviation; this is equivalent to a boost in the regulatory rate league table from the median (rank 41) to the top quarter. This earlier data also provide no support for the homevoter hypothesis and the implications of the welfare economics view at the MSA level. Political ideology, as measured by the *share democratic votes* in the two preceding presidential elections, is seemingly unrelated to regulatory restrictiveness in the late 1970s and 1980s. Finally, in line with our first stage findings, the Kleibergen-Paap statistics suggest that our instruments are weaker than in the baseline case. We thus re-run our regressions with the reduced set of statistically significant excluded instruments: *average January temperature* and *historical population density*.[26] The estimated coefficients are stable and those on *SDL* remain significant at the one percent level; the Kleibergen-Paap statistic also increases, as expected. In short, with the exception of the role of ideology, the patterns that we have uncovered for the 1990/2000 data were already present in the 1970/1980 data.

5.6.   Additional robustness checks

We performed a variety of additional robustness checks. Since they reinforce our main results with only minor qualifications, we report them only briefly; also, we include the associated tables (not intended for publication) in this version of the paper for the sake of completeness.

The contemporaneous sample of cities has 93 MSAs and the one pertaining to the late 1970s has 81. Yet, only 63 MSAs are included in both samples. In Table U5, we replicate the simple OLS as well as the first stage and second stage TSLS and LIML results for both contemporaneous regulation (as measured by *WRLURI*) in Panel A and older regulation (as measured by *SAKS*) in Panel B. The striking result is that the key coefficients of the influential landowner hypothesis remain precisely estimated and of the correct sign, despite the sample size being quite small. This caveat bites only in producing relatively weak Kleibergen-Paap statistics.

NYC and many Californian cities have rent controls. As these controls might affect the nature of the political economy game, we replicate our analysis excluding either or both sets of cities. Table U6 reports the first stage results (Panel A) and the OLS and second stage TSLS and LIML results (Panel B). It is readily verified that the influential landowner hypothesis finds strong and stable support throughout. Interestingly, the homevoter hypothesis now finds support in the TSLS and LIML specifications as well.

---

[26] See Panel A of Table U4 (not intended for publication).

Finally, we drop the explanatory variable *share of democratic votes* and re-run our central specification one more time. The results are reported in Table U7, columns (4) to (7). The estimated coefficients on *SDL* drop somewhat but remain statistically significant. Political ideology clearly reinforces the identification of the influential landowner hypothesis in addition to playing an important role on its own in explaining land use regulation patterns.

6. Land use regulations and the supply of housing

The workings of the economic mechanism central to our influential landowner hypothesis rests partly on the assumption that land use regulations increase the cost and reduce the quantity of further developments as per (7). To check whether this is a feature of the regulatory index in our data, we run the following with OLS:

$$g_{j,t}^{H} = \gamma_0 + \gamma_1(regulation_{j,t}^{e}) + \gamma_2(g_{j,t-1}^{H}) + \zeta_{j,t} \tag{18}$$

where $g_{j,t}^{H}$ is the growth rate of the housing stock in MSA j between time period t and $t+1$ (in number of housing units), $g_{j,t-1}^{H}$ is the equivalent growth rate between time period t-1 and t, γ_0 is the common trend, $regulation_{j,t}^{e}$ is the estimated level of regulatory restrictiveness in MSA j at time t, and $\zeta_{j,t}$ is the error term. We include $g_{j,t-1}^{H}$ to allow for persistence in local housing markets. Our theoretical prior leads us to expect $\gamma_1 < 0$ by a dynamic version of (7). To reduce the importance of high frequency shocks, and for data availability reasons, we compute the growth rates over 10-year time periods. By this token, we will have to wait until 2020 to assess the effect of *WRLURI* on the ten-year growth of the housing stock. However, we can readily assess the effect of the (estimated) *SAKS* measure.

We back out the fitted values for $regulation_{j,t}^{e} = SAKS_{j,t}^{e}$ from the estimations of Table 7, Panel A. We report the regression results for (18) in Panel B. Because the regulatory index measures are estimated values, we report bootstrapped (and robust) standard errors using 1,000 replications. They are also in line with our priors: in all four specifications the predicted regulatory tax has a negative and statistically highly significant effect on the growth rate in housing supply in the 1990s.[27] An increase of the predicted *SAKS* index by one standard deviation reduces the growth rate of housing supply by 2.8 percentage points. This is equivalent to a drop in the growth rate league table from rank 41 to 54. Overall, our results in Table 7 imply

---

[27] The qualitative results go unaltered if we use only *average January temperature* and *historical density* in our set of excluded instruments. See Panel B in Table U4 (not intended for publication).

that while more desirable MSAs grew more quickly in the past when little land was developed and regulation was lax, their growth rate has later slowed down significantly compared with less desirable MSAs. We attribute this effect to tighter land use controls. To fit the spirit of our linear model more closely, we also run (18) in first differences (results not reported), with no effect on the qualitative results that we report here.

7. Concluding remarks

Land use regulations vary tremendously in shape and scope across space and have become more widespread and stringent over time. They can, in principle, raise welfare by correcting market failures. Recent evidence, however, casts doubt on this proposition: land use regulations impose – via increasing housing costs – an enormous gross cost on households that is unlikely to be matched by welfare gains arising from correcting market failure. Understanding the effects and causes of these regulations is thus of primary economic policy importance. Yet, perhaps because a large part of these costs are indirect and the underlying political-economy processes (such as lobbying) are difficult to observe empirically, this area of research remains relatively under-explored.

Our study contributes to the understanding of political economics considerations that shape land use restrictions. We focus exclusively on residential land use by the nature of the regulatory data available. In practice, zoning also separates incompatible land uses and the business districts from residential areas. With this caveat in mind, our results point to land-based-interests explanations and suggest that the tightness of residential land use regulations goes beyond welfare economics considerations. Thus, the outcome is suboptimal. Specifically, regulation in highly desirable and highly developed places like New York City and San Francisco may be grossly over-restrictive while less attractive metro areas may be too little regulated. As a result, too few people appear to be living in the desirable and productive cities and too many people may be living in less desirable places relative to the social optimum.

References

Alonso, W. 1964. *Location and Land Use.* Cambridge, MA: Harvard University Press.

Angrist, J.D., G.W. Imbens, A.B. Krueger. 1999. 'Jackknife instrumental variables estimation.' *Journal of Applied Econometrics* 14(1): 57-67.

Angrist, J.D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics.* Princeton, NJ: Princeton University Press.

Anderson, S. P., A. De Palma, and J. F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. Cambridge, MA: MIT Press.

Austen-Smith, D. 1995. 'Campaign contributions and access.' *American Political Science Review*, 89(3): 566-581.

Baldwin, R.E., and F. Robert-Nicoud. 2007. 'Entry and asymmetric lobbying: Why governments pick losers.' *Journal of the European Economic Association*, 5(5): 1064-1093.

Bernheim, Douglas B., and Michael D. Whinston. 1986. 'Menu auctions, resource allocation, and economic influence.' *Quarterly Journal of Economics,* 101(1): 1-31.

Brueckner, J. K. 1987. 'The structure of urban equilibria: A unified treatment of the Muth-Mills model.' In *Handbook of Regional and Urban Economics*, vol. 2, ed. E.S. Mills, 821-846. Amsterdam: Elsevier.

Brueckner, J. K. 1995. 'Strategic control of growth in a system of cities.' *Journal of Public Economics*, 57(3): 393-424.

Burchfield, M., H.G. Overman, D. Puga, and M.A. Turner. 2006. 'Causes of sprawl: A portrait from space.' *Quarterly Journal of Economics*, 12(1): 587-633.

Cheshire, P.C., and C.A.L. Hilber. 2008. 'Office space supply restrictions in Britain: The political economy of market revenge.' *Economic Journal*, 118(529): F185-F221.

Combes, P.P., G. Duranton, and L. Gobillon. 2010. 'Estimating agglomeration economies with history, geology and worker effects' in E.L. Glaeser (ed) *Agglomeration economics*. Chicago: University of Chicago Press, and NBER.

Dehring, C., C. Depken, and M. Ward. 2008. 'A direct test of the homevoter hypothesis', *Journal of Urban Economics*, 64(1): 155-170.

Dixit, A., G. Grossman, and E. Helpman. 1997. 'Common agency and coordination: General theory and application to government policy making.' *Journal of Political Economy*, 105(4): 752-769.

Epple, D., T. Romer, and R. Filimon. 1988. 'Community Development with Endogenous Land Use Controls.' *Journal of Public Economics*, 35: 133-162.

Evenson, B., and W.C. Wheaton. 2003. 'Local variations in land use regulations.' *Brookings Wharton Papers on Urban Affairs*: 221-260.

Fischel, W. A. 2001. *The Homevoter Hypothesis. How Home Values Influence Local Government Taxation, School Finance, and Land-Use Policies*. Cambridge (MA): Harvard University Press.

Fischel, W.A. 2004. 'An economic history of zoning and a cure for its exclusionary effects.' *Urban Studies*, 41(2): 317-340.

Fujita, M., and J.-F. Thisse. 2002. *Economics of Agglomeration. Cities, Industrial Location, and Regional Growth*. Cambridge: Cambridge University Press.

Glaeser, E.L. 2008. *Cities, Agglomeration and Spatial Equilibrium.* Oxford and New York: Oxford University Press.

Glaeser, E.L., J. Kolko, and A. Saiz. 2001. 'Consumer city.' *Journal of Economic Geography*, 1(1): 27-50.

Glaeser, E. L., J. Gyourko, and R. E. Saks. 2005a. 'Why is Manhattan so Expensive? Regulation and the Rise in House Prices.' *Journal of Law and Economics*, 48(2): 331-369.

Glaeser, E. L., J. Gyourko, and R. E. Saks. 2005b. 'Why Have Housing Prices Gone Up?' *American Economic Review Papers and Proceedings*, 95(2): 329-333.

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

Goldberg, P., and G. Maggi. 1999. 'Protection for sale: An empirical investigation.' *American Economic Review*, 89(5): 1135-1155.

Gyourko, J., C. Mayer, and T. Sinai. 2006. 'Super Star Cities.' NBER Working Paper 12355.

Gyourko, J., A. Saiz, and A. Summers. 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.

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

Hilber, C.A.L., and C.J. Mayer. 2009. 'Why do households without children support local public schools? Linking house price capitalization to school spending.' *Journal of Urban Economics*, 65(1): 74-90.

Hilber, C.A.L., and F. Robert-Nicoud. 2006. 'Owners of developed land versus owners of undeveloped land: Why land use is more constrained in the Bay Area than in Pittsburgh. CEP Discussion Paper no. 760 (Centre for Economic Performance, LSE).

Kakutani, S. 1941. A generalization of Brouwer's fixed point theorem. *Duke Mathematical Journal* 8(3): 457-459.

