Land use regulation and welfare[§]

Andrew Haughwout [‡]
NYFRB

Matthew A. Turner[†]
University of Toronto

Wilbert van der Klaauw [*]
NYFRB

This draft: 20 July 2010

ABSTRACT: We estimate the effect of land use regulation on the value of land by exploiting variation in exposure to municipal land use regulation across municipal boundaries. Since the value of land gives us the market's measure of the attractiveness of a location, our estimates allow us to draw conclusions about the welfare implications of land use regulation. Our results suggest that marginal reductions in municipal land use regulation are generally welfare improving.

Key words: land regulation, zoning, urban economics, regulation.
JEL classification: R5,H2.

[§]Preliminary, comments welcome. We are grateful to Kailin Clarke for excellent research assistance, to Carmen Carrion-Flores, Gilles Duranton and Christian Redfern for helpful comments, and to seminar participants the Paris School of Economics and the University of California at Berkeley.

[‡]Research and Statistics Group Federal Reserve Bank of New York 33 Liberty Street New York, NY 10045. Andrew.Haughwout@ny.frb.org.

[†]Department of Economics, University of Toronto, 150 Saint George Street, Toronto, Ontario, Canada M5S 3G7. mturner@chass.utoronto.ca , http://www.economics.utoronto.ca/mturner/index.htm.

[*]Research and Statistics Group Federal Reserve Bank of New York 33 Liberty Street New York, NY 10045. Wilbert.VanDerKlaauw@ny.frb.org.

1. Introduction

We estimate the effect of land use regulation on the value of land by exploiting variation in exposure to municipal land use regulation across municipal boundaries. Since the value of land gives us the market's measure of the attractiveness of a location, our estimates allow us to draw conclusions about the effect of land use regulation on welfare.

Any estimate of the effect of regulation on the land market must confront the fact that many of the same factors that determine land use regulation also determine land price: a parcel on the beach is valuable for its proximity to the beach and may be occupied by people with a taste for stringent regulation to protect their views. Thus, a naive estimate of the relationship between land price and regulation will conclude that more highly regulated land is more valuable when there need not be a causal relationship between the two. The resolution of this inference problem is our main concern.

To begin, we observe that land use regulation should have two types of effects on land value. The first, an 'own-lot effect' causes the value of any parcel to decrease because any binding regulation constrains the owner from using the parcel in his most preferred way. The second, an 'external effect', may change the value of a parcel by constraining the owners of nearby parcels from taking harmful (or beneficial) actions. For example, binding minimum lot size regulation will decrease the value of a parcel by constraining the developer to use larger lots than would otherwise be optimal. It will also lead to low density neighborhoods which potential buyers may (or may not) prefer.

We estimate the own-lot and external effects of regulation separately. Loosely, we estimate the own-lot effect of regulation by comparing nearby parcels on opposite sides of a municipal boundary. If all other determinants of land value vary continuously across the border then the difference in prices across the border reflects the own-lot effect of regulation. We estimate the external effect of regulation by comparing interior and peripheral parcels in the same municipality. The peripheral parcels are 50% exposed to a landscape shaped by regulation in the neighboring municipality, and 50% to a landscape shaped by regulation in the home municipality. Interior parcels, on the other hand, are entirely exposed to a landscape shaped by regulation in the home municipality. Since both parcels experience the same own-lot effect of regulation, if the external benefits or costs of proximity to a regulated landscape decay quickly enough then the difference between the price of interior and peripheral parcels reflects the external effect of land use regulation on land value.

Our results are of interest for three main reasons. First, land is among the most important assets in the US economy. According to 2007 estimates, the aggregate market value of US land was about 22 trillion dollars. The market for land is also highly regulated. Therefore, understanding the impact of land use regulation on land value is an economic problem of the first order.

Second, the policy debate surrounding land use regulation attracts many competing interest groups with conflicting agendas. "[T]he Sierra Club urges planning and policies which stimulate: ... "Infill" residential and commercial development on unused or under-used land within city boundaries..."[1] The National Association of Home Builders, on the other hand opposes "urban growth boundaries, which restrict the amount of developable land and contribute to increased housing prices..."[2] Given the value of land and the durability of development, land use policy should be based on the careful analysis of high quality data, and not on interest group politics.

Third, while the intuition behind our estimations is simple, one can imagine confounding effects. For example, the sorting of different demographic groups across municipalities on the basis of factors that are correlated with, but not caused by land use regulation. We devote much of our attention to these issues. Much of the growing literature which exploits policy discontinuities at administrative borders confronts similar inference problems and the framework we develop should clarify these issues beyond the current application.

Land use regulation and land rent in a simple spatial equilibrium

We first develop a simple spatial model of land use regulation, land rent, and location. Our object is to motivate our approach to understanding the effects of land use regulation, and more particularly, to motivate our econometric strategy.

Consider two municipalities, L and R, occupying the real line between $-\overline{\overline{x}}$ and $\overline{\overline{x}}$. The two municipalities share a border at the origin, and the left municipality consists of points to the left of zero, while the right municipality occupies all to the right. Let x denote locations on the real line and let $m \in \{L,R\}$ index the municipalities.

The two municipalities are populated from a pool of agents who all; earn wage w, pay $p(x)$ for their residential location, and derive utility $V(x;\cdot)$ from their housing. The utility of each resident is $u(x) = e^{w-p(x)}V(x;\cdot)$. We discuss V in more detail below. The opportunity cost of land in both municipalities is zero. Mobility is costless, and agents may move to an alternative city where they receive a reservation utility, e^{θ}. In equilibrium all residents are indifferent between all locations in either municipality and the alternative city. Thus $\ln(u(x)) = \theta$ for all x. This implies that land rent is $p(x) = w - \theta + \ln(V(x;\cdot))$ for all x. Every location in each municipality is subject to development, but development in both municipalities is subject to regulation. Let z^m denote regulation in municipality m and let increasing values of z^m reflect increasingly stringent regulation.

We would like to know how $p(x)$ varies with location and regulation. We suppose that land use regulation has two effects on the value of land. The first is to decrease land values by constraining how a landowner to develops his land. Call this effect of regulation an 'own-lot effect'. This effect might operate in many ways; minimum lot size constraints may lead to houses and lots that are 'too large', waiting times for permits may increase financing or design costs, building codes may increase construction costs. For our purpose, the mechanics of how the own-lot effect impacts the value of land are not important, only that regulation z^m is binding and affects land values. Formally, let $v_O(z^m) \in R$ denote the magnitude of the own-lot effect as a function of z^m. Consistent with the discussion above, assume $v_O(z^m) > 0$ and $v'_O < 0$. Since regulation may

[1]Sierra club conservation policies, Adopted by the Board of Directors, February 1, 1986, http://www.sierraclub.org/policy/conservation/urban.aspx.

[2]National association of homebuilders, November 19, 2007, http://www.nahb.org/page.aspx/category/sectionID=633

vary at the municipal boundary, we define an own-lot effect function for the entire area of the two municipalities as,

$$V_O(x,z^L,z^R) = \begin{cases} v_O(z^L) & \text{if} \quad x \leq 0 \\ v_O(z^R) & \text{if} \quad x > 0. \end{cases}$$

Unless regulation is the same in both municipalities, i.e., $z^L = z^R$, V_O is step function with a discontinuity at zero.

We also suppose that land use regulation has an 'external effect'. That is, the value of any given location is affected by the regulation to which nearby locations are subject. Possible examples of this effect are: the value of a parcel will vary with density permitted at nearby locations if residents have a taste for low (or high) density neighborhoods; minimum setback requirements decrease the risk of fire spreading from one house to another; a regulation requiring that neighbors' garages not open facing the street increases the value of parcels to nearby residents who prefer not to see their neighbors' cars. As for the own-lot effect, the mechanics of how regulation causes an external effect are not important to our analysis, only that such an external effect impacts land values. Formally, let $v_E(z^m) \in R$ denote the magnitude of the own-lot effect as a function of z^m. Consistent with the discussion above, suppose that $v_E(z^m) \geq 0$ and $v_E' > 0$. That is, an increase in z^m causes v_O and v_E to move in opposite directions so that regulation which imposes larger own-lot costs has larger external benefits.

By construction, the external effect of regulation affects locations that are near other regulated locations. Thus, locations near $x = 0$ are exposed to the regulations of both municipalities. In particular, parcels in the right municipality but very close to zero are equally exposed to locations subject to z^L and to locations subject to z^R. The same statement is true for locations close to zero in the left municipality. Locations progressively further from the border are progressively more affected by the regulation of their own municipality. To formalize this intuition, define a continuous increasing function $\delta(x)$ which satisfies

$$\delta(x) = \begin{cases} -1 & \text{if } x \leq -\overline{x} \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x \geq \overline{x}. \end{cases}$$

We then write the utility derived from the external affects of regulation as

$$V_E(x,z^L,z^R) = \frac{1-\delta(x)}{2} v_E(z^L) + \frac{1+\delta(x)}{2} v_E(z^R).$$

For $x \leq -\overline{x}$ the external effects of regulation are entirely due to regulation in the left municipality. For $x \geq \overline{x}$ the external effects of regulation are entirely due to regulation in the right municipality. As we move closer to the municipal boundary the utility derived from external effects of regulation is a weighted sum of exposure to regulation in both municipalities. When we are precisely at the municipal border, the regulations of each municipality are equally weighted. Since δ is continuous in x, V_E is also continuous in x.

That the external costs or benefits of proximity to a regulated landscape decays with distance is one of the principal assumptions that we must make to identify the external effects of regulation.

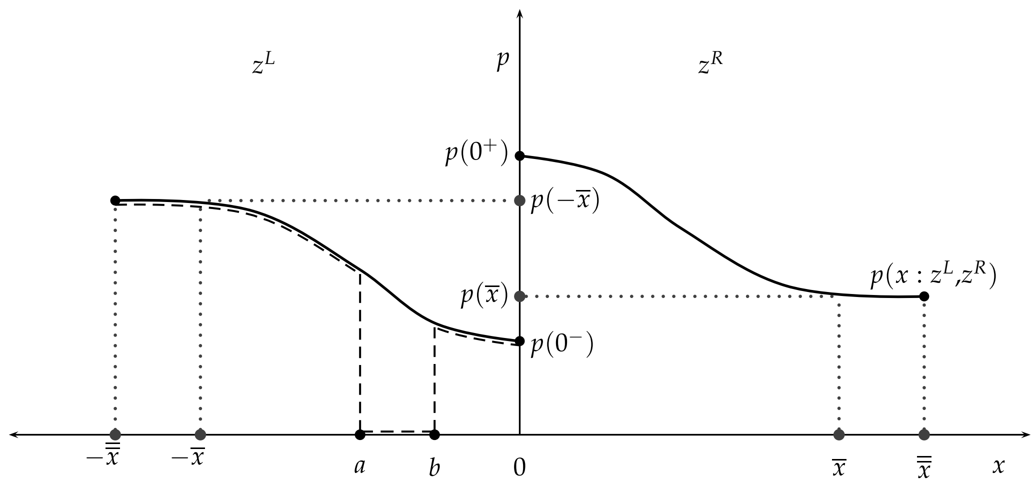


Figure 1. Land rent gradient across the municipal border.

We assume that the decay function operates over spatial scales that are small relative to the size of municipalities. This is consistent with the literature that estimates the effects of open space on residential housing prices and typically finds that the effects of open space attenuate over distances of less than one mile, e.g., Irwin and Bockstael (2002) or for a survey, McConnell and Walls (2005).

