

Dascher, Kristof

Conference Paper

City Silhouette, World Climate

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik
- Session: Environmental Aspects of International Trade, No. A16-V3

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Dascher, Kristof (2014) : City Silhouette, World Climate, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Environmental Aspects of International Trade, No. A16-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100538>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

City Silhouette, World Climate

– January 20th, 2014 –

Kristof Dascher¹

Abstract: A country's urban silhouettes prophesy its future climate policy, or so this paper argues. The more its city silhouettes are skewed to the periphery, the more likely a country is to implement the carbon tax. This is why the effect of a country's urban form on greenhouse gas emissions – a bone of contention in the recent literature – cannot be separated from that country's choice of carbon tax. From this paper's perspective, a country with greater city silhouette skews may emit less greenhouse gases not so much because its cities are more compact but because it places a higher price on carbon consumption.

Keywords: Urban Silhouette, Political Economy, Climate Policy, Silhouette Skewness

JEL-Classifications: R12, Q54, H41

Kristof Dascher
Touro College Berlin
Department of Management
Am Rupenhorn 5
14 055 Berlin
Germany
Email: dascher@euv-frankfurt-o.de

¹Stimulating comments on an earlier version, by Rainald Borck, Alexander Haupt, Stephan Heblich, Robert Schwager, Matthias Wrede, as well as by seminar and conference participants in Plymouth, Regensburg, Palermo and Essen, are much appreciated. Any remaining errors are mine.

1 Introduction

Some countries take climate change more serious than others. At least this is what a casual glance at the gasoline tax, itself not a trivial tool in the panoply of climate policy instruments, suggests. Currently the gasoline tax stands at \$4.19 per gallon in the Netherlands, \$3.80 in France, or \$3.54 in Italy. At the same time this tax comes down to \$1.20 in New Zealand, \$0.96 in Canada or only even \$0.49 in the US. These simple figures (all taken from Knittel (2012), earlier yet related figures are found in Pucher (1988)) do provide some motivation for this paper's premise. This premise is that there are stark differences in carbon taxation across countries, and that these differences are not easily explained by differences in income, differences in climate change exposure (documented in Desmet/Rossi-Hansberg (2012)), or differences in carbon-based resources. Even less can carbon tax variation be put down to countries' common incentive to free-ride.

Understanding carbon taxes (or any other climate policy equivalent to it) could benefit, so this paper argues, from studying *city silhouettes*. The intuition underlying this idea unfolds in five small consecutive steps: (i) A carbon tax raises the cost of carbon intensive commutes. (ii) More expensive commutes have urban residents compete more for living in the city center. (iii) Growing city center rents capture landlords' imagination. And so (iv) where tenants will always resent the tax, (v) landlords may actually support it . . . provided that housing is more plentiful at the city center than at the periphery. – In our words, the more a country's city silhouettes are skewed towards the periphery, the more likely that country is to implement the carbon tax. Literally, Europe puts in greater efforts into tackling greenhouse gas emissions than the US not because Europeans are more environmentally aware but because European silhouettes are more strongly skewed.

Our understanding of carbon taxes can do without assuming different beliefs. Hepburn/Stern (2008, p. 260), for instance, have suggested that “significant proportions of citizens in both Britain and America still do not believe that the world is warming owing to human activity”. Our understanding of carbon taxes could also offer a hiatus in the controversy over whether international climate policy negotiations are ridden by strategic behavior. Carraro/Siniscalco (1998), for example, have argued that a stable grand coalition of carbon taxing countries does not exist. In a sense here we deemphasize the role of strategic behavior, much as pointing to climate policy's potential ancillary benefits (Altemeyer-Bartscher/Markandiyā/Rübelke (2011)) deemphasizes it. Suppose that, in the extreme, only city silhouettes mattered. Then national efforts would be as much set in stone as the buildings those silhouettes are composed of, and certainly not susceptible to strategic behavior.

Our urban contour based explanation has support for the carbon tax come from the electorate's important subset of landlord voters (a majority of the electorate in most countries). We show that landlords may support the tax never mind the fact that they, too, must confront those higher travel to work costs induced by the carbon tax. In the context of a Ricardian city we identify an (easily verifiable) condition as to when the representative city's landlord class benefit from the carbon tax. This condition involves that city's “silhouette”, introduced as a relative of both the city's density profile and its skyline as perceived by a not-too distant observer. *The landlord class will benefit from the*

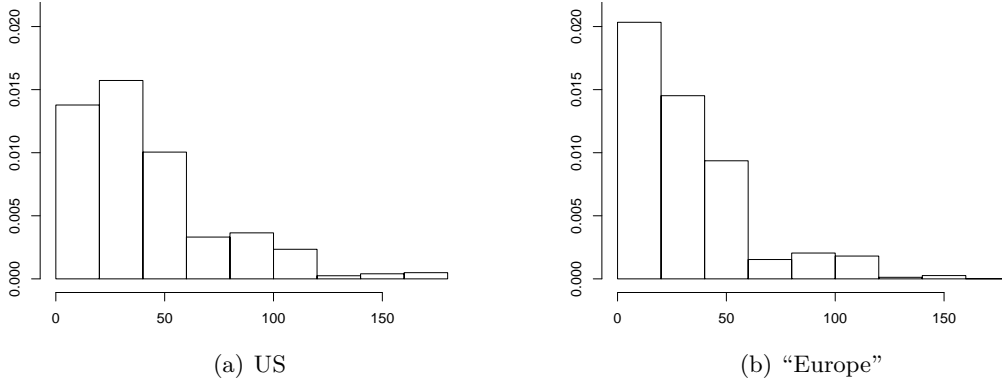


Figure 1: Two Distributions of Daily Commuting Time (in min.)