Kleibergen, F. and R. Paap. 2006. 'Generalized reduced rank tests using the singular value decomposition.' *Journal of Econometrics*, 133: 97-126.

Krishna, P. 1998. 'Regionalism and multilateralism: A political economy approach.' *Quarterly Journal of Economics*, 113(1): 227-250.

Lohmann, S. 1995. 'Information, access, and contributions: A signaling model of lobbying.' *Public Choice*, 85: 267-284.

Manski, C. 1977. 'The structure of random utility models.' *Theory and Decision*, 8: 229-254.

Molotch, H. 1976. 'The city as a growth machine: Toward a political economy of place.' *American Journal of Sociology*, 82(2): 309-332.

Oates, W.E. 1969. 'The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the Tiebout hypothesis.' *Journal of Political Economy*, 77(6): 957-971.

Ortalo-Magné, F., and A. Prat. 2007. 'The political economy of housing supply: Homeowners, workers and voters.' Mimeo, University of Wisconsin-Madison and LSE.

Palmon, O., and B.A. Smith. 1998. 'New evidence on property tax capitalization.' *Journal of Political Economy*, 106(5): 1099-1111.

Persson, T., and G. Tabellini. 2000. *Political Economics: Explaining Economic Policy.* Cambridge MA: MIT Press.

Quigley, J. M., and S. Raphael, S. 2005. 'Regulation and the High Cost of Housing in California.' *American Economic Review Papers and Proceedings*, 95(2): 323-329.

Rappaport, J. 2007. 'Moving to nice weather.' *Regional Science and Urban Economics* 37(3): 375-398.

Roback, J. 1982. 'Wage, rents, and the quality of life.' *Journal of Political Economy*, 90(6): 1257-1278.

Rosen, S. 1979. 'Wage-based indexes of urban quality of life.' In *Current Issues in Urban Economics*, ed. P. Mieszkowski and M. Straszheim. Baltimore: John Hopkins University Press.

Rosenthal, S.S. and W. Strange. 2004. 'Evidence on the nature and sources of agglomeration economies', in J.V. Henderson and J.F. Thisse (eds) *Handbook of Regional and Urban Economics*, Volume 4. Amsterdam: North Holland.

Rossi-Hansberg, E. 2004. 'Optimal land use and zoning.' *Review of Economic Dynamics* 7, 69-106.

Rudel, T. K. 1989. *Situations and Strategies in American Land-Use Planning*. Cambridge, England: Cambridge University Press.

Saks, R. E. 2008. 'Job Creation and Housing Construction: Constraints on Metropolitan Area Employment Growth.' *Journal of Urban Economics*, 64(1): 178-195.

Saiz, A. 2010. 'The Geographic Determinants of Housing Supply.' *Quarterly Journal of Economics*, forthcoming.

Stock, J.H., and M. Yogo. 2005. 'Testing for weak instruments in linear IV regression.' In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. D.W.K. Andrews and J.H. Stock, 80-108. Cambridge: Cambridge University Press.

Thurstone, L. 1927. 'Psychological analysis.' *American Journal of Psychology*, 38: 369-389.

Summary Statistics and Regression Tables

Table 1
Summary statistics

Variable	N	Mean	Std. Dev.	Min	Max
Wharton regulatory index (**WRLURI**), 2005 [a]	93	0.117	0.702	-1.25	2.07
Homeownership rate, 1990 [b]	93	0.628	0.070	0.325	0.739
Developed residential land as % of developable non-industrial land ("share developed residential"), 1992 [c]	93	0.124	0.122	0.0119	0.761
Alternative measures for robustness checks					
Share developed (incl. industrial developments), 1992 [c]	93	0.154	0.134	0.0198	0.847
Share developed, 20km radius, 1992 [c]	93	0.363	0.208	0.0580	1
Share developed residential, 20km radius, 1992 [c]	93	0.309	0.204	0.0436	1
Share developed, excluding parks, 1992 [c]	93	0.149	0.139	0.0204	1
Share developed residential, excluding parks, 1992 [c]	93	0.121	0.132	0.0124	1
Share developed of average place, 1992 [c]	93	0.462	0.164	0.0976	1
Share developed residential of average place, 1992 [c]	93	0.419	0.168	0.0822	1
Share developed of median place, 1992 [c]	93	0.463	0.188	0.0947	1
Share developed residential of median place, 1992 [c]	93	0.414	0.190	0.0793	1
Aggregate property value per m^2 of developable land, 1990 [b]	93	17.9	35.4	1.3	237.5
Population density in developed residential area (per m^2), 1990 [d]	93	0.00264	0.00125	0.00116	0.0107
%Democratic votes in state, av. presidential elections 1988/92 [e]	93	48.8	4.9	34.4	58.8
Average household wage, 1990 [b]	93	30.0	5.2	17.0	46.5
Region = Midwest (omitted) [f]	93	0.215	0.413	0	1
Region = North East [f]	93	0.183	0.389	0	1
Region = South [f]	93	0.376	0.487	0	1
Region = West [f]	93	0.226	0.420	0	1
Metro area has major border with coast [f]	93	0.247	0.434	0	1
Average temperature in January, measured between 1941-1970 [g]	93	38.2	12.5	11.8	67.2
Share land in topography classification that consists of plains [g]	93	0.546	0.432	0	1
Population density in metro area (per m^2), 1880 [i], x 10^{-6}	93	125.5	490.3	0.1	4698.6
% Households with married couples and no children, 1990 [b]	93	0.291	0.028	0.236	0.427
Total open land in MSA in thousand km^2	93	9.11	12.89	0.0436	102.2
Total open land in MSA in km^2 per person	93	0.00915	0.0152	0.0000789	0.138
Saks-index of housing supply regulation (**SAKS**), late 1970s/80s [h]	81	0.00544	0.997	-2.399	2.211
Homeownership rate, 1980 [b]	81	0.636	0.0739	0.278	0.764
Developed residential land as % of developable land, 1976 [c]	81	0.118	0.104	0.0119	0.501
Population density in the developed area (per m^2), 1980 [d]	81	0.00231	0.00113	0.000254	0.00896
%Democratic votes in state, av. presidential elections 1972/76 [e]	81	43.8	3.1	31.6	51.9
Average household wage, 1980 [b]	81	16.0	2.4	9.2	22.0
Region = Midwest (omitted) [f]	81	0.210	0.410	0	1
Region = North East [f]	81	0.160	0.369	0	1
Region = South [f]	81	0.383	0.489	0	1
Region = West [f]	81	0.247	0.434	0	1
Metro area has major border with coast [f]	81	0.284	0.454	0	1
Average temperature in January, measured between 1941-1970 [g]	81	39.4	12.8	11.8	67.2
Share land in topography classification that consists of plains [g]	81	0.534	0.431	0	1
Population density in metro area (per km^2), 1880 [i]	81	75.3	110.0	0.0278	647.7
% Households with married couples and no children, 1990 [b]	81	0.299	0.0357	0.203	0.451
Percent change, housing units, 1990-2000 [b]	81	0.135	0.0874	0.0266	0.484
Percent change, housing units, 1980-1990 [b]	81	0.207	0.154	0.0134	0.637

Sources: [a] Saiz (2008); [b] US Census and Neighborhood Community Database (NCDB); [c] National Land Cover Data (NLCD) 1976 and 1992 from the U.S. Geological Survey; Missing map cells for 1976 were obtained from Diego Puga at http://diegopuga.org/data/sprawl/. Map data was unavailable for Santa Cruz, California and a mis-projected map for Erie, Pennsylvania necessitated the removal of fourteen affected census tracts; [d] Derived from NLCD and NCDB; [e] Dave Leip's Atlas of Presidential Elections; [f] Derived from ESRI's Census 2000 MSA-level shape file; [g] Natural Amenity Scale Data from the Economic Research Service, United States Department of Agriculture; [h] Saks (2008); [i] Interuniversity Consortium for Political and Social research (ICPSR) study #2896. Measure is based on historical MSA boundary definitions.

	OLS	First-stage			Second-stage			
					TSLS	TSLS	LIML	JIVE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent:	WRLURI	SDL 1992	SDL 1992	HOR 1990	WRLURI	WRLURI	WRLURI	WRLURI
Share developed residential land (SDL), 1992	1.981***				**1.993****	**2.299****	**2.372****	**2.599***
	(0.512)				(0.953)	(0.911)	(0.985)	(1.424)
Homeownership rate (HOR), 1990	0.594	-0.221			0.599	**3.674**	**4.248***	**10.25***
	(0.977)	(0.175)			(0.842)	(2.275)	(2.556)	(5.958)
Population density in developed residential area (POPD), 1990	-48.29	-3.480	0.207	-6.623	-48.69	42.35	58.58	253.5
	(75.63)	(9.956)	(13.22)	(12.74)	(83.91)	(106.1)	(114.0)	(236.1)
Share democratic votes in state, average 1988 and 1992	0.0429**	-0.00254	-0.00232	-0.000687	0.0429**	0.0463**	0.0471**	0.0514**
	(0.0196)	(0.00262)	(0.00255)	(0.00140)	(0.0192)	(0.0195)	(0.0199)	(0.0246)
Household wage (in thousand dollar), 1990	0.0167	0.00307***	0.00346***	-0.000953	0.0167	0.0229	0.0240	0.0373*
	(0.0141)	(0.00101)	(0.00116)	(0.00117)	(0.0138)	(0.0165)	(0.0172)	(0.0218)
Region = Northeast	0.458*	0.0338	0.0389	-0.0271	0.457*	0.486**	0.490**	0.577*
	(0.253)	(0.0227)	(0.0247)	(0.0178)	(0.237)	(0.227)	(0.228)	(0.291)
Region = South	0.323*	-0.109***	-0.104***	-0.00546	0.323**	0.436***	0.457***	0.691**
	(0.168)	(0.0314)	(0.0326)	(0.0145)	(0.159)	(0.169)	(0.173)	(0.289)
Region = West	0.833***	-0.0808*	-0.0706	-0.0315	0.833***	0.991***	1.020***	1.334***
	(0.169)	(0.0448)	(0.0435)	(0.0218)	(0.159)	(0.206)	(0.217)	(0.354)
Metro area has major border with coast		0.0786***	0.0809***	-0.0207*				
		(0.0227)	(0.0219)	(0.0114)				
Average temperature in January, 1941-1970		0.00510**	0.00523***	-0.00101				
		(0.00189)	(0.00189)	(0.000674)				
Share metro area that is classified as consisting of plains		0.0201	0.0220	9.63e-07				
		(0.0148)	(0.0143)	(0.0138)				
Population density in 1880		105***	112***	-4.04e-05*				
		(18.2)	(22.)	(2.18e-05)				
Share households with married couples and no children in 1990			-0.147	1.324***				
			(0.350)	(0.263)				
Constant	-3.361**	0.119	-0.0214	0.384***	-3.368***	-6.011***	-6.508***	-11.60**
	(1.429)	(0.197)	(0.214)	(0.141)	(1.284)	(2.214)	(2.463)	(5.249)
Adjusted *R*-squared	0.377	0.594	0.586	0.658				
Kleibergen-Paap rk Wald F-statistic					12.0	5.6	5.6	

Notes: Robust standard errors in parentheses (observations are clustered by US state). *** p<0.01, ** p<0.05, * p<0.1. **Bold** coefficients are instrumented.