We now define the $V(x;\cdot)$ term in our original formulation of utility as $V(x;\cdot) = V_O(x,z^L,z^R)^{A_1} V_E(x,z^L,z^R)^{A_2}$ and write utility as

$$u(x) = e^{w-p(x)} V_O(x,z^L,z^R)^{A_1} V_E(x,z^L,z^R)^{A_2}. \tag{1}$$

With free mobility, land rent adjusts so that all agents are indifferent between every location in the two target municipalities and their reservation location. Recalling that e^{θ} is the reservation utility level, the resulting land rent gradient is

$$p(x) = w - \theta + A_1 \ln(V_O(x,z^L,z^R)) + A_2 \ln(V_E(x,z^L,z^R)) \tag{2}$$

The solid line in figure 1 illustrates a land rent gradient of the sort described by equation 2. In the more highly regulated left municipality, land rent increases with distance from the border as exposure to the less regulated right municipality drops, eventually reaching the level associated with full exposure to z^L at distance $\overline{x}$ from the border. We see the opposite pattern in the right municipality. As we move to the interior of this municipality, land rent drops as exposure to z^L decreases and exposure to z^R increases. Land rent varies discretely at the border. Locations on either side of the border face the same external effects of regulation, since both are equally exposed to regulation of each municipality, but municipalities on the right pay the lower own-lot effect associated with z^R, while those on the left pay the higher own-lot effect of more stringent z^L.

$$\begin{aligned}
p(-\overline{x}) &= w - \theta + A_1 \ln(V_O(-\overline{x},z^L,z^R)) + A_2 \ln(V_E(-\overline{x},z^L,z^R)) \\
p(0^-) &= w - \theta + A_1 \ln(V_O(0^-,z^L,z^R)) + A_2 \ln(V_E(0^-,z^L,z^R)) \\
p(0^+) &= w - \theta + A_1 \ln(V_O(0^+,z^L,z^R)) + A_2 \ln(V_E(0^+,z^L,z^R)) \\
p(\overline{x}) &= w - \theta + A_1 \ln(V_O(\overline{x},z^L,z^R)) + A_2 \ln(V_E(\overline{x},z^L,z^R)).
\end{aligned} \tag{3}$$

Figure 1 makes our approach to the problem clear. Locations close to, but on opposite sides of the border are each exposed to the same landscape. By comparing parcels on opposite sides of the border we compare parcels that experience the same external effects of regulation but different own-lot effects. Using equation 2 we have,

$$p(0^+) - p(0^-) \quad = \quad A_1 \ln(V_O(0^+, z^L, z^R)) - A_2 \ln(V_O(0^-, z^L, z^R)). \tag{4}$$

In words, we can infer the relationship between changes in regulation and changes in the magnitude of the own-lot effect from change in the land rent gradient across a municipal border. This intuition will be the basis for our 'own-lot effect' estimation, one of our two principal estimating equations.

If we instead compare a parcel near the border with a parcel far away from the border, then we are comparing parcels subject to the same own-lot effect, but different external effects. In particular, the boundary parcel is equally exposed to both types of regulation, while the interior parcel is wholly exposed to the external effect of its own municipality's regulation. From equation 2 we have,

$$p(-\overline{x}) - p(0^-) = A_2 \ln(V_E(-\overline{x}, z^L, z^R)) - A_2 \ln(V_E(0^-, z^L, z^R)). \tag{5}$$

This suggests that we infer the effect of a change from equal exposure to two levels of regulation, to sole exposure to one level of regulation by looking at changes in the land rent gradient as we move from a point very near a municipal border, to a point well on the interior. This intuition will be the basis for our 'external effect' estimation, our second principal estimating equation.

It is useful to note that, if we assume that V_E is symmetric around $x = 0$ for given z^L and z^R then using 2 and a little bit of algebra, we have

$$p(-\overline{x}) - p(\underline{x}) = ((p(0^+) - p(0^-)) + 2(p(-\overline{x}) - p(0^-)). \tag{6}$$

Therefore, estimations of 4 and 15 together allow an estimate of the total effect of changes in land use regulation on land rent.

Before we turn to the details of our econometric model, we note two issues which complicate the intuition behind our approach. First, it may be that land use regulation affects rent by affecting the supply of developable land. In a technical appendix argue that such 'supply' effects are not consistent with a spatial equilibrium in which at agents are free to choose their most preferred location. Second, it is also possible that different demographic groups value regulation differently. The technical appendix also considers this issue and argues that such preference heterogeneity by itself should not confound our estimates of the value of regulation. We note that other motivations for sorting into municipalities on the basis of demographic characteristics may be more problematic and are addressed below.

Related literature

Land use regulation and real estate markets

There is a large literature which investigates the determinants of zoning. Two recent contributions, Saiz (2008) and Hilber and Robert-Nicoud (2009), find evidence that regulation is more stringent where developable land is scarce. Wallace (1998) also finds that, although zoning does not perfectly follow the market, many of the same factors determine zoning and prices. McMillen and McDonald (1991) find somewhat more direct evidence that land prices affect zoning decisions. In all, this literature establishes that many of the same factors which determine land use regulation also determine land prices. It follows that land use regulation must generally be regarded as an endogenous explanatory variable in regressions to predict land price or other measures of land market behavior.

While there is a large literature which looks at the relationship between land use regulation and the real estate market, only a small subset of these papers correct for the endogenous determination of regulation. Thus, only this handful of papers can claim to find a causal effect of land use regulation on real estate markets. Among these, Mayer and Sommerville (2000) predict MSA level housing starts as a function of current and lagged housing prices and regulation. They instrument for regulation using historical demographic characteristics and find that housing starts respond more slowly to price changes as regulation is more stringent. Ihlanfeldt (2007) looks at the effect of municipal land use regulation in Florida, and like Mayer and Sommerville (2000) uses historical demographic variables as instruments for regulation. He finds that regulation increases house prices and decreases land prices. Saiz (2008) estimates a system of equations in which housing demand, housing supply and regulation are all endogenously determined. He concludes that regulation is more stringent where developable land is scarce and that regulation serves to increase housing prices.

Wallace (1998) studies the impact of zoning on King County, Washington, and concludes that zoning generally has a large negative effect on land values. As in the other papers cited above, she estimates a two equation model, one for zoning and one for prices. Unlike other papers in this literature she identifies her model by including a variable in the price equation, whether the parcel has been 'platted', that is excluded from the zoning equation. This exclusion restriction appears reasonable, a parcel cannot be divided into lots until zoning is settled. Zhou, McMillen, and McDonald (2008) consider a 1957 change to Chicago zoning, and compare matched parcels on either side of a zoning border, before and after the change. They find that Chicago's change in zoning does not affect the value of residential property and increases the value of commercial property. That is, the change in zoning that Zhou *et al.* (2008) examine appears to increase the value of land.

Suzuki (2008) considers the effect of land use regulation, on competition in the Texas hotel market. His results suggest that regulation is not welfare improving. To deal with the possible endogeneity of regulation, he uses a bound estimator. This approach exploits the researcher's prior about the slope of a demand or supply curve to bound the causal effects of an endogenous regressor.

Libecap and Lueck (2009) consider the effect of two different land demarcation schemes on land prices. These two schemes, 'metes and bounds' and the 'rectangular system', respectively define parcel boundaries on the basis of irregular natural features of the landscape and a regular rectangular grid laid out by a surveyor. Comparing prices across adjacent Ohio counties subject to the two different demarcation schemes, they find measurably land higher prices land prices in areas where parcels are described by the rectangular system than by metes and bounds.

Mayer and Sommerville (2000), Ihlanfeldt (2007), and Saiz (2008) conclude that land use regulation restricts supply and drives up the price of housing. Each relies on an instrumental variable or exclusion restriction to identify the effect of land use regulation. Wallace (1998) relies on an exclusion restriction that does not appear to be open conventional objections. Zhou *et al.* (2008) exploit a plausible natural experiment. Suzuki (2008) employs a bound estimator to deal with the endogeneity of regulation. Zhou *et al.* (2008), Suzuki (2008), Ihlanfeldt (2007) and Wallace (1998) are concerned with small geographic areas and specific regulations. Zhou *et al.* (2008) and Wallace (1998) are concerned with the effect of specific regulations. Zhou *et al.* (2008) and Libecap and Lueck (2009) find convincing evidence that particular regulations can increase land value. All other papers suggest that regulation is harmful.

While the literature does not generally attempt to distinguish the costs and benefits of regulation, it probably fair to say that the literature cited above is more interested in measuring the cost of regulatory constraints on landowners, our own-lot effect. There is a large complementary literature which looks at the effects of nearby open space and amenities (see McConnell and Walls (2005) for a survey). Arguably, this literature is more interested in external effects, but to our knowledge none of this research corrects for the probable endogeneity of amenities and open space to housing or land prices. Two papers, however, explicitly consider the external effects of regulation. First, Cheshire and Sheppard (2002) conduct a calibration exercise, based on the city of London, which explicitly describes the costs and benefits of regulation. They conclude that the net welfare impact of open space regulation in London is negative. Rossi-Hansberg, Sarte, and Owens (2009) consider the effect on nearby house prices of an exogenous government policy to invest in certain houses, but not others. This allows them to calculate the external effect of these investments, and is the closest analog in the literature to our efforts to measure the external effects of regulation.

In sum, there is a small literature which attempts to estimate the causal effect of land use regulation on land markets and housing markets. This literature generally, but not always, concludes that land use regulation is harmful. This literature also does not generally recognize that regulation should be expected to have two effects, one a harmful constraint on particular landowners, the other a (hopefully) beneficial consequence of constraints on a landowner's neighbors. We contribute to the debate by suggesting a new identification strategy, by disentangling the own-lot and external effects of regulation, and by using rich national level data on regulation and land sales.

Other border studies

Our two main econometric exercises are variants of a regression discontinuity design. The Regression Discontinuity design is increasingly popular and is used to investigate, for example, the effect of class sizes on educational attainment (Angrist and Lavy, 1999), the effect of changes in social assistance programs on employment (Lemieux and Milligan, 2008), or the effect of mayoral party affiliation on municipal policies (Ferreira and Gyourko, 2009). Theory and best practice are described in Hahn, Todd, and Van der Klaauw (2001) and Imbens and Lemieux (2008).

The method has also been used to investigate the effect of policies which vary over physical space as one crosses from one administrative unit to another. In this case, the cut-off of interest is an administrative boundary. Holmes (1998) looks at the impact of changes in right-to-work laws on manufacturing employment near state borders. Black (1999) and Bayer, Ferreira, and McMillan (2007) look at the effect of changes in property values near school district boundaries. Duranton, Gobillon, and Overman (2008) look at the effect of changes in municipal taxation across municipal boundaries on the behavior of firms near these boundaries. In effect, these authors identify the effect of their chosen policy in a two step process. In the first, they estimate the discontinuity of interest at many borders, and in the second they examine the correlation between the magnitude of these cross-border discontinuities and the corresponding cross-border change in the policy variable of interest. This is also the intuition behind our 'own-lot effect' regression. To our knowledge, the intuition behind our 'external effect' regression is novel.

Econometric model

To estimate the own-lot effect and the external effect of land use regulation on the basis of the intuition developed in section 1 we must develop a credible empirical description of land rent gradients near municipal borders. Such a description of the land rent gradient should reflect the following possibilities: that locations may differ in their intrinsic attractiveness; that members of different demographic groups may have a taste for proximity to others in their own group; that people may sort on the basis of their tastes for local public goods not related to land use; that different demographic groups may value land use regulation differently.

To derive estimating equations consistent with these possibilities, we begin with a stylized description of the process which determines individual location choices, land use regulation, and land prices in any given pair of adjacent municipalities. In particular, we imagine that at an initial time period, 'time zero', both municipalities are sparsely populated by a first wave of heterogenous immigrants drawn to heterogenous locations. These initial immigrants choose land use regulation in municipal elections. At 'time one', a second wave of heterogenous immigrants choose locations on the basis of location attributes, regulation and possibly proximity to members of their own demographic group. The simple spatial equilibrium of section 1 analyzes land prices resulting from the location choices of the second wave of immigrants when land is homogenous and agents do not care about the identity of their neighbors.

We now describe this model more thoroughly and develop the notation necessary to describe the resulting land rent gradient. We maintain the same basic description of physical space developed earlier: there are two municipalities, L and R which occupy the intervals $[-\overline{\overline{x}},0)$ and $(0,\overline{\overline{x}}]$ respectively. However, we now suppose that each location x has an intrinsic attractiveness, $a(x)$, and that this intrinsic attractiveness can be decomposed into two parts. A deterministic component $f(x)$ and a stochastic component $\phi(x)$. We suppose that f is positive and, to fix ideas, decreasing in x. For one of our two main estimations we will require that f be continuous at 0. We suppose that for all x, $\phi(x)$ is mean zero, that $E(\phi(x))$ exists, and that $\mathrm{Cov}(f(x),\phi(x)) = 0$. Intuitively, there is more sunshine or a shorter commute as we move from right to left with some noise around the trend. On average, the left municipality is nicer than the right. The assumption that f is continuous at zero requires that the municipal border does not separate qualitatively different types of land. We require that $a(x)$ be exogenous in both of our main regressions and that the location of borders be exogenous in one of our main regressions.

We imagine that the municipalities are populated in two stages. At time zero, measure zero of immigrants locate in the two municipalities.[3] Each immigrant has a type $N \in [0,1]$. N can describe any demographic characteristic, but to ease exposition we call it education. Types match to locations on the basis of their attractiveness, and we let $N_0(a(x))$ describe this matching. We suppose that $N_0' > 0$, so that more highly educated people match to nicer places.

Time zero immigrants choose land use regulation for their respective municipalities democratically. Let N_0^L denote the demographic characteristics of the mean resident at $x \leq 0$ and N_0^R the mean demographic characteristic for residents located at $x > 0$. As a stylized way to describe the choice of regulation, let $z(N_0^m)$ describe the relationship between the mean voter and the resulting choice of regulation. We suppose that $z > 0$ and that z is continuous and increasing. It follows from our assumptions on N_0, f, $z(.)$ and ϕ that $N_0^L > N_0^R$ and hence that $z^L > z^R$. Thus, nicer places attract better educated people and are more intensively regulated. Let $\hat{z} = [z(N_0^L),z(N_0^R)]$ denote the observed pair of regulatory intensities in our two municipalities.