carbon tax if (and only if) the representative city silhouette is skewed towards the periphery. This statement is independent of how landlords and tenants are assigned to city rings. It effectively pins a country’s climate policy down to the skewness of its representative city silhouette.

All of this promises to uncover a latent connection between a country’s urban form and the political economy of its climate policy. But how general is this result? Most importantly, not all landlords benefit from the carbon tax. While properties close to the city center increase in value properties close to the urban boundary decrease in value, as the plight of commuting between these peripheral plots and the city center intensifies. The carbon tax in fact makes worse off landlords owning properties close to the city’s pre-tax periphery. It might not be sensible to assume that landlords act collectively. However, even if landlords do not act collectively we continue to find that the city silhouette has a powerful role in predicting the strength of the political support for the carbon tax. In essence we observe that: *The city silhouette’s skew also bounds from below the number of landlord beneficiaries.* The greater the representative city’s skew, the more confident can we be of individual landlord voters’ support being strong.

Cities are rarely representative. Cities differ in size, income, etc. If only we recognize the national commuting distribution as the city silhouette’s federal sibling, silhouettes continue to matter to the political economy of climate policy even in a system of heterogeneous cities. We show that: *In a heterogenous urban system, the landlord class’ attitude towards the carbon tax is predicted by both the national commuting time distribution’s skew and the aggregate share of tenants in the overall population.* A country with (i) greater commuting distribution skew and/or (ii) a larger national tenant share sets a higher carbon tax. Despite all that city level heterogeneity, national climate policy may be read off simple national aggregates. The reader permitting, we put this idea to a quick and rough “test”. Figure 1 shows the histograms of commuting time for “Europe” (i.e. the 34 countries participating in the European Survey of Working Conditions 2010) and the US (data from the American Community Survey 2011). Here “Europe”’s commuting distribution surely looks more skewed, while tenant shares are similar.²

²Tenant shares are 0.29 for the EU-29 (Eurostat) and 0.33 for the US (US Census Bureau). Appendix B’s Part (iii) has more detailed information on the commuting data that enter Figure 1.

Brueckner (2005), Glaeser/Kahn (2010), Glaeser (2011), Kim/Brownstone (2013) and de Lara et al. (2013) all have recently suggested that more “compact”, i.e. more densely populated, cities emit less greenhouse gases (GHG). This view also resonates with environmentalists (e.g. Lopate (2004)) and architects or urban planners (e.g. Roaf/Crichton/Nichol (2009)). In contrast, Gaigné et al. (2012) argue that more compact cities may drive up GHG emissions in the transport sector should greater compactness come along with unfavorable adjustments in city sizes, and Borck (2014) argues that less compactness may reduce GHG emissions in the residential sector should less compactness be brought about by tighter building height restrictions. Somewhat surprisingly, all papers party to this important controversy treat the carbon tax as being orthogonal to urban form. Yet from this paper’s perspective, urban form’s effect on GHG emissions cannot be separated from the carbon tax. *An analysis of urban form’s impact on GHG emissions must account for urban form’s simultaneous impact on the politics of the carbon tax.* Countries with more compact cities may emit less GHG not so much because they are more compact but because they place a higher premium on carbon consumption.

If the silhouette skew is important we must ask why city silhouette skews differ across countries. Intuitively, features of the natural terrain play a role here, as must institutions such as the city’s historical zoning record. Anything forcing individuals to reside further away from the city center contributes to reducing the city’s skew. Building height restrictions prevent property development near the CBD. Green building ordinances have a similar effect, by putting up the cost of remaking the city generally. *Building height restrictions, low emission construction standards and similar types of green zoning are at cross-purposes with both (i) cities’ transformation into greener form and (ii) voters’ support for the carbon tax.* This observation’s first part conforms with earlier work on zoning, emissions and climate change (e.g. Bertaud (2004), Glaeser (2011)) and thus is not, unlike the other italicized results in this introduction, new.

Our focus on the existing international variation of current climate policies should not distract from the fact that none of these policies stand up to the optimal policy response (IPCC (2013), Nordhaus (2013)). Yet even if our focus is not normative, our analysis nonetheless may provide a small step towards understanding the institutions that give rise to policies that align better with the optimal policy. The emerging pattern of urbanization in the world’s two most populous countries, India and China, has frequently been emphasized in this context. These two countries’ silhouettes are now in their formative years. For many years to come their ultimate form will not just determine commuting distances there (Glaeser (2011)), but will also shape, so this paper argues, carbon tax choices there. From this paper’s perspective, both silhouette skew and tenant share – and their underlying determinants – deserve climate policy analysts’ attention.

