Institut d'Economia de Barcelona

Document de treball de l'IEB 2010/33

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

Christian A. L. Hilber, Frédéric Robert-Nicoud

Cities and Innovation

Document de treball de l'IEB

Documents de Treball de l’IEB 2010/33

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

Christian A. L. Hilber, Frédéric Robert-Nicoud

The **IEB** research program in **Cities and Innovation** aims at promoting research in the Economics of Cities and Regions. The main objective of this program is to contribute to a better understanding of agglomeration economies and 'knowledge spillovers'. The effects of agglomeration economies and 'knowledge spillovers' on the Location of economic Activities, Innovation, the Labor Market and the Role of Universities in the transfer of Knowledge and Human Capital are particularly relevant to the program. The effects of Public Policy on the Economics of Cities are also considered to be of interest. This program puts special emphasis on applied research and on work that sheds light on policy-design issues. Research that is particularly policy-relevant from a Spanish perspective is given special consideration. Disseminating research findings to a broader audience is also an aim of the program. The program enjoys the support from the **IEB-Foundation**.

The **Barcelona Institute of Economics (IEB)** is a research centre at the University of Barcelona which specializes in the field of applied economics. Through the **IEB-Foundation**, several private institutions (Caixa Catalunya, Abertis, La Caixa, Gas Natural and Applus) support several research programs.

Postal Address:
Institut d’Economia de Barcelona
Facultat d’Economia i Empresa
Universitat de Barcelona
C/ Tinent Coronel Valenzuela, 1-11
(08034) Barcelona, Spain
Tel.: + 34 93 403 46 46
Fax: + 34 93 403 98 32
ieb@ub.edu
http://www.ieb.ub.edu

The IEB working papers represent ongoing research that is circulated to encourage discussion and has not undergone a peer review process. Any opinions expressed here are those of the author(s) and not those of IEB.

Documents de Treball de l'IEB 2010/33

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

Christian A. L. Hilber, Frédéric Robert-Nicoud

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.

JEL Codes: H7, Q15, R52

Keywords: Land use regulations, zoning, land ownership, housing supply.

Christian A. L. Hilber
Spatial Economics Research Centre (SERC)
London School of Economics
Houghton Street
London, WC2A 2AE
E-mail: c.hilber@lse.ac.uk

Frédéric Robert-Nicoud
Economics Department
University of Geneva
E-mail: frederic.robert-nicoud@unige.ch

[*] This paper replaces and extends our 2006 unpublished paper (Hilber and Robert-Nicoud 2006) in a variety of directions. We are especially grateful to Gilles Duranton and Henry Overman for thorough discussions, detailed feedback and encouragement. For discussions and suggestions, we also thank Alex Anas, John Antonakis, Jan Brueckner, Marius Brülhart, Jean Cavailhès, Paul Cheshire, Steve Gibbons, Yoshitsugu Kanemoto, Gerrit Knaap, Rafael Lalive, Tim Leunig, Jim Markusen, Florian Mayneris, Muriel Meunier, Mark Partridge, Esteban Rossi-Hansberg, Michael Storper, Will Strange, Jacques Thisse, Wouter Vermeulen and conference/seminar participants at the Erasmus University of Rotterdam, ERES annual conference (Krakow), HEC Lausanne, IEB workshop on urban economics (Barcelona), INRA, LSE, Princeton, and RSAI annual conference (Toronto). We are grateful to Raven Saks for kindly providing regulatory index data. Alejandra Castrodad-Rodriguez, Ralph Van Doorn, Felix Weinhardt and especially Ben Dachis provided excellent research assistance. The remaining errors are the sole responsibility of the authors. Part of this paper was written when Frédéric Robert-Nicoud was a Peter B. Kenen visiting scholar at Princeton University; the hospitality of this institution is gratefully acknowledged.

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. 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. *We conclude that the current pattern of land use regulations is suboptimal.*

The paper also sheds light on the evolution and 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.