A second wave of immigrants, of measure $2\overline{\overline{x}}$, now settles the remaining locations. These immigrants match to the location which gives them the highest utility. Denote the resulting distribution of immigrants by $N_1(x)$. By assumption, the initial agents occupied measure zero of the available land, so that N_0 does not affect the supply of land available to the second wave of immigrants. The first wave of immigrants, N_0, affects the second wave, N_1, only through its choice of regulation, $\hat{z}$.

As in our earlier theoretical model, we suppose that each immigrant chooses between locations in the left or right municipality, and the alternative city. If they locate in the left or right municipality, they receive a wage that does not vary with location. Generalizing the earlier theoretical model, we allow wages and outside options to vary with type and denote them by $w(N_1)$ and $\theta(N_1)$. We suppose that both functions are continuous and increasing: wages and outside options increase smoothly with education.

Land use regulation contributes to the welfare of immigrants in exactly the same way as described in our earlier theoretical model, i.e., equation 1.

Finally, we allow the possibility that immigrants derive utility from proximity to other immigrants whose types are 'close' to their own. This effect will be determined by the immigrant's own location x, the immigrant's type $N_1(x)$, and the distribution of other agents, N_1. Let $\gamma(x,N_1(x), N_1)$ denote the utility derived from proximity to other people.[4]

We can imagine two mechanisms by which proximity to immigrants of particular types can affect utility. In the first, immigrants sort into municipalities on the basis of their tastes for local public goods. In this case, the value of γ is determined in much the same way as is regulation. It is

[3]This assumption that time zero immigrants are measure zero simplifies exposition but is not essential to the intuition we develop.

[4]Since N_0 involves measure zero of immigrants, provided that γ is constructed by integrating any continuous objective function on some real interval, N_0 does not affect the value of γ.

based on the levels of public services and local taxes that are determined by election. In this case, we expect that γ varies discretely with the level of public services at $x = 0$. Note that in this case, the level of γ in the two municipalities depends on mean demographic characteristics in the two municipalities. Alternately, the value of γ may reflect a preference for proximity to people with similar characteristics.[5] In this case, γ is determined by the whole distribution of N_1, but it should be expected to vary continuously at the municipal border.

With this notation established, we write the utility of the agent at location x as the product of the different components described above.[6] That is,

$$u(x) = e^{w(N_1(x))-p(x)} V_O(x,\hat{z})^{A_1} V_E(x,\hat{z})^{A_2} \gamma(x,N_1(x),N_1)^{A_3} e^{a(x)}. \tag{7}$$

An equilibrium is an arrangement of types and a land rent gradient such that all agents are indifferent between their own location and their reservation location and no agent would prefer another agent's location. Thus we have, for all x, that $\ln(u(x)) = \theta(N_1(x))$. Together with equation 7, we have the generalization of the land rent gradient corresponding to equation 2,

$$p(x) = w(N_1(x)) - \theta(N_1(x)) + A_1 \ln(V_O(x,\hat{z})) + \\ A_2 \ln(V_E(x,\hat{z})) + A_3 \ln(\gamma(x,N_1(x),N_1)) + a(x). \tag{8}$$

In the theoretical model of section 1, only regulation changes when we cross the municipal border. In the more realistic model of prices given by equation 8, it is at least possible that other determinants of land rent change discretely at the border. This means that it will not generally be possible to identify the effect of regulation on land rents by looking at only a single border. Instead, we must look for a relationship between land rent and regulation across a set of many municipal borders. An analysis of many borders requires that we generalize our notation.

Let $j \in \{1,...,J\}$ index municipal borders and let a j superscript indicate a scalar or function that is particular to a border. We will adopt the convention that the more stringently regulated municipality is the left municipality, and refer to individual municipalities as 'left' or 'right' of border j. Informally, each j refers to a replication of figure 1.

We also generalize our earlier description to allow heterogeneity across border pairs. Let $a^j(x) = f(x,\mu^j) + \phi(x)$ be the intrinsic attractiveness of location x of border j. As before, f describes a trend around the border, but we now suppose that it is parameterized by the pair $\mu = (\mu_1,\mu_2) \in R^2$ with $f(x,\mu) = \mu_1 + \mu_2 x$. We suppose that μ is a random variable with density $g_\mu : R^2 \longrightarrow [0,1]$, and that each border draws a single μ^j. Thus, the μ's are a generalization of a fixed effect and parameterize changes in the intrinsic attractiveness in a neighborhood of border j. Further suppose that $\phi(x)$ is a real valued random variable with density $g_\phi : R \longrightarrow [0,1]$. We suppose that $\phi(x)$ is identically distributed for all j and x, that $E(\phi(x)) = 0$ and that $Cov(\phi(x),\phi(y)) = 0$ for all $x,y \in [-\bar{\bar{x}},\bar{\bar{x}}]$ and all j. Note that the intuition behind $a^j(x)$ is not changed from our initial

[5]One such γ is $\gamma(x,N_1(x),N_1) = \int_{-\bar{\bar{x}}}^{\bar{\bar{x}}} e^{-\rho|x-y|}|N_1(x) - N_1(y)|dy$ for ρ a positive real 'decay rate'.

[6]With the multiplicative specification of utility, the marginal utility of changes in z vary with income. This allows us to rationalize our observation that different municipalities choose different regulations. In an additive specification, the marginal utility of regulation does not vary with income, so it is hard to rationalize the heterogeneity of observed regulation.

discussion. $a^j(x)$ describes the fact that as we move from one municipality to another, locations may become systematically more attractive and there will be some noise around this trend.

Our data describe transactions of particular parcels. To describe these data, as opposed to hypothetical rent gradients, we let i index parcels in a border pair j. We refer to the sale price of a particular parcel as p_i^j, with other parcel attributes indexed similarly. We refer to the location of parcel i in border pair j as x_i^j. The magnitude of x_i^j is the distance from border j, with negative distances indicating displacement into the more intensively regulated municipality, and positive displacements indicating displacements into the less regulated municipality.

We will sometimes need to distinguish between the municipalities that form a border pair. To do this, we recall that municipalities within a border pair are indexed by $m \in \{L,R\}$ and introduce an extra subscript. Thus, p_i^{mj} refers to the price of parcel i in the m municipality of border pair j. Similarly, p_i^{-mj} refers to a parcel in the other municipality of border pair j.

Naive regression

One obvious approach is to exploit cross-municipality variation in mean land rent and regulation. If we ignore edge effects, then equation 8 lets us write the price for either municipality (say R) as,

$$p^j(x) = w(N_1^j(x)) - \theta(N_1^j(x)) + A_1 \ln(V_O(z(N_0^{Rj}))) \\ + A_2 \ln(V_E(z(N_0^{Rj}))) + A_3 \ln(\gamma(x,N_1^j(x),N_1^j)) + a^j(x). \quad (9)$$

This is an ordinary hedonic regression and is the basis for much of the extant research on land use regulation.

The problem with this approach is clear. Recalling that $N_0^{Lj} = E(N_0(a^j(x)|\overline{\overline{x}} < x < 0)$ and that $z^j = z(N_0^{Lj})$, we see that z depends on the distribution of the municipalities' initial attractiveness, and in particular on μ^j. It follows that if intrinsic attractiveness is not observed by the econometrician, as must surely be at least partly the case, then $z(N_0^R)$ must be correlated with the error term. Equation 9 also makes clear the problem with using historical demographic characteristics as instruments as in Ihlanfeldt (2007). Since these variables are themselves functions of the intrinsic attractiveness of the location, they fail the requirement that they be orthogonal to $a(x)$.

Own lot effect

To overcome the endogeniety problem that affects the cross-municipality regression described by equation 9 we exploit the intuition developed in section 1 to separately estimate the own-lot and external effects of regulation. In this section we develop our own-lot effect estimating equation, the empirical counterpart of equation 4.

To begin, substitute the more realistic description of the land rent gradient provided in equation 8 into description of the cross-border land rent differential of equation 4. Recall that $a^j(x) \equiv f^j(x) + \phi(x)$. By construction, V_E is continuous at 0 and we here assume f^j is continuous at zero.

Thus we have,

$$p^j(0^-) - p^j(0^+) = \begin{bmatrix} w(N_1^j(0^-)) - \theta(N_1^j(0^-)) + A_1 \ln(V_O(0^-,\hat{z})) \\ +A_3 \ln(\gamma(0^-,N_1^j(0^-),N_1^j)) + \phi(0^-) \end{bmatrix} - \begin{bmatrix} w(N_1^j(0^+)) - \theta(N_1^j(0^+)) + A_1 \ln(V_O(0^+,\hat{z})) \\ +A_3 \ln(\gamma(0^+,N_1^j(0^+),N_1^j)) + \phi(0^+) \end{bmatrix}. \quad (10)$$

Given a sample of municipal borders, this expression describes the relationship between the land rent gap at the border and the cross-border difference in regulation, along with several possible confounding factors. Our problem is to develop estimating equations that isolate the relationship between regulation and price.

To begin, we assume that only regulation varies discontinuously at the border. While we will relax this assumption in what follows, note, that if γ reflects peoples' preference for being near others in their own demographic group, then we expect it to be described by a 'potential' function of the sort given in footnote 5, and consequently to be continuous around zero. Under this continuity assumption, we are left with,

$$p^j(0^-) - p^j(0^+) = A_1 \ln(V_O(0^-,\hat{z})) - A_1 \ln(V_O(0^+,\hat{z})) + \phi(0^-) - \phi(0^+)$$

That is, if w,θ,and γ are all continuous then any discontinuity in land rent across the municipal border entirely reflects differences in the own-lot effect in the two adjoining municipalities. These own-lot effects, in turn, are functions of regulation. To proceed, parameterize $A_1 \ln(V_O(0^+,\hat{z})) - A_1 \ln(V_O(0^-,\hat{z}))$ as a linear function of the difference in municipal regulations. That is,

$$A_1 \ln(V_O(0^+,\hat{z})) - A_1 \ln(V_O(0^-,\hat{z})) = B_0(z^{Lj} - z^{Rj}).$$

B_0 measures the relationship between regulation and land rent and is the parameter of interest.

We have assumed that w, θ and γ are all continuous at the border. If we restrict attention to transactions that are 'close enough' to the border we can treat w,θ and γ as constant. Thus, in our narrow interval, prices on each side of the border are described by a constant and by parcel specific errors. The only systematic difference between cross-border parcels is due to regulation. Therefore, defining χ_i^{Lj} to be an indicator variable that is one if parcel i lies in the left municipality of border pair j, we have the estimating equation,

$$p_i^j \quad = \quad A_0^j + \chi_i^{Lj} B_0(z^{Lj} - z^{Rj}) + \phi_i^j. \quad (11)$$

Conditional on our other assumptions, this estimating equation will give us unbiased estimates of B_0 provided that z^{Lj} and z^{Rj} are orthogonal to ϕ_i^j. This orthogonality follows from the fact that regulation is a function of mean municipal characteristics. Loosely, since we are considering only a small section of the municipality, and since $\phi(0)$ gives us no information about ϕ at other values of x, it follows that $\phi(0)$ must be orthogonal to any function of municipal mean characteristics, land use regulation in particular.

While equation 11 allows us to estimate how the own-lot effect of regulation varies with regulation, this estimation relies on two strong assumptions. First, that we restrict attention to parcels close enough to the border that everything except own-lot effects and idiosyncratic error are constant. Second, that only regulation varies discontinuously at the border. In particular, we require that the systematic part of the intrinsic attractiveness of the parcels, f, does not vary discontinuously at the border. This is exactly analogous to the continuity assumption required for a standard RDD estimation (Hahn *et al.*, 2001) and requires that municipal borders not divide one 'quality' of land from another.

To relax the assumption that factors other than regulation are constant in a neighborhood of the border, we allow the slope of the rent gradient to vary at each border and we parameterize $w - \theta + \gamma + f$ as,

$$w(N_1^j(x_i^j)) - \theta(N_1^j(x_i^j)) + \gamma(0^-, N_1^j(x_i^j), N_1^j) + f(x_i^j, \mu^j) = C_0^j + C_1^j x_i^j.$$

In this case, equation 11 becomes,

$$p_i^j \quad = \quad (A_0^j + C_0^j) + C_1^j x_i^j + \chi_i^{Lj} B_0 (z^{Lj} - z^{Rj}) + \phi_i^j. \tag{12}$$

A few comments are in order. First, if equation 12 is correct, and we estimate equation 11, we expect the resulting estimates of B_0 to be biased. In this case, since regulation is correlated with μ^j, regulation is endogenous. Second, note that 12 implicitly assumes that the slope of the rent gradient does not change at the border. It is straightforward to generalize this equation to allow the slope of the rent gradient to vary across the border. Finally, note that equation 12 requires that for each border pair of municipalities, we estimate two constants, an intercept and a slope, thus this estimation is likely to be very demanding of our data.