There is a large body of literature relating the cost of commuting to urban form (e.g. Glaeser/Kahn (2004), Brueckner (2005), Bento/Franco/Kaffine (2006)). This literature’s interest is in the important effect of the cost of commuting on urban form. I.e., cheaper petrol facilitates the decentralization of population. At the same time this literature generally shuts out the possibility of cities’ urban form looping back into the carbon tax. Ultimately causality runs both ways between urban form and carbon tax. In this paper the carbon tax affects urban form not just because urban residents frequently flock to

more central parts of their cities in response (as in, say, de Lara et al. (2013)) – but also because large cities contract. So even with an endogenous distribution of city sizes, in our setup it is true that: *Not only do individual cities compactify in response to the carbon tax; the urban system compactifies, too.*

We briefly address three further strands of the literature this paper connects to also. Evidently this paper fits into the vast literature on the redistributive side effects of public good provision. Typically there some decisive subset of society has government provide a public good even as this makes that subset’s complement worse off. Then the paper also accords well with the literature on commuting subsidies. Much as a climate tax raises urban travel costs do commuting subsidies reduce it (Brueckner (2005), Borck/Wrede (2005)). At the same time this literature does not embed its discussion of tax or subsidy into a context of negative externalities. In a context of GHG emissions, discussing a tax is not simply the reverse of discussing a subsidy. Finally this paper is also preceded by the aggregate land rent literature following Arnott/Stiglitz (1982). That literature’s focus is on how aggregate rents and commuting costs relate to each other. Our focus instead is on how rental incomes and commuting costs shape landlord incentives.

The paper comes in eight sections. Section 2 offers a three paragraph starter. Section 3 introduces a representative city’s silhouette in a standard closed-city framework. It is also there that we make more precise the concepts of “compactness”, “contour”, “silhouette” and “density” strewn across this introduction. Section 4 shifts attention away from the landlord class, and towards individual landlords. Section 5 replaces the representative city framework with an urban system awash with heterogeneous cities. Section 6 allows for urban system-wide housing stock adjustment, and for environmental benefits to win over tenants, too. Section 7 considers further extensions, while section 8 concludes.

2 A Night Time Silhouette

We illustrate the paper’s theme by way of a simple linear-city example. Consider three “rings” around the central business district (CBD), and at ever greater distances to it. Let introducing a carbon tax raise the cost of commuting from ring 1 to the CBD by 0 Euro, from ring 2 to the CBD by 2 Euro, and from ring 3 to the CBD by 4 Euro. Once adjustments have taken place, rent in ring 1 must have risen by 4 Euro, and rent in ring 2 must have risen by 2 Euro, while rent in ring 3 will not have changed at all. Following Ricardo, these changes just offset the extra commuting cost advantages rings 1 and 2 enjoy vis-à-vis ring 3. Now consider six units of (equally sized) housing. Three of these units are to be found in ring 1, two in ring 2, and one unit of housing is peripheral, in ring 3. Finally, let one half of society, also referred to as its three landlords, own all six units of housing.

We quickly assess the costs/benefits attached to different allocations of homeowners and tenants. One such allocation is $(\{1, 3\}, \{1, 2\}, \{1, 2\})$, where the interpretation of, say, $\{1, 3\}$, is that a landlord residing in ring 1 herself has her tenant live in ring 3. In this allocation, the landlord wound up in the first match $\{1, 3\}$ clearly loses nothing in commuting costs but also gains nothing in rent. Landlords party to either the second or third match

$\{1, 2\}$ do not suffer from extra commuting costs yet gain 2 in rent. Adding up yields a net aggregate landlord gain of 4. This gain, we note first, never depends on how landlords and tenants are distributed across the housing stock. For example, a different landlord-tenant allocation, of $(\{2, 3\}, \{1, 2\}, \{1, 1\})$, yields an identical net aggregate landlord gain of 4. Moreover, as we emphasize second, this gain may even be read off the city’s physical form:

Suppose the three units of housing in ring 1 were stacked on top of each other, composing a building of three stories. Likewise, the two units of housing in ring 2 could form a two storey house while the single unit of housing in ring 3 is the city’s “bungalow”. Then from a distance an observer would not just make out the urban silhouette lit up against the night time sky but would also immediately recognize this silhouette to be skewed towards the periphery. Intuitively it is this skew to the periphery that underlies the landlord class gain’s being positive. Were this silhouette skewed towards the center then landlords’ aggregate net benefit would be negative. To see this one simply replays our little example with the roles of CBD and periphery reversed.