---

[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.

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, first, a causal link from the increasing urbanization rates (a demographic pattern) to the spread of residential land use regulations (a political pattern) and, second, 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 ρ = .31 is statistically larger than zero. Figure 1 (panel b) suggests that this pattern was already visible in the data in the late

1970s (ρ = .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, ρ = .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. Third, out-of-sample evidence 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 very meaningful. The implications of this set of results can hardly be overstated: *the current*

---

[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.

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. 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

[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).

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' (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

---

[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.

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 and is unique 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 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

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

The common component V_j is deterministic and summarizes the costs and benefits from living in jurisdiction j, expressed in monetary units; think about it as the indirect 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 are heterogeneous in their appreciation of these amenities.[8] In order to get simple, explicit solutions, we assume that the ε's are uniformly distributed over the interval $[-\sigma/2, \sigma/2]$, where σ is proportional to the standard deviation of G. The mean and mode of G are zero, meaning that the average and median households are a-priory indifferent about where to live.[9] As a result of this and (2), the location choice probabilities in (1) are equivalent to

---

[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.

[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] Specifically, $\text{Var}(\varepsilon) = \sigma^2/12$ and $G(\varepsilon) = (2\varepsilon + \sigma)/(2\sigma)$. The deterministic case obtains at the limit $\sigma = 0$. The reader may be more familiar to modeling discrete location decisions with the multinomial logit model. We show in Appendix C (not intended for publication) that the equilibrium properties of the multinomial logit are qualitatively similar to the linear case. We choose to work with the linear model because we obtain simple, closed form solutions.

$$f_j = \frac{1}{J}\left[1 + \frac{V_j - \bar{V}}{\sigma}\right] \tag{3}$$

with f_j = 0 or f_j = 1 in an obvious manner if the 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 indirect utility V_j attract more households than the average jurisdiction. The degree of household heterogeneity σ governs the sensitivity of f_j with respect to the utility differential: heterogeneous populations are less sensitive to differences in the common components $V_1, ..., V_J$.

We assume that the common 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 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]

In this paper, the land market outcomes play the central role, so we treat a_j and w as parameters but we endogenize the cost of living as follows. 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 $\tau H_j + m_j$.[11] Substituting this expression for c_j in (4) yields:

There is a tradition in the Political Economics literature that uses linear discrete choice models to model voters' decisions to cast their ballot for a candidate or another (Persson and Tabellini, 2000).

[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 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

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

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 congestion 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 $\tau\mathbf{H}$ 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. Defining $\bar{H} \equiv H / J$ and using (5) and (3), we may define the spatial equilibrium as

$$\forall j \in \Im: \qquad f_j = \frac{H_j}{H} \quad \Leftrightarrow \quad \frac{H_j}{\bar{H}} = 1 + \frac{\left(\omega_j - t_j - \tau H_j\right) - \left(\bar{\omega} - \bar{t} - \tau \bar{H}\right)}{\sigma}. \tag{6}$$

That is, the fraction of people living in j is increasing in the local well-being net of the regulatory tax and congestion costs and decreasing in the well-being net of regulatory tax and congestion 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:

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

---

conversion costs m_j) is $r(d) = \tau (H_j - d)$. As a result, the cost of living is $c_j = \tau H_j + m_j$, the aggregate land rent is equal to $\tau H^2/2$ and $\omega_j = a_j + w - m_j$, as in (5).