Table 3

Robustness check: Drop one instrument or any combination of two instruments at a time *(N=93)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	PANEL A: First-stage / Dependent variable: Share developed residential land (SDL), 1992										
Excluded instruments:	Baseline	Drop one excluded instrument at a time				Drop two excluded instruments at a time					
Major border with coast	0.0809***		0.110***	0.0838***	0.0796***				0.117***	0.105***	0.0862***
	(0.0219)		(0.0351)	(0.0220)	(0.0257)				(0.0386)	(0.0368)	(0.0263)
Average January temperature	0.00523***	0.00639***		0.00541***	0.00447**		0.00674***	0.00563**			0.00479**
	(0.00189)	(0.00229)		(0.00194)	(0.00199)		(0.00242)	(0.00237)			(0.00208)
Share plains	0.0220	0.0334**	0.0394*		0.0470**	0.0620*		0.0580**		0.0588**	
	(0.0143)	(0.0163)	(0.0232)		(0.0199)	(0.0354)		(0.0218)		(0.0255)	
Population density in 1880	112***	111***	95***	118***		87***	121***		105***		
	(22.3)	(25.1)	(23.3)	(22.9)		(30.3)	(26.4)		(22.4)		
Share married and no children	-0.147	0.178	0.202	-0.220	0.119	0.799	0.0843	0.438	0.0905	0.389	-0.0132
	(0.350)	(0.454)	(0.319)	(0.354)	(0.377)	(0.606)	(0.474)	(0.485)	(0.331)	(0.355)	(0.412)
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted *R*-squared	0.586	0.534	0.499	0.587	0.507	0.392	0.531	0.458	0.492	0.445	0.494
	PANEL B: Second-stage / TSLS / Dependent variable: WRLURI										
Share developed residential land (SDL)	**2.299**	**1.963**	**2.962**	**2.114**	**2.865***	**2.869**	**1.604**	**2.566**	**2.598**	**3.805***	**2.463***
	(0.911)	(0.994)	(1.186)	(0.921)	(0.940)	(1.369)	(1.074)	(1.064)	(1.155)	(1.340)	(0.953)
Homeownership rate (HOR)	**3.674**	**3.947***	**3.395**	**3.563**	**2.800**	**3.453**	**3.939**	**3.102**	**3.384**	**2.259**	**3.082**
	(2.275)	(2.319)	(2.211)	(2.329)	(2.445)	(2.246)	(2.408)	(2.494)	(2.268)	(2.496)	(2.423)
Population density in developed residential area	42.35	67.28	2.013	47.13	-13.87	8.301	83.66	10.32	18.55	-76.06	14.47
	(106.1)	(102.8)	(121.1)	(105.0)	(126.9)	(123.6)	(100.5)	(126.1)	(118.7)	(149.8)	(119.4)
Share democratic votes	0.0463**	0.0444**	0.0503**	0.0451**	0.0494***	0.0498**	0.0421**	0.0477**	0.0480**	0.0550***	0.0471**
	(0.0195)	(0.0199)	(0.0196)	(0.0197)	(0.0187)	(0.0199)	(0.0205)	(0.0188)	(0.0200)	(0.0191)	(0.0192)
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap statistic	5.6	7.4	12.4	6.9	11.2	6.6	9.8	7.5	16.4	8.2	12.4
	PANEL C: Second-stage / LIML / Dependent variable: WRLURI										
Share developed residential land (SDL)	**2.372**	**1.962***	**3.081**	**2.128**	**2.962***	**3.007***	**1.604**	**2.608**	**2.619**	**3.872***	**2.478**
	(0.985)	(1.082)	(1.315)	(0.944)	(1.036)	(1.708)	(1.075)	(1.197)	(1.188)	(1.403)	(0.987)
Homeownership rate (HOR)	**4.248***	**4.448***	**3.719**	**3.739**	**3.079**	**3.764**	**3.941**	**3.406**	**3.477**	**2.288**	**3.203**
	(2.556)	(2.565)	(2.379)	(2.419)	(2.683)	(2.485)	(2.409)	(2.755)	(2.315)	(2.561)	(2.506)
Population density in developed residential area	58.58	84.49	7.540	52.48	-8.812	12.52	83.76	18.78	20.70	-78.16	17.88
	(114.0)	(110.7)	(128.6)	(107.5)	(137.2)	(143.5)	(100.5)	(139.1)	(120.9)	(154.0)	(122.9)
Share democratic votes	0.0471**	0.0446**	0.0512**	0.0453**	0.0502***	0.0508**	0.0421**	0.0481**	0.0482**	0.0555***	0.0472**
	(0.0199)	(0.0204)	(0.0200)	(0.0198)	(0.0189)	(0.0209)	(0.0205)	(0.0191)	(0.0201)	(0.0193)	(0.0193)
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap statistic	5.6	7.4	12.4	6.9	11.2	6.6	9.8	7.5	16.4	8.2	12.4

Notes: Robust standard errors in parentheses (observations are clustered by US state). *** $p<0.01$, ** $p<0.05$, * $p<0.1$. **Bold** coefficients are instrumented.

Table 4

Robustness check: Endogenize population density in developed residential area *(N=93)*

Dependent variable:	First-stage			Second-stage		
				TSLS	LIML	JIVE
	SDL	HOR	PDDR	WRLURI		
	(1)	(2)	(3)	(4)	(5)	(6)
Share developed residential land (SDL)				**3.114***	**3.432**	**3.185***
				(0.939)	(1.372)	(1.634)
Homeownership rate (HOR)				**0.618**	**-0.188**	**4.846**
				(3.802)	(5.885)	(7.876)
Population density in developed residential area (POPD)				**-185.6**	**-252.8**	**-17.36**
				(216.3)	(355.8)	(424.8)
Share democratic votes	-0.00232	-0.000785	1.47e-05	0.0502***	0.0518***	0.0526***
	(0.00251)	(0.00136)	(1.66e-05)	(0.0179)	(0.0182)	(0.0194)
Household wage (in thousand dollar)	0.00346***	-0.000866	-1.31e-05	0.0155	0.0133	0.0244
	(0.00118)	(0.00113)	(1.82e-05)	(0.0188)	(0.0237)	(0.0213)
Region = Northeast	0.0389	-0.0265	-9.34e-05	0.437*	0.418	0.484*
	(0.0246)	(0.0165)	(0.000289)	(0.243)	(0.265)	(0.262)
Region = South	-0.104***	-0.000763	-0.000708**	0.269	0.221	0.442
	(0.0321)	(0.0158)	(0.000262)	(0.227)	(0.322)	(0.407)
Region = West	-0.0706	-0.0294	-0.000307	0.847***	0.808**	1.059**
	(0.0433)	(0.0221)	(0.000264)	(0.231)	(0.316)	(0.426)
Major border with coast	0.0810***	-0.0244	0.000558**			
	(0.0231)	(0.0153)	(0.000258)			
Average January temperature	0.00523***	-0.00112*	1.66e-05*			
	(0.00187)	(0.000642)	(9.44e-06)			
Share plains	0.0219	0.00300	-0.000453**			
	(0.0148)	(0.0127)	(0.000212)			
Population density in 1880	113***	-51***	1.6***			
	(9.8)	(5.9)	(.11)			
Share married and no children	-0.150	1.427***	-0.0155***			
	(0.307)	(0.270)	(0.00398)			
Constant	-0.0200	0.342**	0.00646***	-3.456	-2.796	-7.069
	(0.206)	(0.138)	(0.00156)	(3.469)	(5.202)	(6.579)
Adjusted *R*-squared	0.591	0.657	0.672			
Kleibergen-Paap statistic				2.0	2.0	

Notes: Robust standard errors in parentheses (observations are clustered by US state). *** p<0.01, ** p<0.05, * p<0.1. **Bold** coefficients are instrumented.

Table 5

Robustness check: Use alternative measures to proxy for relative influence of owners of developed land *(N=93)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	PANEL A: First stage / Dependent variable: Alternative proxy measures for the relative influence of owners of developed land									
	Alternative measures of share developed land (SDL)									Aggr. property value per m^2 developable land
	Developed residential + industrial	20 km radius, residential + industrial	20 km radius, residential only	Excluding parkland, res. + ind.	Excluding parkland, res. only	Average place in MSA, res. + ind.	Av. place in MSA, res. only	Median place in MSA, res. + ind.	Median place in MSA, res. only	
Major border with coast	0.0896***	0.127**	0.120**	0.102***	0.0932***	0.107**	0.110***	0.113**	0.121***	21.33**
	(0.0236)	(0.0563)	(0.0536)	(0.0284)	(0.0275)	(0.0392)	(0.0387)	(0.0452)	(0.0443)	(8.653)
Average January temperature	0.00557***	0.00622	0.00694	0.00619*	0.00614*	0.00606*	0.00684**	0.00735**	0.00785**	0.749***
	(0.00203)	(0.00445)	(0.00444)	(0.00322)	(0.00314)	(0.00320)	(0.00337)	(0.00338)	(0.00360)	(0.225)
Share plains	0.0244	0.0725	0.0763	0.0227	0.0219	0.0523*	0.0450	0.0717*	0.0615	-3.828
	(0.0165)	(0.0525)	(0.0517)	(0.0200)	(0.0174)	(0.0301)	(0.0306)	(0.0392)	(0.0386)	(4.819)
Population density in 1880	115***	141***	163***	60**	61**	130***	148***	127***	147***	28,700***
	(22.6)	(48.5)	(50.6)	(24.)	(23.2)	(34.8)	(36.9)	(43.1)	(46.6)	(8,820)
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted *R*-squared	0.579	0.242	0.278	0.376	0.357	0.322	0.359	0.325	0.357	0.755
	PANEL B: Second-stage / Dependent variable: WRLURI / TSLS									
Share developed residential land (SDL)	**2.205***	**2.115***	**1.875***	**2.184***	**2.217***	**1.770***	**1.586***	**1.556***	**1.416***	**0.0105*
	(0.852)	(0.768)	(0.686)	(0.797)	(0.807)	(0.642)	(0.591)	(0.564)	(0.530)	(0.00614)
Homeownership rate (HOR)	**3.797	**5.485**	**5.055**	**3.469	**3.332	**2.926	**2.883	**2.676	**2.607	**4.118*
	(2.324)	(2.277)	(2.325)	(2.397)	(2.352)	(2.091)	(2.070)	(2.111)	(2.108)	(2.371)
Population density in developed residential area	38.37	106.4	105.5	65.44	69.49	72.93	74.21	72.02	71.52	-40.17
	(108.0)	(110.3)	(110.4)	(105.5)	(103.7)	(101.8)	(101.8)	(101.1)	(102.3)	(138.5)
Share democratic votes	0.0483**	0.0498**	0.0467**	0.0500**	0.0474**	0.0438**	0.0413**	0.0473**	0.0437**	0.0455**
	(0.0198)	(0.0215)	(0.0208)	(0.0204)	(0.0198)	(0.0196)	(0.0196)	(0.0199)	(0.0201)	(0.0215)
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap statistic	4.9	4.4	5.4	2.8	2.7	8.9	12.3	8.9	11.6	8.4
	PANEL C: Second-stage / Dependent variable: WRLURI / LIML									
Share developed residential land (SDL)	**2.289**	**2.323**	**2.054**	**2.309**	**2.344**	**1.897***	**1.690**	**1.665***	**1.513***	**0.0122*
	(0.924)	(0.911)	(0.799)	(0.923)	(0.935)	(0.721)	(0.657)	(0.622)	(0.583)	(0.00687)
Homeownership rate (HOR)	**4.377*	**6.040**	**5.629**	**4.046	**3.902	**3.257	**3.255	**2.934	**2.887	**4.977*
	(2.604)	(2.425)	(2.487)	(2.667)	(2.615)	(2.262)	(2.263)	(2.263)	(2.281)	(2.648)
Population density in developed residential area	53.82	115.1	116.5	80.81	85.13	80.60	83.83	77.85	78.28	-42.83
	(115.9)	(114.4)	(115.4)	(114.0)	(112.1)	(107.5)	(108.3)	(106.5)	(108.4)	(153.5)
Share democratic votes	0.0492**	0.0517**	0.0484**	0.0513**	0.0486**	0.0448**	0.0422**	0.0485**	0.0447**	0.0480**
	(0.0203)	(0.0223)	(0.0213)	(0.0211)	(0.0204)	(0.0197)	(0.0197)	(0.0200)	(0.0202)	(0.0223)
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap statistic	4.9	4.4	5.4	2.8	2.7	8.9	12.3	8.9	11.6	8.4