Equation 12 continues to rely on the assumption that w,θ and γ all vary continuously at the border. To relax this assumption, note that each of these functions depends solely on demographic characteristics. Thus, we can parameterize a border discontinuity as a function of demographic characteristics. To accommodate this, let w^{Lj} denote a vector of municipal mean demographic characteristics for the left municipality, and w^{Rj} the corresponding vector for the right municipality. We can then write

$$\Big[w(N_1^j(0^-)) - \theta(N_1^j(0^-)) + \gamma(0^-, N_1^j(0^-), N_1^j)\Big] - \Big[w(N_1^j(0^+)) - \theta(N_1^j(0^+)) + \gamma(0^+, N_1^j(0^+), N_1^j)\Big] = D_0 + D_1(w^{Lj} - w^{Rj}).$$

If we include this term in equation 12, we have

$$p_i^j \quad = \quad (D_0 + A_0^j + C_0^j) + D_1 \chi_i^{Lj} (w^{Lj} - w^{Rj}) + C_1^j x_i^j + \chi_i^{Lj} B_0 (z^{Lj} - z^{Rj}) + \phi_i^j. \tag{13}$$

Estimating this equation will allow us to assess whether border discontinuities are partly due to changes in demographics across borders. Two comments about this regression are in order. First, if we estimate equation 12 when equation 13 is correct, since regulation and demographics are correlated (by construction) we will attribute to regulation part of the border gap that equation 13 attributes to demographics. If regulation does not cause the difference in regulation, this means that equation 12 overstates the effects of regulation. On the other hand, if demographic sorting occurs because of land use regulation, then equation 12 estimates a long run or total effect of regulation, while equation 13 estimates a partial effect.

Note that parameterizing the cross-border gap with municipal mean demographics is not strictly correct. w,θ and γ are all functions of $f(x,\mu)$, the attractiveness of particular places, rather than of municipalities. It would be better to parameterize the cross border gap as a function of very local demographics as well. Data limitations prevent this. However, we can control for parcel specific measures of intrinsic attractiveness, measures of geography and commuting distance in particular. To the extent that these variables measure location specific heterogeneity correlated with demographics and the cross-border gap in land rents, they improve our estimates. With this in mind, let y_i^j denote a parcel specific vector describing geography and commuting distance. We then write our final estimating equation based on equation 4 as,

$$p_i^j = (D_0 + A_0^j + C_0^j) + D_1\chi_i^{Lj}(w^{Lj} - w^{Rj}) + D_2 y_i^j + C_1^j x_i^j + \chi_i^{Lj} B_0(z^{Lj} - z^{Rj}) + \phi_i^j. \qquad (14)$$

There are two other objections to our estimation strategy. First, suppose that municipalities also choose a level of public service like the frequency of trash collection, denoted by z^{*L} and z^{*R}, that may vary discretely at the municipal border and that this other regulation also impacts land prices. Allowing for this sort of regulation in equation 4 we have,

$$\begin{aligned} p^j(0^-) - p^j(0^+) = {} & A_1 \ln(V_O(0^+,\hat{z})) - A_1 \ln(V_O(0^-,\hat{z})) + \\ & A_2 \ln(V_O(0^+,\hat{z^*})) - A_1 \ln(V_O(0^-,\hat{z^*})) + \phi(0^-) - \phi(0^+). \end{aligned}$$

It is clear that if the two types of regulation are correlated, our approach will generally confound the effects of the two types of regulation. The exception to this is if z^* is itself a function of z. That is, if zoning for large lots leads to a community with twice weekly trash collection, then our estimator will estimate the total effect of zoning, including the effect of the induced high rates of trash collection. With this said, we expect that, given a democratic process for determining these other regulations, that the cross-border change in other regulation will be systematically related to the cross border change in observed demographics. Thus, equations 13 and 14 ought to substantially correct for these problems. If there are unobserved changes in other regulation at municipal borders then we require that, conditional on control variables, these changes be uncorrelated with land use regulation. [7]

External effect

We now turn to estimating the external effect of regulation by looking at the price difference between peripheral and interior parcels. To develop this 'external effect' estimation we substitute the description of the price gradient from equation 8 into the expression describing the price difference between interior and peripheral parcels, equation 5. Recalling that the own-lot effect

[7] A final objection to these estimates of the own-lot effect is that they do not allow the value of regulation to vary systematically with the intrinsic attractiveness of a place. To address this issue, one could allow regulation to interact with measures of landscape, climate or topography, and test whether these interaction terms predict changes in land prices across borders. In practice, our data do not allow us to identify such an effect.

of regulation must be the same for two parcels in the same municipality, this gives,

$$p^j(-\overline{x}) - p^j(0^-) = \left[\begin{array}{l} w(N_1^j(-\overline{x})) - \theta(N_1^j(-\overline{x})) + A_2 \ln(V_E(-\overline{x},\hat{z})) \\ +\gamma(\overline{x},N_1^j(-\overline{x}),N_1^j) + f(-\overline{x},\mu^j) + \phi(-\overline{x}) \end{array} \right] - \left[\begin{array}{l} w(N_1^j(0^-)) - \theta(N_1^j(0^-)) + A_2 \ln(V_E(0^-,\hat{z})) \\ +\gamma(0^-,N_1^j(0^-),N_1^j) + f(0^-) + \phi(0^-) \end{array} \right]. \quad (15)$$

This equation describes the relationship between cross-border changes in regulation and the price difference between peripheral and interior parcels, together with a detailed description of possible confounding factors. Our problem in this section is to develop estimating equations which isolates the relationship between changes in regulation and changes in land prices.

If γ reflects the value of local goods that results from a given sorting of population then these public goods should be provided equally to the whole municipality. In this case, $\gamma(0^-,N_1^j(0^-),N_1^j) = \gamma(\overline{x},N_1^j(-\overline{x}),N_1^j)$, and the two terms involving γ drop out of equation 15. If we also suppose that w and θ are also constant within the left municipality and recall that $f(x,\mu) = \mu_1 + \mu_2 x$, then we are left with,

$$p^j(-\overline{x}) - p^j(0^-) = A_2 \ln(V_E(-\overline{x},\hat{z})) - A_2 \ln(V_E(0^-,\hat{z})) - \mu_2^j \overline{x} + [\phi(-\overline{x}) - \phi(0^-)]. \quad (16)$$

Next we parameterize $A_2 \ln(V_E(-\overline{x},\hat{z})) - A_2 \ln(V_E(0^-,\hat{z}))$ as

$$A_2 \ln(V_E(-\overline{x},\hat{z})) - A_2 \ln(V_E(0^-,\hat{z})) = B_0(z^L - z^R).$$

B_0 is the parameter of interest and describes the external effect of regulation on land prices. Finally, restrict attention to parcels that are either in a narrow interval near a municipal border, or in a narrow interval around -$\overline{x}$. and define χ_i^{Imj} to be an indicator variable that is one when a parcel i is an interior parcel lying within a narrow band around $\overline{x}$ and zero otherwise. We can now re-write estimating equation (16) as,

$$p_i^{mj} = A_0^{mj} + \chi_i^{Imj} B_0(z^{Lj} - z^{Rj}) + \chi_i^{Imj} \mu_2^j + \phi_i^{mj}. \quad (17)$$

Inspection of this equation shows that our strategy of comparing interior and boundary parcels eliminates the influence of the level of f, but not of its slope. However, since neither μ_2^j nor $z^{Lj} - z^{Rj}$ varies within a municipality, we cannot separate the effects of these two components on p. To resolve this problem, we first introduce controls y_i^{mj} that allow us to estimate $\mu_2^j(-\overline{x})$ explicitly. These controls will include measures of physical geography and commuting distance. More formally, we parameterize $\mu_2^j = C_1 y_i^{mj} + \epsilon_i^{mj}$. Substituting in 18 gives,

$$p_i^{mj} = A_0^{mj} + \chi_i^{mj} B_0(z^{Lj} - z^{Rj}) + C_1 y_i^{mj} + \epsilon_i^{mj} + \phi_i^{mj}. \quad (18)$$

In this equation 18, the municipality specific constant measures the level of land rent at the municipal border, and all variation between the border and interior points is attributed either to regulation or to changes in intrinsic attractiveness between interior and boundary locations. We have already established that regulation is orthogonal to ϕ. The additional orthogonality assumption required here is that regulation is orthogonal to ϵ. In words this orthogonality condition is that the unobservable component of the slope of f is uncorrelated with regulation. This does not seem like a strong assumption, provided we have reasonably strong controls.

To allow for the possibility that w,θ, and γ are not constant between out two intervals, let w^{mj} and w^{-mj} denote a vector of demographic characteristics for municipality mj and for its counterpart $-mj$. We parameterize the difference in these quantities as,

$$\left[w(N_1^j(-\overline{x})) - \theta(N_1^j(-\overline{x})) + \gamma(\overline{x},N_1^j(-\overline{x}),N_1^j)\right] - \left[w(N_1^j(0^-)) - \theta(N_1^j(0^-)) + \gamma(0^-,N_1^j(0^-),N_1^j)\right] = C_0(w^{Lj} - w^{Rj}). \quad (19)$$

With this notation in place, we write

$$p_i^{mj} = A_0^{mj} + \chi_i^{Imj} B_0(z^L - z^R) + \chi_i^{Imj} C_0(w^{Lj} - w^{Rj}) + C_1 y_i^{mj} + \phi_i^{mj}. \quad (20)$$

As for the own-lot effect regression, it would be desirable to control for demographic characteristics exactly at the locations of our transactions. Data limitations prevent this. The specification above approximates this ideal.

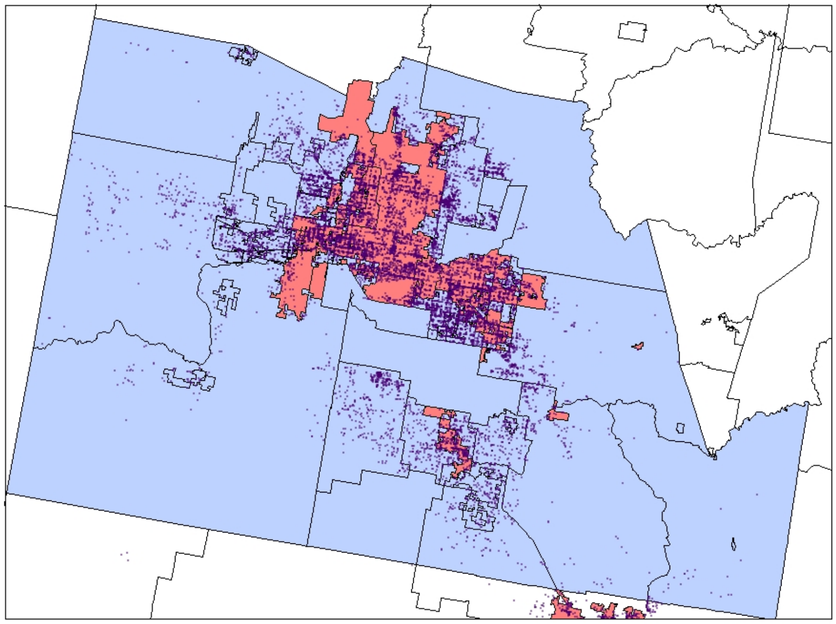
It remains only to determine the widths and locations of the border and interior bins. While theory does not provide any guidance on this issue, the available literature (see the discussion at the end of section 1) suggests that the scale over which should expect the external effects to decay is less than a mile. Thus, we will experiment with different sizes and locations for the interior and peripheral bins. By inspection of figure 1, if increasing the distance of the interior bin from the border affects our estimates, then this bin should be moved further from the border.

Data

Implementing the regressions described in section 1 requires three principal types of data. First, a description of land transactions. In particular, the location, price, and other characteristics of parcels that changed hands. Second, a description of land use regulation by municipalities. Third, a map which allows the parcel and regulation data to be integrated and border distances calculated.

To measure land prices, we use the 'COSTAR' data. This is a proprietary data set describing all land transactions in 25 major metropolitan areas between 1983 and 2009 for which the sale price was more than 250,000 dollars. In addition to recording the latitude and longitude of each parcel, the COSTAR data records transaction price and date, parcel size, and many other details about the parcel. Most of our empirical analysis restricts attention to transactions that occurred during 2000-2009, although we also experiment with samples of transactions that occurred during 1990-2009 and during 2002-2009.