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 :

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

Together with (5), (7) implies

$$V_j = \left(1 - \kappa_0\right)\left(\omega_j - t_j\right) + \kappa_0\left(\overline{\omega} - \overline{t}\right) - \tau \overline{H}, \qquad \kappa_0 \equiv \frac{\tau \overline{H}}{\sigma + \tau \overline{H}}. \tag{8}$$

At the spatial equilibrium, households obtain a real income that is a weighted average of the local fundamental desirability net of regulatory taxes (the first term in the RHS above) and of the average of the same object (the second term), minus the average congestion cost. The weight of *j*-specific characteristics (relative to the average) increases in σ, our measure of the intensity of idiosyncratic household preferences for local specificities. Several aspects of (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, the effect of either variable on the spatial outcome is decreasing in the heterogeneity of households σ and in the overall crowding of jurisdictions $\tau \overline{H}$. Third, *all jurisdictions yield about the same welfare ex post*: congestion and labor mobility between jurisdictions together ensure that in 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 the difference of the real incomes at the spatial equilibrium for two arbitrary jurisdictions *j* and *k*; using (8), we obtain $V_j - V_k = (1-\kappa_0)\left[(\omega_j - \omega_k) - (t_j - t_k)\right]$. At the limit $\sigma \to 0$ (homogeneous population), all populated jurisdictions yield *exactly* the same welfare. Otherwise, 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 enjoy the most.

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.[12] 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 $\Lambda \in \{\text{owners of developed land, land developers}\}$. This objective function conveys the idea that the planning board caters only to the interests of land stakeholders. 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.[13] Thus, the

---

[12] 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).

[13] 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.

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.[14] As a result of these assumptions, the planning board maximizes total land rents plus the regulatory tax revenue:

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

where $H_j(\mathbf{t})$ is given by (7). Two aspects of the program $\max_{t_j} R_j\left(\mathbf{t}\right)$ 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 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 $t_j \in \Re_+$ so as to maximize (9) subject to (7) taking the vector $\mathbf{t^0_{-j}} \equiv \{t_k^0\}_{k \in \Im \setminus \{j\}}$ as given (the superscript '0' pertains to equilibrium values). Then the first order condition for this program may be written as

$$\left.\frac{\partial}{\partial t_j} R_j(t_j, \mathbf{t^0_{-j}})\right|_{t_j = t_j^0} = H_j^0 + \left(\tau H_j^0 + t_j^0\right)\left.\frac{\partial}{\partial t_j} H_j(t_j, \mathbf{t^0_{-j}})\right|_{t_j = t_j^0} \le 0, \qquad t_j^0 \ge 0, \tag{10}$$

with complementary slackness; H_j^0 is the equilibrium jurisdiction size, namely (7) evaluated at the Nash tax vector $\mathbf{t^0} \equiv [t_1^0, \dots, t_J^0]'$ (also $\mathbf{H^0} \equiv [H_1^0, \dots, H_J^0]'$). The RHS consists of respectively the direct effect (keeping the population constant) and the indirect effect (allowing it to vary in

[14] 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.

response) of an increase in the specific regulatory tax. At an interior equilibrium, the cost and benefits of raising t_j are equal. Let

$$\kappa \equiv \frac{J-1}{J}\kappa_0 = \frac{J-1}{J}\frac{\tau\overline{H}}{\sigma+\tau\overline{H}} \tag{11}$$

be a measure of the toughness of fiscal competition among jurisdictions: κ is increasing in the number of jurisdictions J and decreasing in the heterogeneity of households σ. Then, using (7) and (11) and assuming that the parameters of the model are such that the resulting H_j^0 's in (10) are all interior (precise conditions to follow in (16)), we may develop the first order condition (10) to get a non-negative equilibrium relationship between population size H_j^0 and the regulation tax t_j^0 :

$$t_j^0 = \frac{1-\kappa}{\kappa}\tau H_j^0 . \tag{12}$$

Together, (7) and (12) compose a system of 2J equations in $\mathbf{H^0}$ and $\mathbf{t^0}$. Solving for individual regulatory taxes and populations yields:

$$t_j^0 = \frac{1-\kappa}{\kappa}\tau\overline{H} + \frac{1-\kappa}{2-\kappa-\frac{1}{J}}(\omega_j - \overline{\omega}) , \tag{13}$$

where the coefficient of $(\omega_j - \overline{\omega})$ is strictly positive and smaller than unity, and

$$H_j^0 = \overline{H} + \frac{1}{\tau}\frac{\kappa}{2-\kappa-\frac{1}{J}}(\omega_j - \overline{\omega}) , \tag{14}$$