Notes: Robust standard errors in parentheses (observations are clustered by US state). *** p<0.01, ** p<0.05, * p<0.1. **Bold** coefficients are instrumented.

Table 6

Additional robustness checks: Base specification but with controls for total open land or open land per capita *(N=93)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Dependent variable: WRLURI				
	TSLS	TSLS	LIML	JIVE	TSLS	TSLS	LIML	JIVE
Share developed residential land (SDL), 1992	**1.946****	**2.325****	**2.452****	**2.921***	**1.883****	**2.373****	**2.532****	**3.027***
	(0.926)	(0.909)	(1.010)	(1.643)	(0.951)	(0.947)	(1.062)	(1.828)
Homeownership rate (HOR), 1990	0.288	**3.758***	**4.562***	**12.02**	0.106	**4.012***	**4.884***	**11.92***
	(0.746)	(2.255)	(2.621)	(8.051)	(0.781)	(2.223)	(2.578)	(6.965)
Population density in developed residential area (POPD), 1990	-66.56	38.02	60.55	300.8	-72.37	45.09	69.14	297.6
	(81.28)	(106.9)	(117.2)	(307.0)	(82.04)	(106.0)	(115.8)	(278.6)
Share democratic votes in state, average 1988 and 1992	0.0405**	0.0453**	0.0466**	0.0546*	0.0381**	0.0452**	0.0471**	0.0571**
	(0.0192)	(0.0200)	(0.0206)	(0.0276)	(0.0194)	(0.0203)	(0.0211)	(0.0275)
Household wage (in thousand dollar), 1990	0.0155	0.0226	0.0241	0.0405	0.0121	0.0218	0.0238	0.0420
	(0.0144)	(0.0173)	(0.0182)	(0.0260)	(0.0137)	(0.0171)	(0.0181)	(0.0262)
Region = Northeast	0.454*	0.485**	0.489**	0.582*	0.474**	0.494**	0.494**	0.562*
	(0.233)	(0.223)	(0.226)	(0.321)	(0.234)	(0.226)	(0.230)	(0.313)
Region = South	0.305*	0.434***	0.463***	0.752**	0.295*	0.442***	0.473***	0.750**
	(0.159)	(0.168)	(0.174)	(0.364)	(0.159)	(0.166)	(0.172)	(0.333)
Region = West	0.908***	1.051***	1.083***	1.403***	0.916***	1.055***	1.084***	1.352***
	(0.170)	(0.208)	(0.218)	(0.397)	(0.170)	(0.203)	(0.213)	(0.339)
Total open land in MSA in thousand km^2	-0.00744**	-0.00469*	-0.00401	0.00153				
	(0.00360)	(0.00250)	(0.00259)	(0.00784)				
Total open land in MSA in km^2 per person					-7.217***	-3.259	-2.294	4.204
					(2.646)	(2.916)	(3.243)	(7.470)
Constant	-2.906**	-5.966***	-6.684**	-13.19*	-2.557**	-6.140***	-6.953***	-13.30**
	(1.199)	(2.270)	(2.606)	(7.236)	(1.254)	(2.298)	(2.636)	(6.436)
Kleibergen-Paap statistic	19.0	5.2	5.2		12.1	5.1	5.1	

Notes: Robust standard errors in parentheses (observations are clustered by US state). *** $p<0.01$, ** $p<0.05$, * $p<0.1$. **Bold** coefficients are instrumented.

Table 7

Robustness Check: Use data on land use and regulation from late 1970s/early 1980s and explain growth rate in housing supply *(N=81)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PANEL A: First-stage and second-stage results (Explaining restrictiveness of land use regulation)						
	OLS (First-stage)			TSLS	LIML	TSLS	LIML
Dependent variables::	SDL 76	SDL 76	HOR 80	SAKS Regulatory Index			
Share developed residential land, 1976 (SDL 76)				**6.851***	**7.506***	**6.437***	**6.824***
				(1.450)	(1.704)	(1.492)	(1.597)
Homeownership rate, 1980 (HOR 80)	-0.113			-0.724	-0.568	**-3.269**	**-3.281**
	(0.121)			(1.997)	(2.076)	(2.372)	(2.424)
Population density in developed residential area, 1980	1.534	4.425	-9.262	-26.81	-43.18	-96.34	-109.4
	(11.63)	(12.20)	(7.434)	(130.3)	(134.4)	(147.3)	(152.2)
Share democratic votes in state, average 1972 and 1976	-0.00224	-0.00151	-0.00248	0.0245	0.0267	0.00936	0.0101
	(0.00315)	(0.00314)	(0.00156)	(0.0294)	(0.0294)	(0.0296)	(0.0296)
Major border with coast	0.0314	0.0301	-0.0265				
	(0.0265)	(0.0226)	(0.0160)				
Average January temperature	0.00451**	0.00440**	-0.00172*				
	(0.00175)	(0.00176)	(0.000908)				
Share plains	0.0388	0.0355	0.0186				
	(0.0427)	(0.0428)	(0.0154)				
Population density in 1880	0.000316**	0.000365***	-0.000261***				
	(0.000129)	(0.000116)	(8.78e-05)				
Share married and no children		0.145	1.203***				
		(0.220)	(0.298)				
Other controls (incl. household wage) and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted *R*-squared	0.307	0.305	0.642				
Kleibergen-Paap statistic				4.2	4.2	3.6	3.6
	PANEL B: Explaining growth rate in housing supply						
Dependent variable:	Percent change in housing units 1990 to 2000						
Predicted Saks-Index of Housing Supply Regulation				-0.0310***	-0.0293***	-0.0280***	-0.0271***
				(0.00882)	(0.00826)	(0.00890)	(0.00898)
Percent change, housing units, 1980-1990				0.364***	0.364***	0.360***	0.360***
				(0.0810)	(0.0825)	(0.0897)	(0.0855)
Constant				0.0600***	0.0600***	0.0608***	0.0608***
				(0.0133)	(0.0140)	(0.0149)	(0.0139)
Adjusted *R*-squared				0.468	0.467	0.453	0.453

Notes for both Panels: *** p<0.01, ** p<0.05, * p<0.1. **Bold** coefficients are instrumented. Notes for Panel A: Robust standard errors in parentheses (obs. are clustered by US state). Region fixed effects and constant included (yes). Notes for Panel B: Standard errors are (robust) bootstrap standard errors (using 1000 bootstrap replications and clustering by US state.)