3 Silhouette Skewness and the Carbon Tax

We introduce our basic model of silhouette skewness. At its core we position a circular monocentric city. This city extends from the CBD out to its boundary \tilde{r} . Think of the city as being split into n rings spaced equally far apart from each other. If distance from the CBD is r , the first of these rings extends from the CBD to \tilde{r}/n , the second from \tilde{r}/n to $2\tilde{r}/n$, and so forth. The number of housing units supplied by ring i is s_i . A fraction θ of the overall urban housing s supplied, $s = \sum_{i=1}^n s_i$, are tenant-occupied; the remaining fraction $1-\theta$ are inhabited by these tenants’ landlords. The number s is even. All residents commute to the city center, where they earn the wage ω . For a resident in ring i , round trip commuting costs are tr_i , with r_i the distance from the CBD to the midpoint between ring i ’s outer and inner annulus. Every resident consumes one unit of housing. There is no agricultural hinterland.

Even more specifically, for now we also assume: (i) landlords are resident, not absentee, (ii) the city is representative of every of the urban system’s (many) cities, (iii) the wage is given, (iv) within-ring-travel is costless, and (v) all housing is inherited from the past and fixed. In fact, for now we even assume that (vi) the landlord class pursue its aggregate interest, (vii) the landlord class are decisive, (viii) no one cares about the climate, and (ix) tax revenues are not refunded. Revenues are spent on national public goods that enter household utility in additive fashion, and are suppressed notation wise. The first eight of our nine assumptions we will relax gradually, in the order in which they appear here. We never relax assumption (ix). One might argue that environmental tax revenues are always unlikely to be refunded to the tax payer (directly). Or one might simply consider assumption (ix) to be the model’s hinge.

Our representative city is closed (e.g. Moring (1961), Wheaton (1974), Brueckner (1987)). Any shock rippling through our cities below will occur in every one of them alike, simultaneously. Tenants’ competition for the best location within the city implies that income remaining once commuting cost and rent $q(r)$ are deducted must always be the same,

irrespective of tenants' location. Comparing any intra-urban plot with the last peripheral plot occupied thus yields the fundamental $q(r) + tr = t\tilde{r} + q(\tilde{r})$. Throwing in $q(\tilde{r}) = 0$ (peripheral residents do not need to compete, given the abundance of land just one step beyond the urban fringe) joint with the assumption that all residents are perfectly mobile makes urban rent follow the Ricardian $q(r) = t(\tilde{r} - r)$. Rent in r reflects nothing but commuting cost savings from living in r rather than out in \tilde{r} . Tenants face an urban cost-of-living equal to $q(r) + tr$, or $t\tilde{r}$. This urban cost-of-living is the same at every city location.

It is instructive to start out with every landlord owning two properties: one property to live in, and another one to let. (I.e., so $\theta = 0.5$ for now.) Consider a landlord who resides in ring i yet rents out her or his extra property in ring j . This is a “match” $\{i, j\}$. For the landlord involved in such a match, utility is $a_{ij} = \omega - tr_i + q_j$. Dropping the fixed wage for convenience, the full $n \times n$ matrix of landlord utilities connected to residing in i and renting out in j is a “valuation matrix”, denoted A ,

$$A = t \begin{pmatrix} -r_1 + (\tilde{r} - r_1) & \dots & -r_1 + (\tilde{r} - r_n) \\ & \vdots & \vdots \\ -r_n + (\tilde{r} - r_1) & \dots & -r_n + (\tilde{r} - r_n) \end{pmatrix}, \quad (1)$$

and featuring symmetry, given that $a_{ij} = a_{ji}$ for all i and j . Now consider some arbitrary assignment of landlords and tenants to city rings, i.e. an intra city spatial allocation. Since $q(r) = t(\tilde{r} - r)$ in spatial equilibrium, tenants never have an incentive to relocate. We now add that the same is true for resident landlords. Neither will a landlord want to rent out her or his own dwelling to become tenant elsewhere.³ Nor will a landlord want to exchange his location with her or his tenant, in view of A 's symmetry. Any allocation conforming with spatial equilibrium and the pre-existing distribution of housing units (s_1, \dots, s_n) is a locational equilibrium.

A 's counter diagonal (comprising all the elements on the diagonal stretching from the bottom left corner to the top right hand corner) consists of zeros only, because $r_i + r_{n+1-i} = \tilde{r}$.⁴ Matches for which row index i and column index j sum to $n + 1$ represent those perfect hedges for which the landlord's rental income is always just offset by her or his travel cost. In contrast, entries above (below) the counterdiagonal of A are always strictly positive (negative). Now, to A corresponds quite naturally a second matrix B of identical dimensions collecting the frequencies with which the various matches occur. In this “match matrix” the entry b_{ij} simply represents the number of times the match $\{i, j\}$ applies. The aggregate surplus accruing to the landlord class w_l may then be computed as

$$w_l = \iota' (B \circ A) \iota, \quad (2)$$

where \circ is the entry wise (or Hadamard) product while ι is a commensurate (i.e. $n \times 1$) vector of ones.

In applications we are unlikely to be informed about the precise structure of landlord-tenant matches. Fortunately, these – unobservable – matches are intimately related to the

³A landlord moving out of his owner-occupied dwelling in ring i to become tenant in j gains $tr_i + q(r_i)$ in income yet also expends an extra, and equal sized, $tr(q_j) + q_j$.