$j = 1,...,J$. All the properties of the spatial equilibrium continue to hold at the SPE that (13) and (14) characterize. Four additional properties resulting from strategic interactions are noteworthy. First, the equilibrium regulatory tax increases in own desirability and decreases in the desirability of other jurisdictions; this effect is stronger, the higher the heterogeneity of households σ and the lower the congestion costs $\tau\overline{H}$ are. The former cross-effect arises because a homogenous population is more responsive to any differences in amenities and taxes across locations; the latter arises because, when congestion costs are large, cross-jurisdiction differences along other dimensions matter relatively less for households' location choices. Second, 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; this follows by inspection of (13) and from $\kappa < 1$. Third, equilibrium tax rates are decreasing in the number of jurisdictions and increasing in the heterogeneity of workers, ceteris paribus. This can be interpreted as a pro-competitive effect: as σ rises, each jurisdiction perceives a *lower* land-demand elasticity from households. Finally, the average regulatory tax is increasing in the size of the population, keeping the number of jurisdictions constant: as existing jurisdictions become increasingly crowded and developed, the political balance tilts in favor of the owners of developed land.

The second derivative of R_j with respect to t_j is invariant and negative for *any* **t**:

$$0 > \frac{\partial^2}{\partial t_j^2} R_j(\cdot) = -\frac{\kappa(2-\kappa)}{\tau} . \tag{15}$$

So we may write:

Proposition 2 (existence and uniqueness of the SPE tax setting game). The Subgame perfect equilibrium characterized by (13) and (14) exists and is unique.

Proof. See Appendix A.

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

Proposition 3 (properties of the subgame perfect equilibrium). Assume:

$$\forall j \in \Im: \qquad \bar{\omega} - \frac{2-\kappa-1/J}{\kappa}\tau\bar{H} < \omega_j < \bar{\omega} + (J-1)\frac{2-\kappa-1/J}{\kappa}\tau\bar{H} . \tag{16}$$

Then the SPE characteristics of the model, summarized in (13) and (14), imply:

(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 (16) 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 indirect utility: $\omega_j > \omega_k \Rightarrow V_j^0 > V_k^0$.

Proof. See Appendix A.

Our model can also shed light on the demographic origins of land use regulations:

Corollary 3.2 (urbanization and land use regulations). The overall extent of regulation increases with the size of the urban population in the economy; furthermore, all jurisdictions are affected:

(vi) The *average* regulatory tax is increasing in H and in $\bar{H}$: $\partial \bar{t}^0 / \partial H > 0$ and $\partial \bar{t}^0 / \partial \bar{H} > 0$;

(vii) The *variance* of the regulatory tax is decreasing in the size of the urban population: $\partial \operatorname{Var}(t^0) / \partial H < 0$.

Proof. See Appendix A.

The parameter restriction (16) requires the variation in the desirable fundamentals to be bounded above relative to household heterogeneity and congestion costs – not a very stringent condition – and in doing so ensures that the equilibrium population in any *j* is positive for the interior solution $\mathbf{t}^0$. 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 and uniqueness 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 (15) 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 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. Fortunately, a simple, single step is required to match the two. 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 hold across MSAs. To see this, use the definition for MSA averages above into (12) so as to rewrite the development-regulation relationship as

$$t_m^0 = \frac{1-\kappa}{\kappa} \tau H_m^0$$

and into (14) so as to rewrite the amenity-development relationship as

$$H_m^0 = \bar{H} + \frac{1}{\tau} \frac{\kappa}{2 - \kappa - \frac{1}{J}} (\omega_m - \bar{\omega}) .$$

The former expression implies $\partial t_m^0 / \partial H_m^0 > 0$; the latter implies $\partial H_m^0 / \partial \omega_m^0 > 0$. Thus, we have shown that Proposition 3(i) and Proposition 3(ii) also hold across MSAs.