To measure municipal land use regulation, we use the Wharton Land Use Regulation data (WRLURI) (Gyourko, Saiz, and Summers, 2008). These data result from a 2005 survey of 2729 US municipalities and describe many different aspects of municipal land use regulation: minimum lot size, permit waiting times, growth controls, in addition to an index which summarizes overall regulatory intensity. We rely principally on this summary index as our measure of regulation, although we experiment with measures of particular regulations.



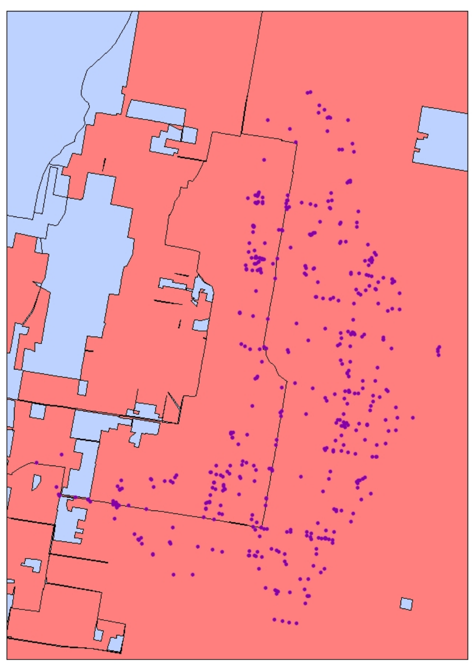


Figure 2. Top: Distribution of parcels in metropolitan Phoenix, with a detail of one particular municipal border. Light blue/light grey indicates counties in the Phoenix MSA, pink/dark grey indicates municipalities and purple/black dots indicate parcel. transactions.Bottom: Detail of Glendale-Phoenix border and parcels matched to this border. Glendale is on the right, Phoenix on the left.

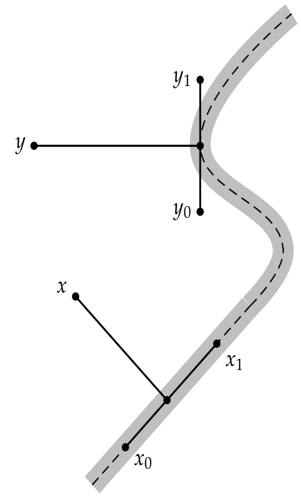


Figure 3. Illustration of our algorithm for identifying straight-line borders. To determine whether a parcel matches to a straight-line border we first calculate the shortest vector which connects the parcel to a municipal border (the length of this vector is our calculated distance to a municipal border). We next calculate the terminal points of the two 1km long vectors originating on the border and orthogonal to the first vector. If these terminal points both lay within 75 meters of the border then the parcel is determined to 'match to a straight-line border', otherwise it 'does not match to a straight-line border'. The figure above illustrates the algorithm. The thin dashed line represents a municipal border and the wider gray line a buffer around this border. Parcel x matches to a straight-line border, parcel y does not.

The WRLURI describes regulation in both incorporated and unincorporated municipalities. Since the US census does not produce a map which shows the boundaries of both types of units, we overlay a 2000 census map of places on the corresponding map of county subdivisions. We then matched each municipality in WRLURI to this map. This allows us to assign regulation to places on the map. We note that the WRLURI samples 2729 municipalities, whereas there are about 55,000 municipalities on our map, so that only a small fraction of US area is covered by the WRLURI. Since the COSTAR data records the latitude and longitude of each transaction, we are also able to locate each COSTAR parcel in our map.

To implement our own-lot effect and external effects estimations we must assign each observed transaction to a border that separates two adjacent municipalities. We do this by calculating the Euclidean distance from each parcel to the nearest municipal boundary and assigning the parcel to this boundary. This process simultaneously selects the neighboring municipality for each parcel. Thus we are able to organize our data around municipal boundaries in conformance with the econometric models developed above.

For a parcel to inform our estimation, we must observe the regulation in its home municipality and in its neighboring municipality. After restricting attention to the period between the last quarter of 1999 and before the second quarter of 2009 and dropping transactions for which no price data is available, we are left with 136,961 transactions. Of these, 9798 are matched to municipal borders for which the WRLURI records regulation in the home and neighboring municipality. Our main sample is restricted to this set of transactions. This sample of parcels lies in 513 of the 2729 WRLURI municipalities, and describes 598 municipal border pairs.

As a more concrete illustration of our data, the top panel of figure 2 maps all COSTAR parcels and municipalities in the Phoenix-Mesa MSA. Light blue indicates the extent of the MSA. Pink indicates the extent of WRLURI municipalities. Purple dots indicate COSTAR transactions. White indicates a county or municipality not in WRLURI or the MSA. The bottom panel provides a detail of the Glendale-Phoenix municipal boundary and of the COSTAR parcels matched to this boundary.

Our regressions require that municipal borders be exogenous. In particular, we require that they not be drawn to systematically separate more attractive land from less attractive land, and in particular, that municipality boundaries not follow natural features where the attractiveness of land changes discretely.

To identify exogenous borders we restrict attention to municipal boundaries which are straight lines, and therefore probably do not follow features of the landscape.[8] We rely on an algorithm to identify parcels which are associated with straight-line municipal boundaries. For each parcel, we identify the shortest vector which reaches from the parcel to the nearest municipal border. We then calculate the two vectors orthogonal to this vector of length 1km, and originating at the intersection of the vector and the border. If both of these orthogonal vectors lie within 75 meters of the municipal boundary, we say that the parcel is associated with a straight-line boundary. Figure 3 illustrates this algorithm.[9]

Applying this algorithm to our data, we find that of the 9798 transactions for which we have data on regulation, 2124 match to straight borders. These 2124 parcels lie in 185 distinct municipalities and provide information about the land price gradient near 190 municipal border pairs.

Table 1 describes the parcels in our sample. Column 1 describes the set of all parcel transactions present in the COSTAR data after we restrict attention to the 1990 to 2009 period and drop transaction without price data. Column 2 further restricts attention to transactions of parcels in WRLURI municipalities and adjacent to WRLURI municipalities, our main sample. Columns 3 restricts attention to parcels in our main sample which match to straight-line boundaries.

The sample of parcels for which we have regulation information is almost certainly not representative of the universe of parcels sampled by COSTAR. Restricting attention to parcels for which WRLURI data is available laves us with parcels that are smaller, more expensive per square foot and closer to the metropolitan region's tallest building. Comparing columns two and three, we see that sample parcels matched to straight boundaries tend to be larger, and marginally more remote and less expensive than an average sample parcel, although all differences are small compared to standard deviations.

Transaction weighted mean municipal characteristics of straight boundary municipalities are also similar to average sampled municipalities. Incomes and share of college educated population are marginally higher in the smaller sample, but the differences are quite small.

[8] We also experimented with excluding municipal boundaries which are also water features from our sample. Such parcels are separated from their cross-border counterparts by a river, and hence it is at least possible that the two relevant municipalities occupy land with discretely different characteristics. However, so few of the borders in our sample were defined by water that we abandoned this exercise.

[9] Almost the entire area of the continental US outside of the original 13 colonies was surveyed in accordance with the Land Ordinance Act of 1785. This act required that nearly all federal lands be surveys and divided into regular 'sections' as a precursor to their eventual settlement. See Libecap and Lueck (2009) for details.

Table 1. Summary statistics for transaction data.

Variable	All	All with WRLURI	All straight with WRLURI
Price ($\$/ft^2$)	29.63	77.34	68.13
	(623.97)	(287.92)	(347.89)
Size (0000 ft^2)	132.31	33.69	45.51
	(1,483.42)	(172.93)	(186.30)
log(Distance to tallest building)	3.43	3.05	3.40
	(0.90)	(1.09)	(1.10)
Share college	0.12	0.12	0.13
	(0.06)	(0.05)	(.06)
Median income	57,760	55,399	58,412
	(18,036)	(18,069)	(18,750)
WRLURI		0.47	0.59
		(0.94)	(0.92)
LPPI		0.67	0.68
		(1.35)	(1.15)
LZAI		1.83	1.79
		(0.70)	(0.59)
LPAI		1.40	1.54
		(0.94)	(0.82)
SRI		0.07	0.06
		(0.44)	(0.40)
DRI		0.17	0.23
		(0.38)	(0.42)
EI		0.70	0.67
		(0.46)	(0.47)
ADI		8.07	8.25
		(3.94)	(3.92)
OSI		.620	0.63
		(0.49)	(0.48)
# transactions	136,961	9,798	2,142
# WRLURI municipalities		513	185
# WRLURI borders		598	190
# school districts		424	182

The WRLURI land use regulation index is constructed from eleven subindexes. Of these, two vary only at the state level and one is relevant only two New England. The eight remaining indices are: the *Local Political Pressure index (*LPPI*)* which is increasing in the propensity of survey respondents to state that local actors are important in the regulation process; the *Local Zoning Approval Index(*LZAI*)* ranges from 0-6 and is the count of the number of entities that must approve a zoning change, the *Local Project Approval Index (*LPAI*)* is analogous to the LZAI but gives the count, from 0-6 of entities that must approve a project, the *Density Restriction Index (*DRI*)* is zero if minimum lot size is less than one acre and one otherwise, the *Exactions Index (*EI*)* is one if the municipality mandates exactions to cover infrastructure costs, the *Approval Delay Index (*ADI*)* reflects survey respondents' statements about the length of time require to get administrative approval for development, the *Open Space Index (*OSI*)* is an indicator variable that is one if a municipality has open space set aside requirements.

In table 1 we see that the WRLURI index is marginally higher in the sample of straight boundaries, though the difference is small compared to both the standard deviations and to the range of the index (from -1.6 to 3.9 in the main sample). We also see that the means of various subindexes are almost identical across the two samples.

Results

Own lot effect regressions

Table 2 reports the results of our estimates of own-lot effects regressions. Each cell of this table reports the results of an estimate of one of our own-lot effect estimating equations (equations 11-14). For a particular specification each cell of the table reports the coefficient B_0 for the WRLURI index, its standard error in parentheses and the number of parcels and borders on which the estimate is based. The table consists of three panels. The top panel reports results based on the entire sample of parcels and borders. In the middle panel we report results based on straight-line borders and parcels that match to these borders. The bottom panel indicates the control variables used for all regressions in each column.

As we move between rows in either of the top two panels we change the sample of parcels used to estimate all regressions in the row. In the top row of each of the top two panels, we use all parcels that match to each border, that is, all parcels within 100 km of the relevant border. In the second row we restrict attention to parcels within 1km of a municipal border. Moving down the remaining three rows we restrict attention to parcels within 500m, 250m and 100m of a border. As expected, the number or parcels and number of borders declines as we move from row one to row five of the table.

In the first five columns of the table we consider transactions which occurred after the last quarter of 1999 and before the second quarter of 2009. In the sixth column we restrict attention to the period 2002Q1-2009Q1. In the seventh column we extend the sample to the period 1990Q1-2009Q1. Note from the first row of column 7 that in this longer sample we increase our sample size to 14,981 parcels and 632 borders. Since the Wharton Index is based on a 2005 survey we would ideally restrict attention to transactions that occurred after this date. However, our sample is too

	[2000-9]	[2000-9]	[2000-9]	[2000-9]	[2000-9]	[2002-9]	[1990-2009]
All							
dist < 100	2.144	-6.136	-6.326	11.594	-7.162	-8.061	-3.654
	(10.227)	(2.974)**	(4.695)	(15.822)	(2.360)**	(3.066)**	(1.727)**
	9798/598	9731/597	9798/598	9798/598	9731/597	6400/520	14887/630
dist < 1	-2.808	-1.351	-2.26	-3.408	-0.674	1.581	1.007
	(1.799)	(2.387)	(1.844)	(1.616)**	(2.369)	-3.222	(2.15)
	4412/547	4383/545	4412/547	4412/547	4383/545	2730/454	6979/580
dist < .5	-5.999	-5.019	-8.28	-1.657	-6.489	-8.472	-3.193
	(3.210)*	(2.182)**	(3.655)**	(1.219)	(2.301)**	(3.588)**	(1.108)**
	2870/480	2861/479	2870/480	2870/480	2861/479	1778/387	4522/519
dist < .25	-8.884	-5.853	-14.364	-1.640	-9.562	-10.561	-5.439
	(6.063)	(3.154)*	(7.547)*	(1.364)	(4.058)**	(5.415)*	(2.211)**
	1872/403	1867/402	1872/403	1872/403	1867/402	1163/323	2963/446
dist < .1	-13.649	-6.213	-23.259	-3.641	-11.470	-17.94	-5.731
	(10.677)	(3.668)*	(13.293)*	(2.437)	(6.173)*	(10.083)*	(2.897)**
	1069/303	1066/303	1069/303	1069/303	1066/303	642/230	1688/356
Straight							
dist < 100	14.302	1.519	3.078	16.299	4.101	9.265	-0.192
	(13.290)	(9.057)	(10.957)	(20.274)	(9.564)	-14.698	(6.429)
	2142/190	2141/190	2142/190	2142/190	2141/190	1409/166	3178/207
dist < 1	-2.967	-2.492	-2.415	-6.803	-2.703	-3.728	-0.594
	(3.083)	(2.859)	(3.352)	(3.710)*	(3.348)	-4.856	(1.995)
	850/153	850/153	850/153	850/153	850/153	540/131	1229/175
dist < .5	-0.826	-0.908	-0.798	-4.313	-0.907	-3.117	-1.075
	(2.243)	(2.277)	(2.287)	(2.181)**	(2.281)	-3.397	(1.885)
	551/121	551/121	551/121	551/121	551/121	351/104	806/137
dist < .25	1.283	1.433	1.092	-1.675	0.987	2.053	0.366
	(2.248)	(2.227)	(2.659)	(1.835)	(2.664)	-3.653	(2.255)
	388/100	388/100	388/100	388/100	388/100	239/85	575/113
dist < .1	-2.445	-3.036	-4.09	-3.032	-4.979	-4.447	-2.953
	(2.120)	(2.617)	(1.675)**	(3.094)	(1.880)**	-3.943	(1.255)**
	252/74	252/74	252/74	252/74	252/74	152/61	386/86
Border pair FE	Y	Y	Y	Y	Y	Y	Y
Parcel controls I	Y	Y	Y	Y	Y	Y	Y
Demographics			Y		Y	Y	Y
School district				Y			
Parcel controls II		Y			Y	Y	Y