⁴This can easily be checked after noting that $r_i = [(i/2) + (i - 1)/2]\tilde{r}/n$.

– observable – structure of housing units they are housed by. Let l_i and m_i denote landlords and tenants in ring i , respectively. Then $-t \sum_{i=1}^n l_i r_i$ captures (the negative of) landlords’ aggregate costs of commuting. At the same time, $t \sum_{i=1}^n m_i (\tilde{r} - r_i)$ captures landlords’ aggregate rental income. Intuitively, these very two aggregates add up to landlord class welfare, so w_l simply becomes $-t \sum_{i=1}^n s_i r_i + t s \tilde{r} / 2$ or, alternatively, the first expression in (3). Proposition 1 restates this expression in its various “apparitions”, also relating it to the two well-known concepts of aggregate land rent $ALR = t \sum_{i=1}^n s_i (\tilde{r} - r_i)$ and aggregate commuting costs $ATC = t \sum_{i=1}^n s_i r_i$. The proposition’s formal proof, delegated to the Appendix A as most of the paper’s proofs, departs from the definition of landlord class welfare in (2).

To appreciate the striking simplicity of the expressions given in (3) note that the number of potential landlord-tenant matches that could possibly be housed by the existing distribution of housing units (s_1, \dots, s_n) is bound to be very large. Yet even so w_l is entirely independent of how landlords and tenants are allocated to this given distribution, and the same is true for tenant welfare w_m (Proposition 1, Part (i)). Effectively none of the expressions in equations (3) feature anything but consolidated ring aggregates. Intuitively, replacing a landlord with some (and not necessarily her or his) tenant has no effect on w_l . Where before replacement it was the landlord’s commuting costs $-tr_i$ that pulled down w_l , after replacement it is the tenant’s commuting costs $-tr_i$ that pull down w_l (by the damage they do to the rent that could be extracted otherwise). Proposition 1’s Part (i) generalizes the spatial invariance theme introduced in the previous section’s little example to any finite number of city rings n and dwellings s .⁵

Proposition 1 (Political Economy and Urban Form)

(i) (*Spatial Invariance*): Both landlord and tenant welfare are invariant w.r.t. how landlords and tenants are allocated to the existing ring specific housing supplies, (s_1, \dots, s_n) .

(ii) (*Tenant Welfare*): Tenant class welfare w_m is independent of urban form, and equals either $-s\tilde{r}/2$ or $-(ALR + ATC)/2$.

(iii) (*Landlord Welfare*): Landlord class welfare w_l is dependent on urban form, and equals any of the following three expressions:

$$t \sum_{i=1}^n \left((\tilde{r}/2) - r_i \right) s_i = (ALR - ATC) / 2 = t \sum_{i=1}^{n/2} \left((\tilde{r}/2) - r_i \right) (s_i - s_{n+1-i}). \quad (3)$$

The practical importance of Proposition 1’s Part (i) is to free us of having to pay attention to resident landlord and tenant location in any of the following. Proposition 1’s Part (ii) proceeds to the issue of aggregate tenant welfare. Each tenant simply incurs those

⁵We briefly pursue an instructive alternative path leading up to the third expression in (3). Even if at first appearance a very special case, let us investigate a city in which the s_i are decreasing in i . (Such a city is illustrated further down, in Figure 2’s panel (a), for the case of $n = 6$.) We assign landlords and tenants to rings 1 through n by making use of the two following rules: (i) First, the s_n landlords to ring n all live in ring 1, the s_{n-1} landlords to ring $n - 1$ all reside in ring 2, etc. And (ii), housing in ring i not occupied yet by the demands of rule (i) is equally shared between remaining tenants and their respective landlords. This special case makes for a particularly simple description of landlord welfare. First, none of the landlords described by (i) receives any match benefit because for these landlords’ matches indices i and j sum to $n + 1$. And second, all of those $(s_i - s_{n+1-i})/2$ landlords in rings $i = 1, \dots, n/2$ addressed by rule (ii) (rather than by rule (i)) receive a utility of $t\tilde{r} - 2tr_i$ each. Aggregating these landlord utilities across the first $n/2$ rings yields the last expression in (3).

familiar costs-of-living of $t\tilde{r}$. Next, according to the second expression in (3), landlord class welfare w_l and tenant class welfare w_m may also be expressed in terms of ALR and ATC . Specifically, landlord welfare w_l may also be written as the difference $(ALR - ATC)/2$ even as of course ALR includes (imputed) rent payments never received, just as ATC also includes commuting costs never incurred, by the landlord class (Part (iii)). ALR and ATC conform with standard urban welfare accounting. For instance, $ALR + ATC = St\tilde{r}$, as in Mohring (1961). More importantly, Part (ii) of Proposition (1) allows us to trace out how this paper differs from Arnott/Stiglitz (1981): Arnott/Stiglitz (1981)'s interest is in whether ALR monitors ATC , whereas our interest is in how the difference between ALR and ATC traces out landlord interests.