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.

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.[15] 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

$$SDL = \frac{developed\ residential\ land\ area}{developable\ residential\ land\ area}, \tag{17}$$

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

[15] 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.

support permanent structures and/or is extremely costly to develop).[16] *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 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.[17] 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

---

[16] 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.

[17] 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.

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 (10): 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 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:

$$WRLURI_m = \beta_0 + \beta_1\left(\boldsymbol{SDL}_m\right) + \beta_2\left(\boldsymbol{HOR}_m\right) + \beta_3\left(POPD_m\right) + \beta_4\left(controls_m\right) + \varepsilon_m, \qquad (18)$$

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 (18) 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.[18] 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.[19] 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.[20] *Region dummies* reveal that broad geographic patterns emerge, 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 (18) 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 (18) 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

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

[19] 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.

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

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 (17) 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 (17).

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:

$$\begin{aligned} SDL_j = \alpha_0 + \alpha_1\left(coast_j\right) + \alpha_2\left(temperature_j\right) + \alpha_3\left(share\ plains_j\right) \\ +\alpha_4\left(historical\ density_j\right) + \alpha_5\left(controls_j\right) + \xi_j\ , \end{aligned} \tag{19}$$

where ξ_j is the error term. Our priors are α_1 ,α_2 , α_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:[21] granting a *border with coast* to a

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

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 (18) 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. It is larger than the OLS coefficient of *SDL*, confirming 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-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 (18) 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 (18) 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 19^{th} 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 (18) 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.[22]

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

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

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 (18) 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 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*.[23] 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.

---

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

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.

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 e.g. (7). To check whether this is a feature of the regulatory index in our data, we run the following with OLS:

$$g^H_{j,t} = \gamma_0 + \gamma_1(regulation^e_{j,t}) + \gamma_2(g^H_{j,t-1}) + \zeta_{j,t} \tag{20}$$

where $g^H_{j,t}$ is the growth rate of the housing stock in MSA j between time period t and t+1 (in number of housing units), $g^H_{j,t-1}$ is the equivalent growth rate between time period t-1 and t, γ_0 is the common trend, $regulation^e_{j,t}$ is the estimated level of regulatory restrictiveness in MSA j at time $t,$ and $\zeta_{j,t}$ is the error term. We include $g^H_{j,t-1}$ 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 $regulation^e_{j,t} = SAKS^e_{j,t}$ from the estimations of Table 7, Panel A. We report the regression results for (20) 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.[24] 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 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 (20) 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 impose – via increasing housing costs – an enormous gross cost on households. 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, 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

[24] 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).

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 Hypothesi*s. *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.

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.

Table 2
Base specification: Determinants of restrictiveness of land use regulations *(N=93)*

	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)*

	First-stage			Second-stage		
				TSLS	LIML	JIVE
Dependent variable:	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.)