Figures

Figure 1. Regulatory restrictiveness and potential determinants.
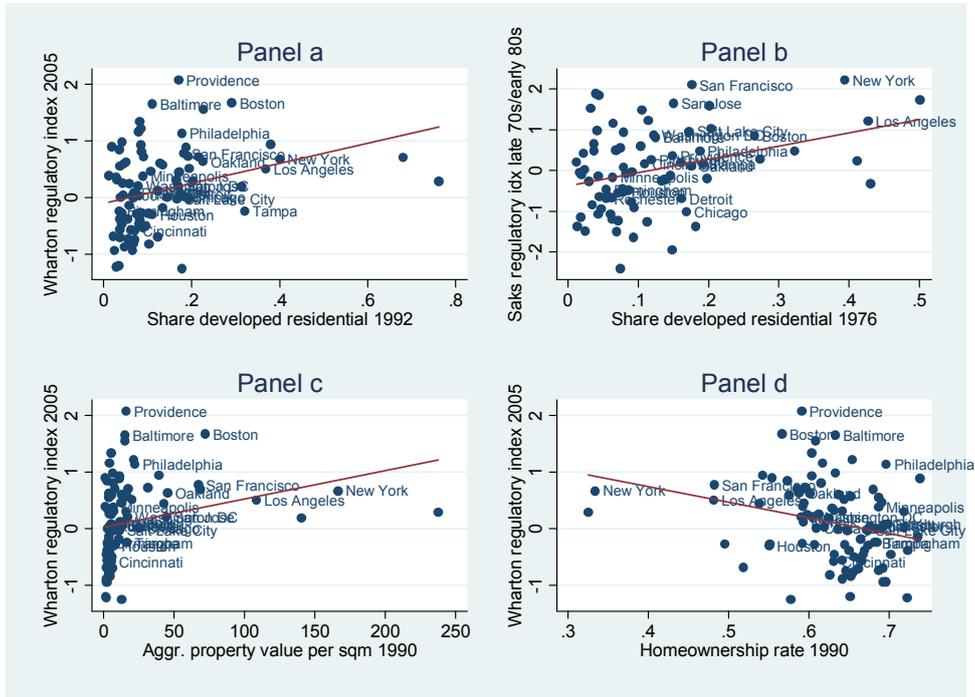
Figure 2. Spatial equilibrium.
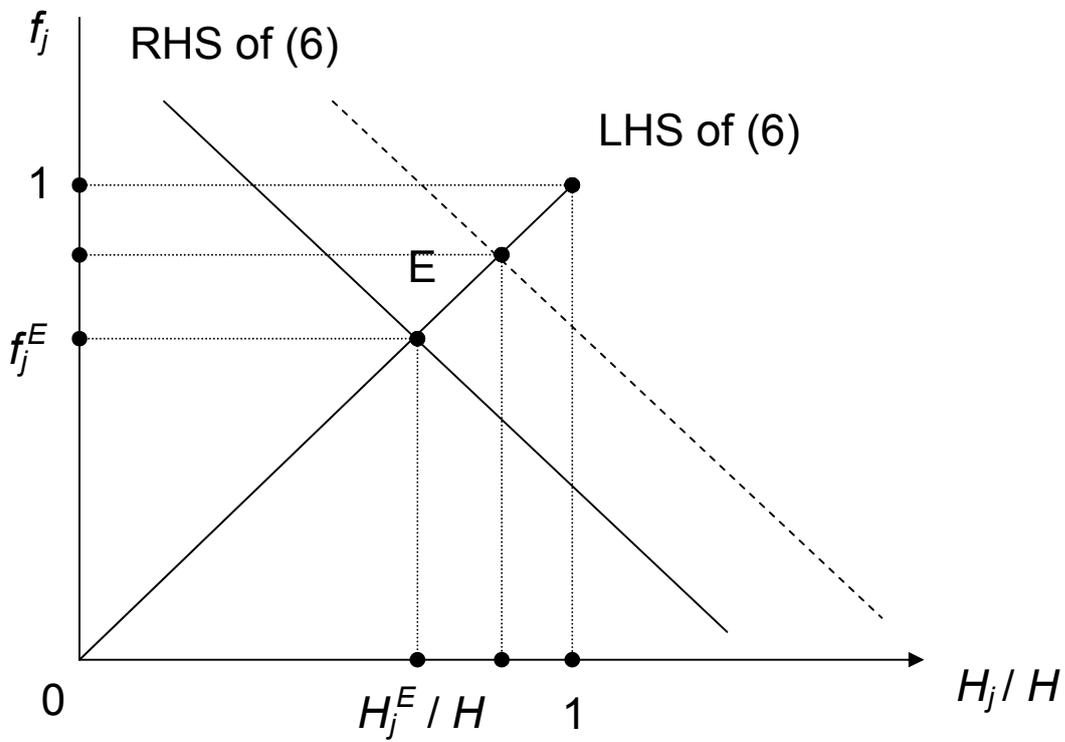
Appendix A. Proofs of Propositions

Proof of Proposition 1. Let us generalize (3) and (6) for any $\sigma > 0$:

$$\forall j \in \mathfrak{I}_+: \qquad f_j = \frac{H_j}{H} \quad \Rightarrow \quad f_j = \frac{H_j}{H} = \frac{\left(\omega_j - t_j - \tau H_j\right)^{1/\sigma}}{\sum_{k \in \mathfrak{I}+} \left(\omega_k - t_k - \tau H_k\right)^{1/\sigma}}.$$

The RHS of this expression is decreasing in H_j by inspection and it belongs to the unit interval for any $H_j \in [0, H]$ (with complementary slackness), so there always exists a jurisdiction j such that a positive mass of households desires to live in j. Conversely, the left-hand-side (LHS) is linearly increasing in H_j and it spans over the unit interval. It follows that there exists at least one $j \in \mathfrak{I}$ such that the LHS and the RHS intersect exactly once in the interior of the unit interval; if (14) holds, then the LHS and the RHS intersect exactly once for all $k \in \mathfrak{I}$. *QED.*

Proof of Proposition 2.[28] SPE in pure strategies characterized by (11) and (12) exist if $R_j(\mathbf{t})$ is quasi-concave in t_j (this is a consequence of Kakutani's (1941) fixed point theorem). Twice differentiating (9) with respect to t_j (keeping $\mathbf{t_{-j}^0}$ constant) yields

$$\frac{\partial^2 R_j}{\partial t_j^2} \equiv \left[2 + \tau \frac{\partial H_j}{\partial t_j}\right] \frac{\partial H_j}{\partial t_j} + \tau H_j \frac{\partial^2 H_j}{\partial t_j^2}.$$

This is negative (and as a consequence R_j is quasi-concave in t_j) if

$$2 + \tau \frac{\partial H_j}{\partial t_j} > 0 \qquad \text{and} \qquad \frac{\partial^2 H_j}{\partial t_j^2} < 0$$

because $\partial H_j/\partial t_j < 0$ by (7). Using (7), (14) and the definition of f_j and rearranging yield

$$2 + \tau \frac{\partial H_j}{\partial t_j} = \frac{2V_j + f_j H \tau \left(1 + f_j\right)}{\omega_j - t_j} > 0 \qquad \text{and} \qquad \frac{\partial^2 H_j}{\partial t_j^2} = -\frac{2H \sum_{k \neq j} \left(\omega_k - t_k\right)}{\left[\sum_{k \in \mathfrak{I}+} \left(\omega_k - t_k\right)\right]^3} < 0 \qquad (19)$$

as was to be shown. *QED.*

Proofs of Proposition 3 and Corollary 3.1.[29] It is convenient to rewrite the SPE relationship (12) as $\omega_j - t_j^0 = \left(1 - f_j^0\right)\left(t_j^0 + f_j^0 H \tau\right)$. Totally differentiating this expression and making use of (11) yields $l_j \mathrm{d} t_j^0 = r_j \left(\mathrm{d}\omega_j - \mathrm{d}t_j^0\right)$, where

---

[28] See Guide to calculations in Appendix C.1 (not intended for publication) for details.

[29] See Guide to calculations in Appendix C.1 (not intended for publication) for details.

$$l_j \equiv \left(V_j^0 + f_j^0 H\tau \right)\left(1 - f_j^0 \right) \quad \text{and} \quad r_j \equiv V_j^0 + f_j^0 \left(1 - f_j^0 \right) t_j^0 + \left(3 - 2 f_j^0 \right)\left(f_j^0 \right)^2 H\tau . \quad (20)$$

Both l_j and r_j are positive by inspection. This implies

$$\frac{d\omega_j - dt_j^0}{d\omega_j} = \frac{l_j}{r_j} > 0 \quad \text{and} \quad 0 < \frac{dt_j^0}{d\omega_j} = \frac{r_j}{l_j + r_j} < 1 . \quad (21)$$

The first property in (21) establishes (iv) and, together with the spatial equilibrium condition (7) (whereby jurisdictions with a higher $\omega_j - t_j^0$ have a higher equilibrium H_j^0) and the location preferences (3) (whereby jurisdictions with a higher V_j^0 command a higher H_j^0), implies (i) and (v), respectively. Finally, turn to (iii): (14) ensures $t_j^0 > 0$ and $V_j^0 > 0$, all j. To see this, recall that the jurisdiction with the minimum equilibrium deterministic utility and the lowest regulatory tax is the jurisdiction with the lowest ω_j – call this jurisdiction 'min'. From (i) and (iv), the least attractive jurisdiction has a less than average equilibrium population share, i.e. $f_{min}^0 < 1/J$ (so that $\kappa_{min}^0 > 1/2$). Using these as well as (12) and (13), we get:

$$V_{min}^0 > \frac{\left(1 - J^{-1} \right)\omega_{min} - J^{-1} H\tau}{2 - f_{min}^0} \quad \text{and} \quad t_{min}^0 > \frac{\omega_{min} - J^{-1} H\tau}{2} . \quad (22)$$

Both lower bounds are positive if $\omega_{min} > H\tau /\left(J - 1 \right)$, i.e., if (14) holds. *QED.*

Appendix B. Appendix Tables not intended for publication

Table U1 (not intended for publication)
MSA-level rankings of measures of regulatory restrictiveness and land scarcity

	(1)		(2)		(3)		(4)		(5)	
Metropolitan Area	Regulatory Tax in % of House Value 1998 [i)]	Rank	Regulatory Index 2005 [ii)]	Rank	Regulatory Index late 1970s & 1980s [ii)]	Rank	Share Developed Land in % 1992 [iii)]	Rank	Share Developed Land in% 1976 [iii)]	Rank
San Francisco	53.1	**1**	0.78	**5**	2.10	**2**	20.8	**10**	25.4	**8**
San Jose	46.9	**2**	0.21	**10**	1.65	**3**	19.9	**12**	25.2	**9**
Los Angeles	33.9	**3**	0.50	**8**	1.21	**4**	44.2	**1**	54.7	**1**
Oakland	32.1	**4**	0.63	**7**	0.10	**14**	26.6	**5**	27.6	**6**
Washington, D.C.	21.9	**5**	0.21	**10**	0.86	**6**	10.2	**17**	16.1	**16**
Newport News, VA	20.7	**6**	0.12	**12**		**(-)**	17.5	**13**	20.5	**14**
Boston	18.6	**7**	1.67	**2**	0.86	**7**	33.2	**4**	33.9	**3**
New York	12.2	**8**	0.67	**6**	2.21	**1**	43.9	**2**	50.7	**2**
Manhattan	>50									
Salt Lake City	11.9	**9**	-0.03	**17**	0.96	**5**	23.3	**6**	28.8	**4**
Chicago	5.7	**10**	0.01	**16**	-1.01	**20**	22.4	**7**	26.2	**7**
Baltimore	1.8	**11**	1.65	**3**	0.80	**8**	14.2	**15**	18.3	**15**
Birmingham	0	**12**	-0.24	**18**	-0.46	**16**	4.8	**21**	7.5	**21**
Cincinnati	0	**12**	-0.58	**21**	0.16	**12**	9.3	**18**	14.8	**18**
Detroit	0	**12**	0.07	**14**	-0.69	**19**	22.3	**8**	22.9	**11**
Houston	0	**12**	-0.30	**20**	-0.52	**17**	14.8	**14**	12.3	**19**
Minneapolis	0	**12**	0.38	**9**	-0.16	**15**	11.1	**16**	9.4	**20**
Philadelphia	0	**12**	1.13	**4**	0.47	**9**	21.1	**9**	27.8	**5**
Pittsburgh	0	**12**	0.08	**13**	0.26	**11**	8.5	**19**	15.4	**17**
Providence	0	**12**	2.07	**1**	0.35	**10**	20.6	**11**	22.2	**12**
Rochester	0	**12**	0.04	**15**	-0.68	**18**	5.3	**20**	21.8	**13**
Tampa	0	**12**	-0.24	**18**	0.16	**13**	35.6	**3**	24.2	**10**

Pair:	Correlation	Rank Correlation		Correlation	Rank Correlation
HOR 1990, (4)	-0.61	-0.44	(2), HOR 1990	-0.24	-0.32
HOR 1980, (5)	-0.69	-0.39	(3), HOR 1980	-0.64	-0.45
(1), (2)	0.12	0.37	(1), (4)	0.33	0.36
(1), (3)	0.68	0.65	(2), (4)	0.28	0.31
(1), HOR 1990	-0.52	-0.68	(3), (5)	0.50	0.34

Sources: [i)] Estimated regulatory tax values are from Glaeser *et al.* (2005a). [ii)] Regulatory index values are from Saiz (2008) and Saks (2008 respectively. [iii)] The share developed land measures are derived from the 1992 and 1976 National Land Cover Data. Homeownership rates (HOR) are from the 1990 and 1980 US Census (tract level data geographically matched to the metropolitan area level).