As a first step towards a city's "silhouette" we compute ring i 's average housing density, d_i , by dividing the stock of ring i 's housing s_i by that same ring's land area a_i (Part (i) of Definitions 1 below). The density profile $d(r)$ then is the set of all ordered pairs of commuting distances and average densities (Part (ii)). In contrast, the city silhouette $s(r)$ is the set of ordered pairs of commuting distances and ring housing stocks (Part (iii)). More prosaically, in the monocentric city the city silhouette coincides with the local distribution of commuting lengths. At first sight it is the density profile that appears to capture best the city's "true silhouette" as witnessed from a distance. Yet note that this "true silhouette" in fact is architects' perspective projection of the upper envelope of the three-dimensional city into two-dimensional space. This projection is not generally the same as the density profile (Part (ii)).⁶

Definitions 1 (City Density, Profile, Silhouette, Skew, Unbalancedness)

- (i) *Ring Density* ... is total housing in ring i divided by i 's area, $s_i/a_i = d_i$.
- (ii) *Density Profile* ... maps distance into density, $\{(r_1, d_1), \dots, (r_n, d_n)\} = d(r)$.
- (iii) *City Silhouette* ... maps distance into ring housing, $\{(r_1, s_1), \dots, (r_n, s_n)\} = s(r)$.
- (iv) *City Skewness* ... is $\sum_{i=1}^{n/2} ((\tilde{r}/2) - r_i)(s_i - s_{n+1-i})/s = \sigma$.
- (v) *City Unbalancedness* ... is $\sum_{i=1}^{n/2} (s_i - s_{n+1-i})$.

We take the liberty to define the city's silhouette in a way that suits our interest in urban form best, i.e. as in Part (iii). Now, typically density decreases as we move out towards the city's periphery because rent, and hence the incentive to build high, diminish – as postulated in theory (e.g., Fujita (1989)) and as observed for many real cities (e.g. Bunting/Filion/Priston (2002)) on Toronto, Montreal, or Ottawa-Hull, or Bertaud (2004, Figure 4) on Barcelona, Warsaw or Bangkok. In contrast, the silhouette, being the product of ring density with ring area, may display much richer behavior. While this product might well decrease (e.g. de Lara et al. (2012, fig. 2) on Paris), it need not decrease at all, and may in fact increase, as we move out, to the extent that built up land rises faster than density falls. For instance, US cities such as Atlanta and L.A. exhibit very flat density gradients, and density profiles for Moscow, Johannesburg and Brasilia are even sloping upwards, and strongly so (Bertaud (2004, Figure 5)). For none of these latter cities do we expect s_i to decrease in i . These cities may be more likely to display a pattern akin to that in panels (d) or (g) in Figure 2 (discussed shortly). Building height

⁶Density profile and perspective projection do coincide if the observer is very far away from the city. Density profile and silhouette do coincide in the simple case of a linear city of unit width (section 2).

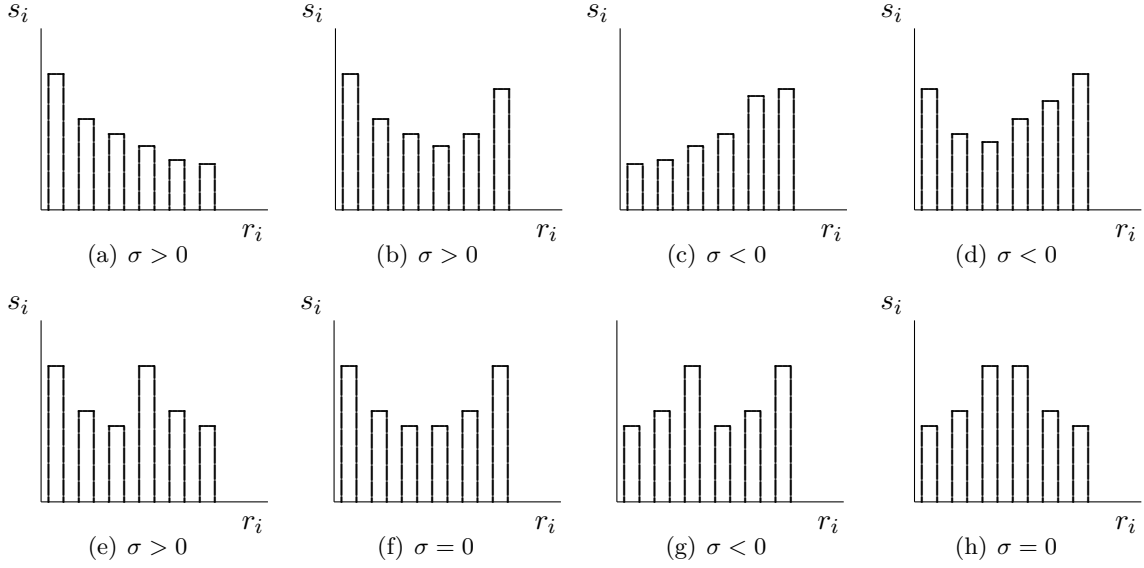


Figure 2: City Silhouette and Silhouette Skewness

restrictions may have interfered with developers’ objectives (L.A.), central land may have disproportionately been set aside for traffic, immigration into the city center could have been prohibited (Johannesburg), or socialist institutions may have eliminated developer incentives altogether (Moscow).