2009

2009/1. Rork, J.C.; Wagner, G.A.: "Reciprocity and competition: is there a connection?"
2009/2. Mork, E.; Sjögren, A.; Svaleryd, H.: "Cheaper child care, more children"
2009/3. Rodden, J.: "Federalism and inter-regional redistribution"
2009/4. Ruggeri, G.C.: "Regional fiscal flows: measurement tools"
2009/5. Wrede, M.: "Agglomeration, tax competition, and fiscal equalization"
2009/6. Jametti, M.; von Ungern-Sternberg, T.: "Risk selection in natural disaster insurance"
2009/7. Solé-Ollé, A; Sorribas-Navarro, P.: "The dynamic adjustment of local government budgets: does Spain behave differently?"
2009/8. Sanromá, E.; Ramos, R.; Simón, H.: "Immigration wages in the Spanish Labour Market: Does the origin of human capital matter?"
2009/9. Mohnen, P.; Lokshin, B.: "What does it take for and R&D incentive policy to be effective?"
2009/10. Solé-Ollé, A.; Salinas, P..: "Evaluating the effects of decentralization on educational outcomes in Spain?"
2009/11. Libman, A.; Feld, L.P.: "Strategic Tax Collection and Fiscal Decentralization: The case of Russia"
2009/12. Falck, O.; Fritsch, M.; Heblich, S.: "Bohemians, human capital, and regional economic growth"
2009/13. Barrio-Castro, T.; García-Quevedo, J.: "The determinants of university patenting: do incentives matter?"
2009/14. Schmidheiny, K.; Brülhart, M.: "On the equivalence of location choice models: conditional logit, nested logit and poisson"
2009/15. Itaya, J., Okamuraz, M., Yamaguchix, C.: "Partial tax coordination in a repeated game setting"
2009/16. Ens, P.: "Tax competition and equalization: the impact of voluntary cooperation on the efficiency goal"
2009/17. Geys, B., Revelli, F.: "Decentralization, competition and the local tax mix: evidence from Flanders"
2009/18. Konrad, K., Kovenock, D.: "Competition for fdi with vintage investment and agglomeration advantages"
2009/19. Loretz, S., Moorey, P.: "Corporate tax competition between firms"
2009/20. Akai, N., Sato, M.: "Soft budgets and local borrowing regulation in a dynamic decentralized leadership model with saving and free mobility"
2009/21. Buzzacchi, L., Turati, G.: "Collective risks in local administrations: can a private insurer be better than a public mutual fund?"
2009/22. Jarkko, H.: "Voluntary pension savings: the effects of the finnish tax reform on savers' behaviour"
2009/23. Fehr, H.; Kindermann, F.: "Pension funding and individual accounts in economies with life-cyclers and myopes"
2009/24. Esteller-Moré, A.; Rizzo, L.: "(Uncontrolled) Aggregate shocks or vertical tax interdependence? Evidence from gasoline and cigarettes"
2009/25. Goodspeed, T.; Haughwout, A.: "On the optimal design of disaster insurance in a federation"
2009/26. Porto, E.; Revelli, F.: "Central command, local hazard and the race to the top"
2009/27. Piolatto, A.: "Plurality versus proportional electoral rule: study of voters' representativeness"
2009/28. Roeder, K.: "Optimal taxes and pensions in a society with myopic agents"
2009/29, Porcelli, F.: "Effects of fiscal decentralisation and electoral accountability on government efficiency evidence from the Italian health care sector"
2009/30, Troumpounis, O.: "Suggesting an alternative electoral proportional system. Blank votes count"
2009/31, Mejer, M., Pottelsberghe de la Potterie, B.: "Economic incongruities in the European patent system"
2009/32, Solé-Ollé, A.: "Inter-regional redistribution through infrastructure investment: tactical or programmatic?"
2009/33, Joanis, M.: "Sharing the blame? Local electoral accountability and centralized school finance in California"
2009/34, Parcero, O.J.: "Optimal country's policy towards multinationals when local regions can choose between firm-specific and non-firm-specific policies"
2009/35, Cordero, J,M.; Pedraja, F.; Salinas, J.: "Efficiency measurement in the Spanish cadastral units through DEA"
2009/36, Fiva, J.; Natvik, G.J.: "Do re-election probabilities influence public investment?"
2009/37, Haupt, A.; Krieger, T.: "The role of mobility in tax and subsidy competition"
2009/38, Viladecans-Marsal, E; Arauzo-Carod, J.M.: "Can a knowledge-based cluster be created? The case of the Barcelona 22@district"