Table 2. Own lot effect regression results for WRLURI index. Each cell reports results from a different regression. **Dependent variable** is Parcel price per square foot. First row of each cell reports coefficient on cross-border change in regulatory index. Second row of each cell reports standard errors in parentheses. Standard errors are clustered by border pair. Third row of each cell reports the number of observations and borders on which the estimate is based. **Parcel controls I** are (parcel ft^2), (parcel ft^2)^2, log of distance to CBD, (log of distance to CBD)^2, quarterly indicators, mean ruggedness of surrounding 0.5km radius disk, mean ruggedness of surrounding 5km radius disk, mean ruggedness of surrounding 10km radius disk. **Demographics** are municipal share black, share asian, share of population under age 17, number of households, share with high school degrees, share with four year degree, median income. **Parcel controls II** are total employment from 1994 zip code business patterns in the 1km, 5km and 10km disk surrounding the parcel and the share of developed land in the corresponding disks from 1992 satellite images. **School district** is an indicator for each secondary or amalgamated school district computed from 2000 census boundary files.

	[LPPI]	[LZAI]	[LPAI]	[SRI]	[DRI]	[EI]	[ADI]	[OSI]
All								
dist < 100	-2.078	-4.163	2.713	-0.221	-4.930	-9.185	-0.895	-5.272
	(1.054)**	(2.030)**	(1.805)	(2.297)	(4.470)	(4.906)*	(0.443)**	(3.716)
	9957/622	9957/622	9905/622	9948/622	9957/622	9803/599	9880/609	9804/599
dist < 1	0.069	-1.432	1.368	-3.068	-3.701	-4.29	1.311	-10.251
	(0.748)	(1.604)	(1.224)	(2.300)	(4.096)	(6.742)	(0.710)*	(6.452)
	4505/568	4505/568	4457/567	4499/568	4505/568	4446/548	4461/554	4447/548
dist < .5	-1.986	-1.186	-2.095	-0.300	-7.64	-4.346	-0.561	-0.269
	(0.877)**	(1.114)	(1.187)*	(0.717)	(3.125)**	(4.149)	(0.378)	(2.146)
	2942/498	2942/498	2914/497	2936/498	2942/498	2904/482	2910/486	2905/482
dist < .25	-3.792	-3.109	-3.322	-0.068	-8.241	-8.584	-0.801	4.655
	(1.499)**	(1.786)*	(1.724)*	(1.541)	(4.355)*	(5.763)	(0.622)	(4.667)
	1923/419	1923/419	1905/418	1919/419	1923/419	1896/405	1901/409	1896/405
dist < .1	-4.034	-4.294	-4.052	-0.669	-8.606	-10.469	-1.094	4.131
	(2.052)**	(2.035)**	(1.797)**	(1.879)	(6.735)	(8.087)	(0.899)	(4.174)
	1086/310	1086/310	1079/309	1082/310	1086/310	1079/305	1080/306	1079/305
Straight								
dist < 100	1.446	-10.729	6.556	15.839	-10.428	1.053	0.150	-3.459
	(3.519)	(6.89)	(5.468)	(10.111)	(8.506)	(12.327)	(1.65)	(12.706)
	2202/200	2202/200	2196/199	2202/200	2202/200	2149/192	2170/196	2149/192
dist < 1	-1.387	1.525	-2.372	-1.385	7.563	1.057	-0.614	3.554
	(1.235)	(2.466)	(1.961)	(2.117)	(4.665)	(4.413)	(0.615)	(3.573)
	867/161	867/161	864/160	867/161	867/161	853/154	855/155	853/154
dist < .5	-0.848	-1.569	-1.373	-1.310	2.156	-1.908	0.063	2.594
	(1.028)	(2.42)	(1.091)	(0.938)	(3.055)	(4.376)	(0.354)	(4.449)
	562/128	562/128	560/127	562/128	562/128	553/122	554/123	553/122
dist < .25	-0.589	-2.13	-2.126	-1.837	3.794	0.548	0.191	0.628
	(0.850)	(2.369)	(1.64)	(2.754)	(3.81)	(3.823)	(0.324)	(6.465)
	394/105	394/105	394/105	394/105	394/105	388/100	389/101	388/100
dist < .1	-2.720	1.376	-3.885	1.066	0.983	0.964	-0.269	-5.257
	(1.026)**	(2.868)	(1.879)**	(4.936)	(2.976)	(4.617)	(0.510)	(2.382)**
	254/76	254/76	254/76	254/76	254/76	252/74	252/74	252/74
Border pair FE	Y	Y	Y	Y	Y	Y	Y	
Parcel controls I	Y	Y	Y	Y	Y	Y	Y	
Demographics	Y	Y	Y	Y	Y	Y	Y	
Parcel controls II	Y	Y	Y	Y	Y	Y	Y	

Table 3. Own lot effect regression results for WRLURI subindexes. Each cell reports results from a different regression. **Dependent variable** is Parcel price per square foot. First row of each cell reports coefficient on cross-border change in regulatory index. Second row of each cell reports standard errors in parentheses. Standard errors are clustered by border pair. Third row of each cell reports the number of observations and borders on which the estimate is based. **Parcel controls I** are (parcel ft^2), (parcel $ft^2)^2$, log of distance to CBD, (log of distance to CBD$)^2$, quarterly indicators, mean ruggedness of surrounding 0.5km radius disk, mean ruggedness of surrounding 5km radius disk, mean ruggedness of surrounding 10km radius disk. **Demographics** are municipal share black, share asian, number of households, share of population under age 17, share with high school degrees, share with four year degree, median income.

small to allow this. By comparing columns 5-7 of table 4 we see that our estimates are qualitatively similar for the different time periods, so we focus our attention on the period 2000Q1-2009Q1. This period is close enough to the Wharton survey date that we should expect the survey to be informative about municipal regulations, and is long enough to take in most of the transactions that we can match to WRLURI municipalities.

All of the regressions reported in this table include a municipal border fixed effect and quarterly indicator variables.

The first five columns of table 2 are based on transactions during 2000Q1-2009Q1. Each estimates a different variant of the own-lot effect estimating equation. As we move from left to right we control for more possible sources of confounding variation. In the first column, in addition to a border pair fixed effect we include our first set of parcel characteristics as control variables. These parcel characteristics are the number of square feet of the parcel and the square of this quantity, the log of the distance to the tallest building in the metropolitan area and the square of this quantity and indicator variables describing the quarter in which the transaction took place. All of these variables are part of the COSTAR data set. In addition, for each parcel, the first set of parcel controls describes the ruggedness of the surrounding 0.5km, 5km and 10km radius disk centered on the parcel. These data are based on data used in Eid, Overman, Puga, and Turner (2008) and are described there.

The second column of table 2 adds our second set of parcel controls. For each parcel, these controls describe the total employment from 1994 zip code business patterns in the 1km,5km and 10km disk surrounding the parcel and the share of developed land in the corresponding disks from 1992 satellite images. These data are based on data used in Eid *et al.* (2008) and are described there.

In the third column of table 2 we include our first set of demographic controls and controls for municipal demographic characteristics. These demographic characteristics are; share black, share asian, number of households, share with high school degrees, share with four year degree, median income. All are calculated for our sample of municipalities from the 2000 census.

The fourth column of table 4 use our first set of parcel controls and a set of indicator variables, one for each secondary or amalgamated school district, computed from 2000 census boundary files. Finally, column 5 includes both out sets of parcel controls and our municipal demographic controls.

In column 6 and 7 we duplicate the regressions of column 3, but use a different sample. In the sixth column we restrict attention. during the period 2002Q1-2009Q1. In the seventh column we extend the sample to the period 1990Q1-2009Q1. Since the Wharton Index is based on a 2005 survey we would ideally restrict attention to transactions that occurred after this date. However, our sample is too small to allow this. By comparing columns 5-7 of table 4 we see that our estimates are qualitatively similar for the different time periods, so we focus our attention on the period 2000Q1-2009Q1.

For both the sample of all borders and the sample of exogenous straight borders, the results of table 2 are unambiguous. The own-lot effect has the expected negative sign. In the top panel of the table the estimated effect of regulation is negative in almost all of the 35 specifications presented. As we move down the rows and consider parcels progressively closer to the border, we more and more nearly approximate the identification strategy described in section 1. In the last row of the top panel, which considers only parcels within 100m of a border, every estimate of the effects of regulation is negative and the errors of these estimates are small enough to allow us to distinguish the estimates from zero at standard levels of confidence in all but two of the seven cases (and these two are close). The bottom panel of table 2 is similar, but estimates are generally less precise. Given the much smaller sample this is not surprising.

Two differences between the top and bottom panels deserve mention. First, where the bottom row of the top panel reports all negative estimates for the own-lot effect, the magnitude of these estimates varies dramatically from one specification to the other. In the bottom panel, where we restrict attention to straight borders, estimates are also negative, but are smaller and estimates do not vary as dramatically across specifications. Second, if we accept that the land on both sides of a straight border is the same in its intrinsic attractiveness, that the estimates of the own-lot effect are smaller in magnitude in the straight borders sample requires that on average the intrinsic attractiveness of land fall discretely when we cross from a less regulated to a more regulated municipality. This is contrary to our prior that more attractive places should be regulated more intensively, but is consistent with table 1 and with Gyourko *et al.* (2008) which suggest that more remote and less valuable land is regulated more intensively.

Finally, from table 1 we see that the standard deviation of the WRLURI index is about 0.9 in sample of straight-line borders, while the average price per square foot of land is about \$68. From the last row of the second panel of table 4 we see that the effect of a one unit change in the WRLURI index on the price of land is about \$4. Multiplying by 0.9 and dividing by \$68, we have that a one standard deviation increase in the WRLURI index decreases the value of an average parcel by about 5%.

In order to assess whether different types of regulation are more or less costly, table 3 duplicates the regressions performed in column 5 of table 2 for each of the eight subindexes of regulation available in the Wharton survey. The structure of this table is similar to that of table 2. In particular, the top panel examines our main sample of borders while the second panel examines straight-line border and moving down rows in either panel restricts attention to parcels close to a municipal border. However, unlike table 2, as we move across columns control variables and samples stay constant while the measure of regulation varies (as noted in the column headings). Looking at the last row of the second panel suggests that increases in Local Political Pressure (LPPI), in the Local Project Approval Index (LPAI) which counts the number of entities which must approve a project, and the Open Space Index (OSI) are responsible for the costs of regulation, while other sorts of local regulation impose cost which are too small to measure in our data.