Part (iv) of Definitions 1 introduces the silhouette skewness σ that is at the heart of this paper. Except for the absence of s and t , this skewness coincides with landlord welfare in (3), or $w_l = st\sigma$. To see why σ is a meaningful measure of the silhouette’s skewness we first refer to $\tilde{r}/2$ as “midtown”, and to $s_i - s_{n+1-i}$ as “ring difference i ”. In that sense σ sums over weighted ring differences, where the deviations of commuting distances r from midtown commuting distance $\tilde{r}/2$ are the positive weights. If $\sigma > 0$ at least one ring difference must be positive. In fact, it must be increasingly so as more and more of those other ring differences turn negative. From the distant observer’s perspective even a single positive ring difference suggests an overall urban skew towards the periphery.

Figure 2 explores σ further. Panels (a) and (b) have ring differences all positive; while all ring differences in panels (c) and (d) are negative. Whenever ring housing is monotonically increasing (decreasing) in r_i then σ is unambiguously positive (negative). Ring differences in panel (e) or (g) no longer carry a uniform sign; one of the ring differences is positive, one zero, and one negative. Since early differences receive greater weight than later ones, panel (e) shows a positive skew, while panel (g) exhibits a negative one. Panels (f) and (h) illustrate two non-skewed, or symmetric, silhouettes. Finally, let “city unbalancedness” refer to the excess of residents in the city’s interior (r less than $\tilde{r}/2$) over residents in the city’s periphery (r beyond $\tilde{r}/2$) (Part (v)). Then panels (e) through (h) illustrate balanced silhouettes. Panel (g) also shows why growing unbalancedness need not reinforce skew. Migration of residents initially in the periphery towards the city’s interior may reduce the city skewness if marginal migrants come from, as well as move to, locations close to midtown (i.e. from the “fourth to the third bar” in the panel).

Having laid out the basic model, let federal government now introduce a carbon tax, equal

to $\Delta t > 0$. Now city costs-of-living $t\tilde{r}$ rise throughout the city, by the extent to which commuting costs at the urban boundary \tilde{r} do, i.e. by $\tilde{r}\Delta t$ just. This is each tenant's loss in utility, irrespective of her or his location in space (Proposition 2, Part (i)). Differentiating the last expression in equation (3) with respect to t also shows that per-dwelling landlord class' welfare change, $(1/s)(dw_l/dt)$, just equals σ . The landlord class welcome the tax if (and only if) silhouette skew is positive (Part (ii)).⁷

This far we have assumed that within-ring commuting is costless. It seems more adequate, and it also turns out more convenient, to reduce ring width Δr further. Let a twice differentiable housing shape function $F(r)$, with $F(\tilde{r}) = s$, summarize all available housing between the CBD and r units of distance out. We approximate the number of dwellings in ring i , s_i , by $F'(r_i) \equiv f(r_i)$, in the sense that $f(r_i)$ indicates available housing in the one-unit-width ring r_i away from the CBD. Put differently, $f(r)$ now captures the city's silhouette.⁸ Proposition 2's Part (iii) makes immediate use of this refined silhouette, stating that the change in the landlord class welfare, dw_l/dt , may more compactly be expressed as the integral on the r.h.s. of the first equation in (4).

Proposition 2 (City Silhouette Skew and Political Economy of Carbon Tax)

- (i) Tenant class welfare w_m is decreasing in the tax, independently of city skew.
- (ii) Landlord class welfare w_l is increasing in the tax if (and only if) the city is skewed.
- (iii) More specifically, if city ring width becomes arbitrarily small then the change in landlord class welfare in response to a one Euro tax may be approximated by

$$dw_l/dt = \int_0^{\tilde{r}} (\tilde{r}/2 - r) f(r) dr = s\sigma = s(\tilde{r}/2 - \rho), \quad (4)$$

where ρ is average commuting cost, $\rho = \int_0^{\tilde{r}} (f(r)/s)r dr = ATC/st$.

In the literature “compactness” often is equated with “high density” (e.g. Riou et al. (2012), Glaeser (2011)). Yet density is a function of CBD distance even in the simplest of cities. If intuitively more “compactness” is meant to capture the idea of less aggregate commuting $s\rho$ then equating “compactness” with the silhouette's skew may be a meaningful alternative. On the one hand, from the last equation in (4) $\rho = \tilde{r}/2 - \sigma$. For given “city width” \tilde{r} , average emission abatement ρ tracks (the negative of) skewness $-\sigma$, and in a one-to-one fashion even: $d\rho = -d\sigma$. It is in this sense that changes in skewness $d\sigma$ also are an indicator of changes in the city's GHG externalities $d\rho$ (again, as long as city width remains the same). To phrase this slightly differently, from the last equation in (4) we also conclude that

$$s\sigma/\tilde{r} = s/2 - s\rho/\tilde{r}. \quad (5)$$

⁷From our expression for landlord welfare, obviously, we may even suspect that the landlord class would wish to introduce a subsidy on carbon consumption if the representative city's skew were negative. We do not discuss the subsidy any further. Ultimately with endogenous city size (as in section 6) a subsidy creates extra sprawl, and hence in our setup produces no aggregate welfare gain.