2010

2010/1, De Borger, B., Pauwels, W.: "A Nash bargaining solution to models of tax and investment competition: tolls and investment in serial transport corridors"
2010/2, Chirinko, R.; Wilson, D.: "Can Lower Tax Rates Be Bought? Business Rent-Seeking And Tax Competition Among U.S. States"
2010/3, Esteller-Moré, A.; Rizzo, L.: "Politics or mobility? Evidence from us excise taxation"
2010/4, Roehrs, S.; Stadelmann, D.: "Mobility and local income redistribution"

2010/5, Fernández Llera, R.; García Valiñas, M.A.: "Efficiency and elusion: both sides of public enterprises in Spain"
2010/6, González Alegre, J.: "Fiscal decentralization and intergovernmental grants: the European regional policy and Spanish autonomous regions"
2010/7, Jametti, M.; Joanis, M.: "Determinants of fiscal decentralization: political economy aspects"
2010/8, Esteller-Moré, A.; Galmarini, U.; Rizzo, L.: "Should tax bases overlap in a federation with lobbying?"
2010/9, Cubel, M.: "Fiscal equalization and political conflict"
2010/10, Di Paolo, A.; Raymond, J.L.; Calero, J.: "Exploring educational mobility in Europe"
2010/11, Aidt, T.S.; Dutta, J.: "Fiscal federalism and electoral accountability"
2010/12, Arqué Castells, P.: "Venture capital and innovation at the firm level"
2010/13, García-Quevedo, J.; Mas-Verdú, F.; Polo-Otero, J.: "Which firms want PhDS? The effect of the university-industry relationship on the PhD labour market"
2010/14, Calabrese, S.; Epple, D.: "On the political economy of tax limits"
2010/15, Jofre-Monseny, J.: "Is agglomeration taxable?"
2010/16, Dragu, T.; Rodden, J.: "Representation and regional redistribution in federations"
2010/17, Borck, R; Wimbersky, M.: "Political economics of higher education finance"
2010/18, Dohse, D; Walter, S.G.: "The role of entrepreneurship education and regional context in forming entrepreneurial intentions"
2010/19, Åslund, O.; Edin, P-A.; Fredriksson, P.; Grönqvist, H.: "Peers, neighborhoods and immigrant student achievement - Evidence from a placement policy"
2010/20, Pelegrín, A.; Bolance, C.: "International industry migration and firm characteristics: some evidence from the analysis of firm data"
2010/21, Koh, H.; Riedel, N.: "Do governments tax agglomeration rents?"
2010/22, Curto-Grau, M.; Herranz-Loncán, A.; Solé-Ollé, A.: "The political economy of infraestructure construction: The Spanish "Parliamentary Roads" (1880-1914)"
2010/23, Bosch, N.; Espasa, M.; Mora, T.: "Citizens' control and the efficiency of local public services"
2010/24, Ahamdanech-Zarco, I.; García-Pérez, C.; Simón, H.: "Wage inequality in Spain: A regional perspective"
2010/25, Folke, O.: "Shades of brown and green: Party effects in proportional election systems"
2010/26, Falck, O.; Heblich, H.; Lameli, A.; Südekum, J.: "Dialects, cultural identity and economic exchange"
2010/27, Baum-Snow, N.; Pavan, R.: "Understanding the city size wage gap"
2010/28, Molloy, R.; Shan, H.: "The effect of gasoline prices on household location"
2010/29, Koethenbuerger, M.: "How do local governments decide on public policy in fiscal federalism? Tax vs. expenditure optimization"
2010/30, Abel, J.; Dey, I.; Gabe, T.: "Productivity and the density of human capital"
2010/31, Gerritse, M.: "Policy competition and agglomeration: a local government view"
2010/32, Hilber, C.; Lyytikäinen, T.; Vermeulen, W.: "Capitalization of central government grants into local house prices: panel data evidence from england"

ieb@ub.edu
www.ieb.ub.edu

Cities and Innovation

Document de treball de l'IEB