External effect regressions

Table 4 reports the results of our estimates of external effects regressions. Each cell of this table reports the results of an estimate of a variant of equation 18 or 20. In particular, for a particular specification each cell of the table reports the coefficient B_0 for the WRLURI index, its standard error in parentheses and the number of parcels and borders on which the estimate is based. The table consists of three panels. The top panel reports results based on the entire sample or parcels

	[2000-9]	[2000-9]	[2000-9]	[2000-9]	[2000-9]	[2002-9]	[1990-09]	[2000-9(res.)]
All								
1>x>.5,.5>x>0	6.215	8.434	8.506	7.793	8.4	15.257	2.961	20.924
	(5.121)	(6.007)	(6.012)	(5.833)	(5.841)	(8.800)*	(2.224)	(11.408)*
	4302/799	4302/799	4302/799	4273/797	4273/797	2668/647	6808/870	1167/436
1>x>.5,.25>x>0	10.417	14.317	14.29	12.816	13.426	21.04	6.137	20.284
	(6.748)	(6.730)**	(6.730)**	(6.492)**	(6.206)**	(8.619)**	(3.331)*	(13.595)
	3331/741	3331/741	3331/741	3306/739	3306/739	2070/593	5293/813	905/383
1>x>.5,.1>x>0	13.204	18.731	18.81	15.86	16.673	27.202	10.008	7.046
	(8.496)	(8.366)**	(8.425)**	(7.666)**	(7.223)**	(8.729)**	(5.504)*	(15.649)
	2549/656	2549/656	2549/656	2526/654	2526/654	1562/515	4053/737	698/319
.5>x>.25,.25>x>0	3.749	7.24	7.065	5.213	5.328	4.716	4.577	-3.618
	(4.49)	(6.76)	(6.561)	(4.437)	(4.547)	(5.123)	(3.996)	(5.061)
	2799/680	2799/680	2799/680	2790/679	2790/679	1738/534	4410/759	717/334
.5>x>.25,.1>x>0	6.683	11.628	11.601	6.099	6.326	6.03	7.205	3.006
	(7.718)	(11.258)	(11.102)	(4.861)	(5.019)	(4.856)	(6.618)	(9.96)
	2017/592	2017/592	2017/592	2010/592	2010/592	1230/447	3170/683	510/276
Straight								
1>x>.5,.5>x>0	0.900	-0.095	-0.081	0.212	-0.479	3.504	0.187	-1.920
	(2.000)	(2.438)	(2.466)	(2.418)	(2.322)	(4.866)	(1.477)	(3.200)
	831/204	831/204	831/204	831/204	831/204	527/174	1203/234	220/103
1>x>.5,.25>x>0	0.584	-1.08	-1.082	0.089	-0.815	2.651	-0.44	-1.242
	(2.504)	(2.758)	(2.753)	(3.354)	(3.297)	(6.732)	(2.061)	(12.421)
	670/185	670/185	670/185	670/185	670/185	417/158	974/213	170/88
1>x>.5,.1>x>0	-3.343	-4.907	-4.759	-4.361	-5.134	-7.463	-2.205	-156.977
	(2.642)	(4.614)	(4.641)	(5.014)	(5.806)	(12.934)	(3.749)	(23.241)**
	539/167	539/167	539/167	539/167	539/167	335/139	791/196	145/81
.5>x>.25,.25>x>0	-3.156	-3.349	-3.569	-3.501	-3.787	-6.739	-2.731	-6000.000
	(1.332)**	(1.457)**	(1.432)**	(1.678)**	(1.846)**	(2.849)**	(1.416)*	(3951.162)
	539/158	539/158	539/158	539/158	539/158	342/134	791/180	128/67
.5>x>.25,.1>x>0	-3.501	-3.507	-3.781	-3.739	-3.93	-7.1	-3.704	-31000.000
	(1.485)**	(1.371)**	(1.396)**	(1.902)**	(2.113)*	(2.806)**	(1.496)**	(25000.000)
	408/133	408/133	408/133	408/133	408/133	260/114	608/154	103/59
Muni.-Border FE	Y	Y	Y	Y	Y	Y	Y	Y
Parcel controls I	Y	Y	Y	Y	Y	Y	Y	Y
Δ Demographics		Y	Y	Y	Y	Y	Y	Y
Interior= 1			Y	Y	Y	Y	Y	Y
Parcel controls II				Y	Y	Y	Y	Y
School district					Y	Y	Y	Y

Table 4. External effect regression results for WRLURI index or regulation. Each cell reports results from a different regression. **Dependent variable** is Parcel price per square foot. First row of each cell reports coefficient on cross-border change in the regulatory index. Second row of each cell reports standard errors in parentheses. Standard errors are clustered by municipality and border. Third row of each cell reports the number of observations and borders on which the estimate is based. **Parcel controls I** are (parcel ft^2), (parcel $ft^2)^2$, log of distance to CBD, (log of distance to CBD)^2, quarterly indicators, mean ruggedness of surrounding 0.5km radius disk, mean ruggedness of surrounding 5km radius disk, mean ruggedness of surrounding 10km radius disk.Δ **Demographics** are cross border changes in municipal share black, share asian, number of households, share with high school degrees, share with four year degree, median income. **Interior=1** is an indicator that is 1 for interior parcels and zero otherwise. **Parcel controls II** are total employment from 1994 zip code business patterns in the 1km,5km and 10km disk surrounding the parcel and the share of developed land in the corresponding disks from 1992 satellite images. **School district** is an indicator for each secondary or amalgamated school district computed from 2000 census boundary files.

and borders. In the middle panel we report results based on straight-line borders and parcels that match to these borders. The bottom panel indicates the control variables used for all regressions in each column.

As we move between rows in either of the top two panels we change the definitions of 'interior' and 'peripheral' bins upon which the regression is based. In the top row of the top panel, the peripheral bin extends from the border to 500m from the border while the interior bin extends from 500m to 1000m from the border. In the second row the peripheral bin extends from the border to 250m from the border. In the third row the peripheral bin extends from the border to 100m from the border. In the fourth row the peripheral bin extends from the border to 250m from the border while the interior bin extends from 250m to 500m from the border. In the bottom row, row the peripheral bin extends from the border to 100m from the border. In each regression we restrict attention to parcels which fall into one of the two bins and construct an indicator that is one for the interior bin and zero otherwise. This indicator corresponds exactly to the indicator variable χ_i^{Imj} in equations 18 and 20. As expected, the number or parcels and number of borders on which our estimates are based declines as we move from row one to row five of the table.

In the first five columns of the table we consider transactions which occurred after the last quarter of 1999 and before the second quarter of 2009. In the sixth column we restrict attention to the period 2002Q1-2009Q1. In the seventh column we extend the sample to the period 1990Q1-2009Q1. Comparing columns 5-7 of table 4 we see that our estimates are qualitatively similar for the different time periods, so we focus our attention on the period 2000Q1-2009Q1. Column 8 of table 4 is based on the same 2000Q1-2009Q1 period as used in columns 1-5, but restricts attention to parcels for which the buyer reported the intention to develop the parcel for residential use. Our intention here to investigate whether regulation affects residential property differently than non-residential property.

Each of the first five columns of table 4 are based on transactions during 2000Q1-2009Q1, but estimates a different variant of equations external effect estimating equation. As we move from left to right we control for more possible sources of confounding variation. In the first column, in addition to a border pair fixed effect we include our first set of parcel characteristics as control variables. These parcel characteristics are the number of square feet of the parcel and the square of this quantity, the log of the distance to the tallest building in the metropolitan area and the square of this quantity and indicator variables describing the quarter in which the transaction took place. All of these variables are part of the COSTAR data set. In addition, for each parcel, the first set of parcel controls describes the ruggedness of the surrounding 0.5km, 5km and 10km radius disk centered on the parcel.

The second column of table 4 also controls for cross-border changes in municipal demographic characteristics as described in equation 20. These demographic characteristics are; share black, share asian, number of households, share with high school degrees, share with four year degree, median income.

In the third column of table 4 we also include and indicator variable that is one for all interior parcels and zero otherwise. The object of this variable is to control for the possibility that the edges of municipalities are systematically less nice than their interiors, independent of regulation.

The fourth column of table 4 adds our second set of parcel controls. For each parcel, these controls describe the total employment from 1994 zip code business patterns in the 1km, 5km and 10km disk surrounding the parcel and the share of developed land in the corresponding disks from 1992 satellite images. These data are based on data used in Eid *et al.* (2008) and are described there. Finally, column five of table 4 is our preferred specification and also includes indicator variable for each secondary or amalgamated school district computed from 2000 census boundary files. Columns 6-8 are based on the same specification as is used in column 5 but different samples.

The results reported in table 4 are striking. For example, in the second column of row three in the top panel, we see that a one unit cross border increase in the WRLURI index is associated with a 18.02$ change in the price per square foot of an average parcel. In fact, this is a very large change. The sample standard deviation of the WRLURI index is 0.47 and the mean price per square foot is $77.34. Thus, a one standard deviation change in the WRLURI index is associated with an 11% change in the price of land as we move from parcels within 250m of a border to one between 500 and 1000m from the interior. This effect is statistically different from zero at standard levels of confidence and this effect persists for each of the five specifications and three samples examined in the first seven columns of table 4. Each of the five rows in the top panel of table 4 tells a similar story, although except for row 2, coefficients tend to be smaller and to have less statistical significance. In sum, the top panel of table 4 suggests that regulation has a significant positive external effect on land prices.

In column 8 of the top panel of 4 we restrict attention to parcels reported to be for residential use. While coefficients here are generally positive (and in row one significant) these estimates are imprecise. Bearing in mind that our estimation strategy requires that we estimate one fixed effect per border, our inclination is to interpret this column as an indication that we do not have enough data to distinguish the effects of regulation on residential from its effects on non-residential land.

In the bottom panel of table 4 we duplicate the regressions reported in the top panel, but restrict attention to straight-line borders and parcels which match to these borders. This panel tells exactly the opposite story as the top panel. There is not a single specification or sample with a statistically significant positive external effect of regulation on land prices. On the contrary. To the contrary, the preponderance of specifications yield negative point estimates for value of the external effect of land regulation and many of these, rows four and five in particular, are statistically significant at standard levels of confidence.

To understand these results recall that the central problem with which we are concerned is the simultaneous determination of land use regulation and land prices. We are concerned that land prices and land regulation are determined by the same unobserved factor. To overcome this problem, our analysis relies on municipal borders which separate similar land into distinct jurisdictions. When we consider the set of all municipal borders, we should be concerned that some of these borders follow landscape features which separate one type of place from another. When we consider straight-line borders, many of which likely date from the first federal surveys of the continent, this is less of a concern. These two observations, together with the different results presented in top and bottom panels of table 4 suggest that, in fact, irregular municipal borders are not exogenous to the process that determines land prices and regulation. Thus, in the sample of all borders we observe a positive external effect of regulation, while the sample of straight borders suggests that the actual external effect of regulation is negative.

That is, if the bottom panel of table 4 correctly isolates the causal effect of regulation, then regulation restricts development of land in a way that is harmful to nearby parcels. Since table 2 shows that regulation also has a negative own-lot effect, this suggests that changes to municipal land use regulation which reduce the WRLURI index are beneficial.

To investigate this further, table 5 duplicates the estimations of column 5 of table 4 for each of these seven subindexes. Except for the Density Restriction Index, each of the regulation sub-indices appears to have positive external effects on the whole sample but these effects are negative or not distinguishable from zero. Only the density restrictions index has a statistically significant external effect on the sample of straight borders. That is, restrictions that limit density appear to benefit surrounding land. Conversely, at the margin a reduction in the other sorts of planning and regulatory effort measured by the various indices is probably beneficial.

The density restriction index is an indicator variable which is one if a municipality has a minimum lot size greater than an acre. From the last row of the DRI column in the bottom panel of table 5 we see that about 250 meters interior from the boundary of a municipality which has a minimum lot size but adjoins one that does not results in an increase in land prices of about $8 per square foot, or about 12% of the value of the land. From footnote 6 we calculate the total change in land value for an average interior parcel as twice estimated the external effect minus the own-lot effect. From the preceding calculation, twice the estimated external effect is $16 per square foot or 24% of the price. From table 3 we see that the estimated own-lot effect for DRI is not distinguishable from zero for any sample and the largest point estimate is less than $11. Thus our point estimate of the benefits of an increase in the minimum lot size is positive.

Conclusion

We estimate the effect of land use regulation on the value of land by exploiting variation in exposure to municipal land use regulation across municipal boundaries. Our estimations are based on a novel theoretical and empirical framework which allows us to separately estimate own-lot and external effects of regulation by looking at variation in land prices and land use regulation across and near municipal boundaries. To implement these estimations we assemble data describing a large sample of land transactions and municipalities and implement an algorithm to identify parcel close to municipal borders which do not follow landscape features.

Since the value of land gives us the market's measure of the attractiveness of a location, our estimates allow us to draw conclusions about the welfare implication of land use regulation. Our results do not support the proposition that municipal land use regulation is generally welfare improving, although minimum lot size regulation may be. With this said, we note that these conclusions are ultimately based on fairly small samples of transactions and municipalities so that some uncertainty remains.