Table U2 (not intended for publication)
Relationship between land scarcity and restrictiveness of land use regulations *(N=93)*
Method: OLS

	Dependent Variable: WRLURI from 2005						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share developed, 1992 (SDL)	1.778***	1.340**	1.411**	2.018***	1.950***	1.950***	1.981***
	(0.598)	(0.621)	(0.554)	(0.498)	(0.555)	(0.555)	(0.512)
Homeownership rate, 1990 (HOR)		-1.729*	-1.976	-1.296	-1.045	-1.045	0.594
		(0.969)	(1.249)	(1.083)	(1.140)	(1.140)	(0.977)
Population density in developed area, 1990			-26.56	-82.67	-78.23	-78.23	-48.29
			(84.99)	(80.12)	(82.71)	(82.71)	(75.63)
Share democratic votes in state, average 1988/1992				0.0625***	0.0586***	0.0586***	0.0429**
				(0.0180)	(0.0189)	(0.0189)	(0.0196)
Household wage (in thousand US dollar), 1990					0.0145	0.0145	0.0167
					(0.0138)	(0.0138)	(0.0141)
North East							0.458*
							(0.253)
South							0.323*
							(0.168)
West							0.833***
							(0.169)
Constant	-0.104	1.036	1.252	-2.155*	-2.557**	-2.557**	-3.361**
	(0.101)	(0.628)	(0.962)	(1.260)	(1.258)	(1.258)	(1.429)
Adjusted R-squared	0.086	0.100	0.092	0.259	0.261	0.261	0.377

Notes: Robust standard errors in parentheses (observations are clustered by US state). *** p<0.01, ** p<0.05, * p<0.1.

Table U3 (not intended for publication)
Robustness check: Replicate Table 5 but with 3 endogenous variables *(N=93)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	PANEL A: First stage / Dependent variable: Alternative proxy measures for the relative influence of owners of developed land									
	Alternative measures of share developed land (SDL)									Aggr. property value per m^2 developable land
	Developed residential + industrial	20 km radius, residential + industrial	20 km radius, residential only	Excluding parkland, res. + ind.	Excluding parkland, res. only	Average place in MSA, res. + ind.	Av. place in MSA, res. only	Median place in MSA, res. + ind.	Median place in MSA, res. only	
Major border with coast	0.0913***	0.115**	0.105*	0.102***	0.0914***	0.0957**	0.0967**	0.102**	0.107**	25.38***
	(0.0247)	(0.0549)	(0.0523)	(0.0297)	(0.0285)	(0.0403)	(0.0396)	(0.0460)	(0.0452)	(8.735)
Average January temperature	0.00562***	0.00587	0.00650	0.00618*	0.00609*	0.00572*	0.00643*	0.00702**	0.00746**	0.869***
	(0.00199)	(0.00446)	(0.00447)	(0.00321)	(0.00316)	(0.00317)	(0.00336)	(0.00333)	(0.00358)	(0.215)
Share plains	0.0231	0.0822	0.0885*	0.0230	0.0234	0.0616*	0.0560	0.0806*	0.0724*	-7.116
	(0.0164)	(0.0532)	(0.0524)	(0.0203)	(0.0180)	(0.0344)	(0.0348)	(0.0433)	(0.0423)	(4.395)
Population density in 1880	120***	107***	121***	59***	56***	98***	109***	96***	109***	40,200***
	(10.8)	(24.8)	(24.6)	(16)	(15.4)	(16.1)	(16.7)	(19.4)	(19.9)	(3300)
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted *R*-squared	0.584	0.245	0.278	0.383	0.365	0.322	0.356	0.328	0.356	0.737
	PANEL B: Second-stage / Dependent variable: WRLURI / TSLS									
Share developed residential land (SDL)	**2.913***	**2.483***	**2.287***	**2.383***	**2.439***	**2.313***	**2.145***	**1.896***	**1.808***	**0.0217***
	(0.886)	(0.865)	(0.749)	(0.823)	(0.824)	(0.765)	(0.708)	(0.598)	(0.577)	(0.00751)
Homeownership rate (HOR)	**0.923**	**3.465**	**2.767**	**2.129**	**1.913**	**-0.337**	**-0.562**	**-0.0288**	**-0.476**	**-0.802**
	(3.757)	(3.433)	(3.578)	(3.607)	(3.640)	(3.850)	(3.903)	(3.714)	(3.901)	(4.684)
Population density in developed residential area	**-177.7**	**-42.97**	**-62.53**	**-22.25**	**-22.52**	**-137.6**	**-148.5**	**-97.10**	**-121.7**	**-590.5***
	(214.5)	(198.5)	(202.5)	(178.5)	(179.5)	(201.8)	(205.8)	(190.4)	(201.3)	(350.4)
Share democratic votes	0.0523***	0.0519**	0.0489**	0.0511***	0.0484***	0.0461**	0.0433**	0.0496***	0.0458**	0.0577**
	(0.0180)	(0.0209)	(0.0202)	(0.0192)	(0.0186)	(0.0185)	(0.0185)	(0.0192)	(0.0194)	(0.0242)
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap statistic	2.0	2.1	2.0	2.4	2.4	1.9	1.8	2.0	1.9	1.4
	PANEL C: Second-stage / Dependent variable: WRLURI / LIML									
Share developed residential land (SDL)	**3.177***	**2.671***	**2.481***	**2.494***	**2.567***	**2.540***	**2.407***	**2.079***	**2.013***	**0.0267***
	(1.241)	(1.044)	(0.937)	(1.115)	(1.152)	(1.007)	(1.016)	(0.759)	(0.779)	(0.0115)
Homeownership rate (HOR)	**0.323**	**3.497**	**2.665**	**2.710**	**2.388**	**-1.101**	**-1.631**	**-0.709**	**-1.431**	**-2.620**
	(5.641)	(3.917)	(4.276)	(5.573)	(5.805)	(5.007)	(5.496)	(4.890)	(5.376)	(6.735)
Population density in developed residential area	**-232.7**	**-59.04**	**-86.54**	**-2.614**	**-8.636**	**-189.9**	**-219.6**	**-142.4**	**-183.0**	**-809.3**
	(341.1)	(233.6)	(251.2)	(296.4)	(308.3)	(270.7)	(301.6)	(257.1)	(285.9)	(553.4)
Share democratic votes	0.0540***	0.0534**	0.0503**	0.0523***	0.0495***	0.0474**	0.0445**	0.0512***	0.0472**	0.0632**
	(0.0183)	(0.0214)	(0.0206)	(0.0199)	(0.0190)	(0.0187)	(0.0187)	(0.0195)	(0.0198)	(0.0284)
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap statistic	2.0	2.1	2.0	2.4	2.4	1.9	1.8	2.0	1.9	1.4

Notes: Robust standard errors in parentheses (observations are clustered by US state). *** p<0.01, ** p<0.05, * p<0.1. **Bold** coefficients are instrumented.

Table U4 (not intended for publication)
Replicate Table 7 with reduced set of (statistically significant) excluded instruments *(N=81)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PANEL A: First-stage and second-stage results (Explaining restrictiveness of land use regulation)						
	OLS (First-stage)			TSLS	LIML	TSLS	LIML
Dependent variables::	SDL 76	SDL 76	HOR 80	SAKS Regulatory Index			
Share developed residential land, 1976 (SDL 76)				**7.015***	**7.091***	**6.620***	**6.628***
				(1.533)	(1.555)	(1.535)	(1.537)
Homeownership rate, 1980 (HOR 80)	-0.0958			-0.685	-0.667	**-2.419**	**-2.419**
	(0.142)			(2.043)	(2.052)	(2.303)	(2.304)
Population density in developed residential area (POPD), 1980	0.141	3.469	-9.679	-30.91	-32.81	-74.58	-74.83
	(12.40)	(13.31)	(7.386)	(130.3)	(130.7)	(139.6)	(139.7)
Share democratic votes in state, average 1972 and 1976	-0.00266	-0.00182	-0.00307**	0.0250	0.0253	0.0145	0.0145
	(0.00277)	(0.00277)	(0.00133)	(0.0299)	(0.0299)	(0.0299)	(0.0299)
Household wage (in thousand US dollar), 1980	0.0117***	0.0132***	0.000599	-0.0210	-0.0213	-0.0266	-0.0267
	(0.00313)	(0.00257)	(0.00330)	(0.0418)	(0.0419)	(0.0410)	(0.0410)
Average January temperature	0.00576***	0.00543***	-0.00228**				
	(0.00162)	(0.00179)	(0.00100)				
Population density in 1880	412***	453***	-312***				
	(104)	(85.3)	(71)				
Share married and no children		0.235	1.151***				
		(0.242)	(0.292)				
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted *R*-squared	0.301	0.303	0.631				
Kleibergen-Paap statistic				5.0	5.0	3.9	3.9
	PANEL B: Explaining growth rate in housing supply						
Dependent variable::				Percent change in housing units 1990 to 2000			
Predicted Saks-Index of Housing Supply Regulation				**-0.0316***	**-0.0314***	**-0.0296**	**-0.0296***
				(0.0107)	(0.0111)	(0.0124)	(0.0115)
Percent change, housing units, 1980-1990				0.365***	0.365***	0.362***	0.362***
				(0.0818)	(0.0814)	(0.0873)	(0.0853)
Housing units (in million units), 1990				2.75e-09	2.73e-09	1.48e-09	1.48e-09
				(1.22e-08)	(1.33e-08)	(1.35e-08)	(1.29e-08)
Constant				0.0580***	0.0580***	0.0595***	0.0595***
				(0.0186)	(0.0193)	(0.0195)	(0.0185)
Adjusted *R*-squared				0.461	0.461	0.451	0.451

Notes for both Panels: *** p<0.01, ** p<0.05, * p<0.1. **Bold** coefficients are instrumented. Notes for Panel A: Robust standard errors in parentheses (obs. are clustered by US state). Region fixed effects and constant included (yes). Notes for Panel B: Standard errors are (robust) bootstrap standard errors (using 1000 bootstrap replications and clustering by US state.)

Table U5 (not intended for publication)
Robustness check: Findings for 'contemporaneous' and '70s/80s' specifications
restricted to joint sample of metropolitan areas *(N=63)*

	First-stage	OLS	Second-stage	
			TSLS	LIML
	(1)	(2)	(3)	(4)
PANEL A				
Dependent variables:	SDL	WRLURI, 2005		
Share developed, 1992 (SDL)		1.979***	**3.565***	**4.677****
		(0.611)	(0.977)	(1.821)
Homeownership rate, 1990 (HOR)		0.593	**2.408**	**3.537**
		(1.183)	(1.867)	(2.367)
Population density in developed area, 1990 (POPD)	23.80**	-19.54	-11.29	-11.55
	(11.63)	(94.60)	(125.7)	(153.2)
Share democratic votes in state, average 1988 and 1992	-0.00504*	0.0257	0.0382*	0.0468**
	(0.00264)	(0.0231)	(0.0206)	(0.0219)
Household wage (in thousand US dollar), 1990	0.00263	0.0248	0.0259	0.0265
	(0.00155)	(0.0178)	(0.0181)	(0.0201)
Major border with coast	0.0831**			
	(0.0359)			
Average January temperature	0.00500**			
	(0.00200)			
Share plains	0.0567**			
	(0.0263)			
Population density in 1880	42			
	(145)			
Share married and no children	-0.112			
	(0.423)			
Region fixed effects	Yes	Yes	Yes	Yes
Constant	0.0709	-2.826*	-4.859***	-6.161***
	(0.226)	(1.626)	(1.854)	(2.383)
Adjusted R-squared	0.451	0.315		
Kleibergen-Paap-statistic			4.4	4.4
PANEL B				
Dependent variables:	SDL	SAKS, late 70s/early 80s		
Share developed, 1976 (SDL 76)		2.362***	**6.911***	**10.16****
		(0.653)	(1.548)	(3.798)
Homeownership rate, 1980 (HOR 80)		-3.406**	**-1.207**	**1.262**
		(1.571)	(3.166)	(4.124)
Population density in developed area, 1980 (POPD 80)	-5.112	-34.94	-87.86	-89.41
	(21.74)	(76.32)	(145.4)	(213.0)
Share democratic votes in state, average 1972 and 1976	-0.00202	0.00159	0.0212	0.0395
	(0.00316)	(0.0312)	(0.0264)	(0.0305)
Household wage (in thousand US dollar), 1980	0.000291	-0.0250	0.0116	0.0452
	(0.00480)	(0.0576)	(0.0605)	(0.0694)
Major border with coast	0.0520*			
	(0.0296)			
Average January temperature	0.00440**			
	(0.00198)			
Share plains	0.0421			
	(0.0480)			
Population density in 1880	266			
	(165)			
Share married and no children	-0.852			
	(0.774)			
Region fixed effects	Yes	Yes	Yes	Yes
Constant	0.278	2.123	-1.195	-4.573
	(0.275)	(2.248)	(3.800)	(5.113)
Adjusted R-squared	0.295	0.270		
Kleibergen-Paap-statistic			1.8	1.8

Notes: Robust standard errors in parentheses (observations are clustered by US state). *** p<0.01, ** p<0.05, * p<0.1.
Bold coefficients are instrumented.

Table U6 (not intended for publication)

Additional robustness checks: Remove Californian metro areas, New York City, or both

	Exclude CA			Exclude NYC			Exclude CA and NYC		
	PANEL A: First-stage / Dependent variable: Share developed land								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Major border with coast		0.0787***			0.0770***			0.0727***	
		(0.0260)			(0.0203)			(0.0252)	
Average January temperature		0.00636***			0.00542***			0.00620**	
		(0.00227)			(0.00181)			(0.00234)	
Share plains		0.0160			0.0191			0.0166	
		(0.0175)			(0.0148)			(0.0181)	
Population density in 1880		108***			137***			136***	
		(25.3)			(21.8)			(32.1)	
Share married and no children		0.0820			-0.143			0.0917	
		(0.333)			(0.349)			(0.322)	
Other controls and constant		Yes			Yes			Yes	
Observations		82			92			81	
Adjusted R-squared		0.621			0.582			0.614	
	PANEL B: Second-stage / Dependent variable: WRLURI								
	OLS	TSLS	LIML	OLS	TSLS	LIML	OLS	TSLS	LIML
Share developed residential land	2.336***	**2.841***	**2.955***	1.981***	**2.395***	**2.490**	2.332***	**2.924***	**3.042***
(SDL)	(0.483)	(0.988)	(1.073)	(0.510)	(0.924)	(1.001)	(0.481)	(1.017)	(1.104)
Homeownership rate (HOR)	0.976	**4.817*	**5.584*	0.598	**3.978**	**4.448**	0.900	**5.466***	**6.039***
	(1.085)	(2.595)	(2.990)	(1.026)	(1.984)	(2.145)	(1.197)	(2.025)	(2.205)
Population density in developed	-31.86	76.94	98.08	-48.43	30.62	39.86	-29.04	75.04	86.22
residential area (POPD)	(82.91)	(103.3)	(111.5)	(77.95)	(87.93)	(89.60)	(89.16)	(83.66)	(83.78)
Share democratic votes	0.0475**	0.0531***	0.0543***	0.0429**	0.0462**	0.0469**	0.0477**	0.0527***	0.0536**
	(0.0189)	(0.0196)	(0.0203)	(0.0198)	(0.0199)	(0.0202)	(0.0191)	(0.0204)	(0.0209)
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	82	82	82	92	92	92	81	81	81
Adjusted R-squared	0.369			0.372			0.362		
Kleibergen-Paap statistic		4.9	4.9		5.2	5.2		4.9	4.9

Notes: Robust standard errors in parentheses (observations are clustered by US state). *** p<0.01, ** p<0.05, * p<0.1. **Bold** coefficients are instrumented.

Additional robustness checks: Add interaction 'share democratic voters' x household wage and drop variable 'share democratic voters' *(N=93)*

	Add interaction SDV x wage			Drop variable 'share democratic voters'		
	PANEL A: First-stage / Dependent variable: Share developed land					
	(1)	(2)	(3)	(4)	(5)	(6)
Major border with coast		0.0797***			0.0784***	
		(0.0224)			(0.0216)	
Average January temperature		0.00519**			0.00541***	
		(0.00197)			(0.00190)	
Share plains		0.0219			0.0262*	
		(0.0141)			(0.0134)	
Population density in 1880		114***			117***	
		(23.7)			(20.7)	
Share married and no children		-0.175			-0.116	
		(0.409)			(0.342)	
Other controls and constant		Yes			Yes	
Adjusted R-squared		0.582			0.586	
	PANEL B: Second-stage / Dependent variable: WRLURI					
	OLS	TSLS	LIML	OLS	TSLS	LIML
Share developed residential land (SDL)	2.039***	2.609**	2.812**	1.571***	**1.699***	**1.740***
	(0.496)	(1.060)	(1.209)	(0.543)	(0.949)	(1.012)
Homeownership rate (HOR)	0.803	4.421	5.438	0.307	**3.437**	**4.009**
	(1.016)	(2.768)	(3.366)	(1.170)	(2.243)	(2.508)
Population density in developed residential area (POPD)	-38.75	70.30	99.07	-37.93	63.27	80.97
	(79.19)	(115.0)	(129.6)	(85.62)	(105.4)	(112.6)
Share democratic votes	0.0897	0.201	0.234*			
	(0.115)	(0.127)	(0.142)			
Share democratic votes x household wage	-0.00154	-0.00509	-0.00612			
	(0.00370)	(0.00405)	(0.00449)			
Household wage	0.0938	0.277	0.329			
	(0.189)	(0.207)	(0.231)			
Other controls and constant	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.371			0.326		
Kleibergen-Paap statistic		3.4	3.4		6.4	6.4

Notes: Robust standard errors in parentheses (observations are clustered by US state). *** p<0.01, ** p<0.05, * p<0.1.
Bold coefficients are instrumented.

Quantitative effects

	Relative change in land scarcity (Change in percent)	Absolute change in land scarcity (Change in percentage points)
	Effect of major access to coast line	
Table 2 (3)	+65.2%	+8.1%
	Effect of one std. dev. increase in January temperature	
Table 2 (3)	+52.7%	+6.5%
	Effect of one std. dev. increase in 'share plains'	
Table 2 (3)	+7.7%	+0.95%
	Effect of one std. dev. increase in historic population density	
Table 2 (3)	+44.3%	+5.5%
	Effect of increase of share developed land by one standard deviation (+12.2 / +10.4 percentage points) on regulatory restrictiveness	
	Change in regulatory index	Change in rank order
Table 2 (1) (OLS)	+0.24	47 → 37
Table 2 (6) (TSLS)	+0.26	47 → 36
Table 2 (7) (LIML)	+0.27	47 → 36
Table 4 (4) (TSLS)	+0.35	47 → 33
Table 4 (5) (LIML)	+0.38	47 → 31
Table 7A (6) (TSLS)	+0.67	41 → 21
Table 7A (7) (LIML)	+0.71	41 → 20
	Effect of an increase in the homeownership rate by one standard deviation (+7.0 / +7.4 percentage points) on regulatory restrictiveness	
	Change in regulatory index	Change in rank order
Table 2 (1) (OLS)	(+0.042) [†]	(47 → 46) [†]
Table 2 (6) (TSLS)	(+0.26) [†]	(47 → 36) [†]
Table 2 (7) (LIML)	(+0.30) [††]	(47 → 35) [††]
Table 4 (4) (TSLS)	(+0.043) [†]	(47 → 46) [†]
Table 4 (5) (LIML)	(-0.013) [†]	(47 - unchanged) [†]
Table 7A (6) (TSLS)	(-0.23) [†]	(41 → 50) [†]
Table 7A (7) (LIML)	(-0.23) [†]	(41 → 50) [†]
	Effect of increase in predicted regulatory index (SAKS) by one standard deviation (+1.0) on the growth rate of housing supply	
	Change in the growth rate of housing supply between 1990 and 2000 (Change in percentage points)	Change in rank order
Table 7B (6)	-2.8%	41 → 54
Table 7B (7)	-2.7%	41 → 54

Notes: The marginal effects are measured at the means of the independent variables. The marginal effect of having major access to a coast is measured for an MSA that does not have major access to a coast compared to one that has access. The change in rank order is calculated for the MSA with the median regulatory index and the median growth rate, respectively. [†] Effect is not statistically significant. [††] Effect is only marginally statistically significant.