	[LPPI]	[LZAI]	[LPAI]	[SRI]	[DRI]	[EI]	[ADI]	[OSI]
All								
1>x>.5,.5>x>0	1.096	6.777	5.504	0.799	9.409	6.495	1.503	-18.57
	-1.276	(3.680)*	(3.364)*	(2.105)	(5.357)*	(8.724)	(0.944)	(11.029)*
	4502/870	4502/870	4409/861	4491/864	4502/870	4396/816	4419/830	1206/449
1>x>.5,.25>x>0	2.564	8.617	6.856	2.423	12.424	17.097	2.121	-15.957
	-1.613	(4.410)*	(3.304)**	(3.151)	(6.474)*	(10.097)*	(1.045)**	(19.678)
	3484/801	3484/801	3412/792	3476/795	3484/801	3398/756	3422/771	930/392
1>x>.5,.1>x>0	3.803	8.591	6.517	4.28	13.164	29.005	2.653	-29.683
	(2.218)*	(5.100)*	(3.732)*	(3.858)	(7.855)*	(13.443)**	(1.235)**	(35.511)
	2647/703	2647/703	2595/695	2639/697	2647/703	2590/669	2611/682	716/326
.5>x>.25,.25>x>0	1.727	4.034	2.974	0.916	2.716	9.921	0.714	-0.16
	(1.373)	(2.598)	(2.128)	(1.44)	(3.707)	(5.509)*	(0.781)	(9.149)
	2939/736	2939/736	2884/727	2930/732	2939/736	2873/696	2877/704	744/345
.5>x>.25,.1>x>0	2.051	4.306	3.414	1.052	3.573	14.552	0.91	10.708
	(1.703)	(3.07)	(2.263)	(1.187)	(4.563)	(8.901)*	(0.727)	(19.634)
	2102/634	2102/634	2067/625	2093/630	2102/634	2065/608	2066/611	530/286
Straight								
1>x>.5,.5>x>0	-0.745	-0.700	-1.393	2.146	10.056	-4.399	-0.674	-1.389
	(1.251)	(3.361)	(1.248)	(3.2650	(5.132)*	(5.314)	(0.378)*	(2.369)
	867/226	867/226	856/223	867/226	867/226	842/207	850/213	842/207
1>x>.5,.25>x>0	-1.141	0.620	-2.66	-3.326	14.673	-6.933	-0.523	-1.839
	(1.477)	(3.497)	(2.357)	(4.321)	(7.861)*	(7.04)	(0.604)	(4.01)
	699/204	699/204	691/202	699/204	699/204	678/187	686/193	678/187
1>x>.5,.1>x>0	-3.33	-0.387	-1.718	-5.557	14.449	-7.039	-1.001	-4.711
	(3.548)	(5.092)	(2.374)	(7.516)	(10.282)	(13.107)	(0.878)	(5.464)
	559/181	559/181	554/179	559/181	559/181	544/169	550/174	544/169
.5>x>.25,.25>x>0	-2.916	1.679	0.213	-0.77	4.64	1.2	-0.817	-3.153
	(1.063)**	(1.557)	(1.072)	(2.15)	(2.905)	(2.898)	(0.365)**	(1.724)*
	562/174	562/174	556/171	562/174	562/174	545/161	549/164	545/161
.5>x>.25,.1>x>0	-2.334	1.53	-0.357	-5.308	8.114	0.098	-0.744	-2.317
	(1.253)*	(1.787)	(1.245)	(1.968)**	(3.764)**	(3.73)	(0.441)*	(2.687)
	422/145	422/145	419/142	422/145	422/145	411/136	413/138	411/136
Muni.-Border FE	Y	Y	Y	Y	Y	Y	Y	Y
Parcel controls I	Y	Y	Y	Y	Y	Y	Y	Y
Δ Demographics	Y	Y	Y	Y	Y	Y	Y	Y
Interior= 1	Y	Y	Y	Y	Y	Y	Y	Y
Parcel controls II	Y	Y	Y	Y	Y	Y	Y	Y
School district	Y	Y	Y	Y	Y	Y	Y	Y

Table 5. External effect regression results for WRLURI subindexes. Each cell reports results from a different regression. **Dependent variable** is Parcel price per square foot. First row of each cell reports coefficient on cross-border change in regulatory index. Second row of each cell reports standard errors in parentheses. Standard errors are clustered by municipality and border. Third row of each cell reports the number of observations and borders on which the estimate is based. **Parcel controls I** are (parcel ft^2), (parcel $ft^2)^2$, log of distance to CBD, (log of distance to CBD$)^2$, quarterly indicators, mean ruggedness of surrounding 0.5km radius disk, mean ruggedness of surrounding 5km radius disk, mean ruggedness of surrounding 10km radius disk.Δ **Demographics** are cross border changes in municipal share black, share asian, number of households, share with high school degrees, share with four year degree, median income. **Interior=1** is an indicator that is 1 for interior parcels and zero otherwise. **Parcel controls II** are total employment from 1994 zip code business patterns in the 1km,5km and 10km disk surrounding the parcel and the share of developed land in the corresponding disks from 1992 satellite images. **School district** is an indicator for each secondary or amalgamated school district computed from 2000 census boundary files.

References

Angrist, J. D. and V. Lavy. 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics* 114(2):533–575.

Bayer, P., F. Ferreira, and R. McMillan. 2007. A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy* 115(4):588–637.

Black, S. E. 1999. Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics* 114(2):577–599.

Burchfield, Marcy, Henry G. Overman, Diego Puga, and Matthew A. Turner. 2006. Causes of sprawl: A portrait from space. *Quarterly Journal of Economics* 121(2):587–633.

Cheshire, Paul and Stephen Sheppard. 2002. The welfare economics of land use planning. *Journal of Urban Economics* 52(2):242–269.

Duranton, G., L. Gobillon, and H. G. Overman. 2008. Assessing the effects of local taxation using microgeographic data. Processed, University of Toronto.

Eid, Jean, Henry G. Overman, Diego Puga, and Matthew A. Turner. 2008. Fat city: Questioning the relationship between urban sprawl and obesity. *Journal of Urban Economics* 63:385–404.

Ferreira, F. and J. Gyourko. 2009. Do political parties matter? Evidence from US cities. *Quarterly Journal of Economics* 124(1):399–422.

Glaeser, E. L., J. Gyourko, and R. Saks. 2005. Why is Manhattan so expensive? Regulation and the rise in housing prices. *Journal of Law and Economics* XLVII:331–69.

Gyourko, Joseph, Albert Saiz, and Anita A. Summers. 2008. A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies* 45:693–729.

Hahn, J. Y., P. Todd, and W. Van der Klaauw. 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69(1):201–209.

Hilber, C. A. L. and F Robert-Nicoud. 2009. On the origins of land use regulations: The 'influential landowner' hypothesis. *Processed, London School of Economics* .

Holmes, T. J. 1998. The effect of state policies on the location of manufacturing: Evidence from state borders. *Journal of Political Economy* 106(4):667–705.

Ihlanfeldt, K. 2007. The effect of land use regulation on housing and land prices. *Journal of Urban Economics* 61:420–35.

Imbens, Guido W. and Thomas Lemieux. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142(2):615–635.

Irwin, E. G. and N. E. Bockstael. 2002. Interacting agents, spatial externalities and the evolution of residential land use patterns. *Journal of Economic Geography* 2(1):31–54.

Lemieux, Thomas and Kevin Milligan. 2008. Incentive effects of social assistance: A regression discontinuity approach. *Journal of Econometrics* 142(2):807–828.

Libecap, Gary D. and Dean Lueck. 2009. The demarcation of land and the role of cordinating institutions. *Processed, Unoversity of California at Santa Barbara* .

Mayer, C. J. and C. T. Sommerville. 2000. Land use regulation and new construction. *Regional Science and Urban Economics* 30:639–62.

McConnell, V. D. and M. A. Walls. 2005. The value of open space: evidence from studies of nonmarket benefits. *Resources for the Future Report* 30:Washington, DC.

McMillen, D.P. and J.F. McDonald. 1991. A Simultaneous-Equations Model of Zoning and Land Values. *Regional Science and Urban Economics* 21(1):55–72.

Quigley, J. and S. Rafael. 2005. Regulation and the high cost of housing in California. *American Economic Review papers and Proceedings* 95(2):323–28.

Rossi-Hansberg, E., P. Sarte, and R. Owens. 2009. Housing Externalities. *Processed, Princeton University* .

Saiz, A. 2008. On local housing supply elasticity. *Working paper* 0(0):0.

Suzuki, J. 2008. Land use regulation as abarrier to entry: Evidnece from the Texas lodging industry. *Working paper* 0(0):0.

Wallace, N. 1998. The market effects of zoning undeveloped land: Does zoning follow the market? *Journal of Urban Economics* 23:307–26.

Zhou, J., D.P. McMillen, and J.F. McDonald. 2008. Land values and the 1957 comprehensive amendment to the Chicago zoning ordinance. *Urban Studies* 45(8):1647–61.

Appendix: Land use regulation, rent, and the supply of developable land

A number of authors argue that land use regulation drives up the price of housing by restricting the supply of developable land, e.g., Quigley and Rafael (2005) and Glaeser, Gyourko, and Saks (2005). To understand how such a restriction operates in the context of this model, consider regulation which prohibits development in an interval $[a,b]$ of the left municipality and suppose that this regulation has no external effects. We see in equation 2 that land rent is determined by each consumer's outside option. In the absence of external effects, restricting the supply of developable land does not affect consumers' willingness to pay to live in municipality L. Thus, a prohibition on development in $[a,b]$ forces land rent in this interval to zero, but does not otherwise affect land rent. In short, free mobility assures that the demand for land is perfectly elastic, and hence that prices do not respond to changes in supply.

The land rent gradient in municipality L with the prohibition on development in $[a,b]$ parallels the gradient without this prohibition, except on the interval where development is prohibited. This land rent gradient is illustrated in figure 1 by the dashed line. In the region where development is prohibited, no transactions occur, so that this region is invisible to the econometrician. This means that in a world with free mobility, regulation that affects land supply (but which has no external effects) will not affect our estimates of land rent in regions where we observe transactions. Thus, we do not further consider the role of 'supply restrictions' separate from other regulation.

This does not mean that land use regulation does not affect the price of *housing*. Regulation which doubles a binding minimum lot size should be expected to increase the price of housing.

However, the effect on welfare of such a change in housing prices is ambiguous. Regulation which constrains developers to use twice as much land per house may drive up housing prices even as it drives down total land rent. Assessing the welfare implications of such land use regulation requires that we calculate the overall change in unit land rent due to own and external effects.

Land use regulation and rent with heterogenous populations

Suppose that there are two types of agents, A and B. If these agents differ only in that $w^A > w^B$, then we see from equation 2 that the type A agents will always outbid the type B agents, and both municipalities will be entirely populated by agents of type A. Similarly if the types differ only in their outside options. Thus, introducing heterogeneity in wages or outside options does not change the analysis in an interesting way. It does require that we interpret the municipal land rent gradients as reflecting the value of regulation to the agents who sort into the regulated municipalities.

More interesting sorting occurs if the two types have different tastes for regulation. Let v_E^A and v_E^B denote the value that type A and type B agents assign to the external benefits of the more regulated left municipality, and that $v_E^A(z) > v_E^B(z)$ for all $z > 0$. We suppose that types A and B are otherwise alike. From the definition of V_E we have that $V_E^A(x,z^L,z^R) > V_E^B(x,z^L,z^R)$ for all $x < \overline{x}$ and are otherwise equal. Analogous to equation 2, with free mobility the bid rent functions for type A and B agents are

$$\begin{aligned} p^A(x) &= w - \theta + A_1 \ln(V_O(x,z^L,z^R) + A_2 \ln(V_E^A(x,z^L,z^R) \\ p^B(x) &= w - \theta + A_1 \ln(V_O(x,z^L,z^R) + A_2 \ln(V_E^B(x,z^L,z^R). \end{aligned}$$

Since $V_E^A > V_E^B$ it follows immediately that the type A agents outbid type B at every location where the external effect of the more stringent regulation of the left municipality is felt. That is, all locations to the left of $\overline{x}$. For locations where the external effect of the stringent left regulation is not felt, we have $p^A(x) = p^B(x)$, and types A and B are assigned to these locations at random. Our econometric strategy will rely on this result: for all types of immigrant heterogeneity, even type specific preferences for regulation, the region $[-\overline{x},\overline{x}]$ is entirely populated by agents of a single type.

This leads us to the following unsurprising conclusions. First, to the extent that regulation induces sorting of heterogenous agents, this must occur because different types place different values on the costs or benefits of regulation. Second, if sorting on the basis of type occurs in the neighborhood of a municipal border, then cross-border variation in land rents reflects the value placed on regulation by the type that lives near the border. Our econometric estimates of the value of land use regulation should be understood in this light.

If regulation induces sorting by income, and if wealthy communities subsequently demand high levels of local public goods whose values are capitalized into land, then the value of these public goods may also be reflected in land rent gradients across the municipal border. Since the initial sorting is precipitated by land use regulation, we can reasonably attribute the value of the public goods to the land use regulation that precipitated their creation. With this said, it is also be desirable to identify the value of public goods separately. In the discussion below we address this issue in more detail.