Appendix C.   Guide to calculations not intended for publication

C.1 Guide to calculations

It is useful to totally differentiate (7) for further reference. Using the definition $f_j \equiv H_j / H$ yields

$$\mathrm{d}f_j = f_j \left(1 - f_j\right) \frac{1}{\omega_j - t_j} \left(\mathrm{d}\omega_j - \mathrm{d}t_j\right). \qquad (23)$$

Derivation of the SPE (12). Using (23) to substitute for $\mathrm{d}H_j^0 / \mathrm{d}t_j^0$ in (10) yields

$$
\begin{aligned}
0 &= \frac{\mathrm{d}R_j^0}{\mathrm{d}t_j^0} \\
&\equiv \left(\tau H_j^0 + t_j^0\right) H \frac{\mathrm{d}f_j^0}{\mathrm{d}t_j^0} + H_j^0 \\
&= \left(\tau H_j^0 + t_j^0\right) f_j^0 H \left[\frac{-1}{\omega_j - t_j^0} + \frac{1}{J}\frac{1}{\bar{\omega} - \bar{t}^0}\right] + f_j^0 H \\
&= f_j^0 H \left\{\left(f_j^0 H \tau + t_j^0\right) \frac{1}{\omega_j - t_j^0}\left(-1 + f_j^0\right) + 1\right\},
\end{aligned}
$$

where the second line follows from (10) and (7), the third one follows from (7) and (11) and the final one follows from rearranging terms and using (11). For cities that are non-empty in equilibrium, this implies that the term in the curly bracket is equal to zero. This fact and imposing (14) together enable us to write

$$\omega_j - t_j^0 = \left(1 - f_j^0\right)\left(t_j^0 + f_j^0 H \tau\right). \qquad (24)$$

Isolating t_j^0 yields (12), as was to be shown. In (12), $\kappa_j^0 \in \left(\frac{1}{2}, 1\right)$ holds by $f_j^0 \in (0,1)$. Imposing (14) ensures that $f_j^0 > 0$ for all j; relaxing it would imply $f_j^0 = V_j^0 = 0$ for some j.

Establishing (19). Using (23) with $\mathrm{d}\omega_j = 0$ and using $\mathrm{d}H_j = H\mathrm{d}f_j$ by definition of f_j yields

$$2 + \tau \frac{\mathrm{d}H_j}{\mathrm{d}t_j} = 2 + \frac{H_j \tau}{\omega_j - t_j}\left(-1 + f_j\right)$$

$$= \frac{2\left(\omega_j - t_j\right) - H_j \tau \left(1 - f_j\right)}{\omega_j - t_j}$$

$$= \frac{2\left(\omega_j - t_j - H_j \tau\right) + H_j \tau \left(1 + f_j\right)}{\omega_j - t_j}$$

$$= \frac{2V_j + H_j \tau \left(1 + f_j\right)}{\omega_j - t_j} > 0,$$

where the fourth equality follows from (4). The inequality follows from $V_j > 0$ and establishes the first part of (19). We now show that $H_j(t_j)$ is concave; differentiating (23) one more time yields:

$$\frac{\mathrm{d}^2 f_j}{\mathrm{d}t_j^2} = -\frac{2\left[J(\overline{\omega} - \overline{t}) - \left(\omega_j - t_j\right)\right]}{\left[J(\overline{\omega} - \overline{t})\right]^3}$$

$$= -\frac{2\sum_{k \neq j}\left(\omega_k - t_k\right)}{\left[\sum_{k \in \Im+}\left(\omega_k - t_k\right)\right]^3} < 0.$$

This implies the second part of (19) by $\mathrm{d}H_j = H\mathrm{d}f_j$, as was to be shown.

Establishing (20) **and** (21). Differentiating (24) yields

$$\mathrm{d}\omega_j - \mathrm{d}t_j^0 = -\mathrm{d}f_j^0 \left(t_j^0 + f_j^0 H\tau\right) + \left(1 - f_j^0\right)\left(\mathrm{d}t_j^0 + H\tau \mathrm{d}f_j^0\right)$$

$$= \left(1 - f_j^0\right)\mathrm{d}t_j^0 + \left[-\left(t_j^0 + f_j^0 H\tau\right) + \left(1 - f_j^0\right)H\tau\right]\mathrm{d}f_j^0$$

$$= \left(1 - f_j^0\right)\mathrm{d}t_j^0 + \left[-\left(t_j^0 + f_j^0 H\tau\right) + \left(1 - f_j^0\right)H\tau\right]f_j^0\left(1 - f_j^0\right)\frac{1}{\omega_j - t_j}\left(\mathrm{d}\omega_j - \mathrm{d}t_j^0\right),$$

where the third equality follows from (23). Isolating the term $\left(\mathrm{d}\omega_j - \mathrm{d}t_j^0\right)$ and rearranging yield:

$$\left(\omega_j - t_j^0\right)\left(1 - f_j^0\right)\mathrm{d}t_j^0 = \left\{\left(\omega_j - t_j^0\right) + \left[\left(t_j^0 + f_j^0 H\tau\right) - \left(1 - f_j^0\right)H\tau\right]f_j^0\left(1 - f_j^0\right)\right\}\left(\mathrm{d}\omega_j - \mathrm{d}t_j^0\right).$$

In turn, using the definition of V_j in (4) and rearranging yield:

$$\left(V_j^0 + f_j^0 H\tau\right)\left(1 - f_j^0\right)\mathrm{d}t_j^0 = \left\{\left(V_j^0 + f_j^0 H\tau\right) + \left[t_j^0 + \left(2f_j^0 - 1\right)H\tau\right]f_j^0\left(1 - f_j^0\right)\right\}\left(\mathrm{d}\omega_j - \mathrm{d}t_j^0\right)$$

$$= \left\{V_j^0 + f_j^0\left(1 - f_j^0\right)t_j^0 + \left[1 + \left(2f_j^0 - 1\right)\left(1 - f_j^0\right)\right]f_j^0 H\tau\right\}\left(\mathrm{d}\omega_j - \mathrm{d}t_j^0\right)$$

$$= \left[V_j^0 + f_j^0\left(1 - f_j^0\right)t_j^0 + \left(3 - 2f_j^0\right)\left(f_j^0\right)^2 H\tau\right]\left(\mathrm{d}\omega_j - \mathrm{d}t_j^0\right),$$

which we rewrite as $l_j \mathrm{d}t_j^0 = r_j \left(\mathrm{d}\omega_j - \mathrm{d}t_j^0 \right)$, where l_j and r_j are defined in (20). Both terms are positive by inspection. This implies (21) and, together with (23):

$$\frac{\mathrm{d}f_j^0}{\mathrm{d}\omega_j} = f_j^0 \left(1 - f_j^0 \right) \frac{1}{\omega_j - t_j^0} \frac{l_j}{r_j} > 0,$$

which is another way of getting property (i).

Establishing (22). Using (12) for the least desirable jurisdiction yields

$$t_{\min}^0 = \kappa_{\min}^0 \omega_{\min} - \left(1 - \kappa_{\min}^0 \right) f_{\min}^0 H \tau$$
$$> \frac{\omega_{\min} - J^{-1} H \tau}{2},$$

which is positive if (and only if) $\omega_{\min} > H\tau / J$; the inequality follows from $f_{\min}^0 < 1/J$ and $\kappa_{\min}^0 > 1/2$. By the same token, applying (13) to the least desirable jurisdiction yields

$$V_{\min}^0 = \left(1 - \kappa_{\min}^0 \right) \omega_{\min} - \kappa_{\min}^0 f_{\min}^0 H \tau$$
$$> \frac{\left(1 - J^{-1} \right) \omega_{\min} - J^{-1} H \tau}{2 - f_{\min}^0},$$

which is positive if (and only if) $\omega_{\min} > H\tau / (J-1)$; the inequality follows from $f_{\min}^0 < 1/J$. As a result, both $t_j^0 > 0$ and $V_j^0 > 0$ hold for all j if (14) holds, as was to be shown.

C.2 Adding agglomeration economies

We claim in the text that adding agglomeration economies or external congestion costs do not affect the essence of our model and empirical strategy. To see this, let local amenities and local wages be a function of local population and a location-specific shifter, which we write $a_j = a\left(H_j, \tilde{a}_j \right)$ and $w_j = w\left(H_j, \tilde{w}_j \right)$ so that $\omega_j = \omega\left(H_j, \tilde{\omega}_j \right)$. Without loss of generality we let $a_2(\cdot) > 0$ and $w_2(\cdot) > 0$ (subscripts denote the variable with respect to which the function is differentiated). It is reasonable to assume $w_1(\cdot) > 0$ (agglomeration economies) and $a_1(\cdot) < 0$ if people prefer to live in places that are not too dense on average. Thus the term $\omega_1(\cdot)$ can be positive or negative a priori. Equation (7) still defines a spatial equilibrium in this context, though it is now an implicit equation. Its solution exists and is locally stable if the RHS of (6) is decreasing, which is the case iff $\omega_1(H_j, \cdot) < \tau$. Imposing $w_1(\cdot) + a_1(\cdot) < \tau$ for all j and all

$H_j \in \mathbb{R}_+$ yields uniqueness. Turning to the regulation game, the first order condition for of the program $\max_{t_j} R_j(\mathbf{t})$ reads (omitting complementary slackness for simplicity)

$$\frac{\mathrm{d}}{\mathrm{d}t_j} R_j(t_j, \mathbf{t_{-j}^0})\bigg|_{t_j = t_j^0} = H_j^0 + \left(H_j^0 \tau + t_j^0\right) \frac{\mathrm{d}}{\mathrm{d}t_j} H_j(t_j, \mathbf{t_{-j}^0})\bigg|_{t_j = t_j^0} = 0,$$

where

$$\frac{\mathrm{d}}{\mathrm{d}t_j} H_j(t_j, \mathbf{t_{-j}}) = -\frac{1}{\bar{\omega} - \bar{t} - \bar{H}\tau}\left(1 - \frac{H_j}{H}\right)\left[1 + \frac{\tau - \omega_1\left(H_j, \tilde{w}_j, \tilde{a}_j\right)}{\bar{\omega} - \bar{t} - \bar{H}\tau}\left(1 - \frac{H_j}{H}\right)\right]^{-1} < 0.$$

Thus, the size of a jurisdiction still decreases with respect to its own regulation in equilibrium. By the same token, it is easy to show that more desirable locations are more populated in the spatial equilibrium by

$$\mathrm{d}H_j(\tilde{w}_j, \tilde{a}_j; \cdot) = \frac{1}{\bar{\omega} - \bar{t} - \bar{H}\tau}\left(1 - \frac{H_j}{H}\right)\left[1 + \frac{\tau - \omega_1\left(H_j, \tilde{w}_j, \tilde{a}_j\right)}{\bar{\omega} - \bar{t} - \bar{H}\tau}\left(1 - \frac{H_j}{H}\right)\right]^{-1}$$

$$\times \left[\omega_2\left(H_j, \tilde{w}_j, \tilde{a}_j\right)\mathrm{d}\tilde{w}_j + \omega_3\left(H_j, \tilde{w}_j, \tilde{a}_j\right)\mathrm{d}\tilde{a}_j\right].$$

Likewise, the rest of the analysis goes through unaltered (adding some regularity conditions).