⁸If we decompose the city's silhouette $f(r)$ into the product of a differentiable land function $a(r)$ with a differentiable housing density function $d(r)$, $f(r) = a(r)d(r)$, we may conveniently revisit our earlier discussion on the relationship between silhouette and density profile. A silhouette decreasing in r amounts to observing $f' < 0$, or $a'(r)/a(r) < -d'(r)/d(r)$. Thus the silhouette $f(r)$ is not decreasing in r just because density $d(r)$ is. Rather, the silhouette is decreasing if (and only if) available land is not growing faster than density is shrinking.

This alternative equation relates emissions (standardized by city width) to skewness (also standardized by city width). Standardized skewness reveals standardized emissions. On the other hand, skewness also predicts landlords' interest in overcoming these emissions (Proposition 2, Part (i)). Thus a large (standardized) skew really captures both: (i) small global externalities joint with (ii) strong local interest in the carbon tax.

4 Skewness and Landlord Beneficiaries, . . . and Zoning

This section provides a different, yet complementary motivation for studying silhouette skewness. Suppose landlords fail to unite as an interest group. For example, in a city with 6 rings a landlord residing, and also renting out to a tenant, in ring 5 will clearly lose more in extra commuting costs than she or he gains in extra rent. In contrast, a landlord resident in ring 5 and renting out in ring 1 enjoys a net gain. If there are no transfers from landlords who gain to landlords who lose, landlords should not be expected to form an interest group. Support for the carbon tax would come from less than one half of the electorate. But from how much less? While we are not able to compute landlord beneficiaries' exact number we nonetheless may place a lower bound on it, by inspecting what are successive cumulative ring differences.

To illustrate the underlying principle we start with housing units in ring 1. Except for those residents matched up with residents in ring n , all of these units are tied up in matches with strictly positive value. In the extreme, every resident in ring n might be linked to some resident in ring 1 (rather than to some resident in any of the remaining rings). In the extreme, moreover, all $(s_1 - s_n)$ remaining residents in ring 1 might be matched up with one another (rather than to residents in any of the remaining rings). Then $(s_1 - s_n)/2$ supplies a lower bound on landlords who are better off strictly. Of course, this lower bound may be negative, in which case it is not particularly informative. But there are many other lower bounds. For instance, $((s_1 + s_2) - (s_{n-1} + s_n))/2$ is another lower bound, as in fact is any partial sum $l(n') = \sum_{i=1}^{n'} (s_i - s_{n+1-i})/2$, with $n' \leq n/2$. Let us pick n' such that l becomes greatest. This greatest lower bound involves the first as well as the last n^* rings in the succession of concentric rings around the CBD, and hence $l^* = \sum_{i=1}^{n^*} (s_i - s_{n+1-i})/2$ provides the minimum number we are looking for (Proposition 3, Part (i)). Reverting to Figure 2 helps illustrate these ideas. In panel (a), $n^* = 3$ and hence $l^* = \sum_{i=1}^3 (s_i - s_{7-i})$. In panel (e), $n^* = 1$ and hence now $l^* = s_1 - s_6$ only. We emphasize that l^* is computed simply by inspecting the representative city's silhouette.⁹

Alternatively, let the city be divided into arbitrarily many rings of correspondingly smaller width. Then $l(r') = [F(r') - (s - F(\tilde{r} - r'))]/2$ gives the lower bound on landlord beneficiaries if both the first r' and last r' rings are included. Maximizing this expression with respect to r' implicitly defines the optimal ring index r^* . The corresponding value function value, $l(r^*)$, is $l^* = [F(r^*) - (s - F(\tilde{r} - r^*))]/2$, and identifies the greatest of all of these lower bounds (Part (ii)). The Proposition's third part relates the city silhouette's skew not just to aggregate landlord welfare, but also to landlord beneficiaries' absolute

⁹By analogy we may extend this idea to also assessing the greatest lower bound on the number of landlords who are strictly worse off with the carbon tax and must be expected to oppose it, l^{**} . In panel 1(g), for instance, we would be certain that this number comes to $l^{**} = (s_6 - s_1)/2$.

number. This part states that, provided that the mean ring difference is positive (satisfied in Figure 2's panels (a), (b), (f) and (h), for instance), the adjusted skew (i.e. skew divided by city width) bounds the number of landlord beneficiaries from below.

Proposition 3 (Silhouette Skewness and Landlord Voting)

- (i) (Greatest Lower Bound) ... on landlord beneficiaries l^* is $l^* = \sum_{i=1}^{n^*} (s_i - s_{n+1-i})/2$, where n^* is the very n' that maximizes $\sum_{i=1}^{n'} (s_i - s_{n+1-i})$. Moreover,
- (ii) (Greatest Lower Bound) ... approaches $l^* = [F(r^*) - (s - F(\tilde{r} - r^*))]/2$ for ring width Δr sufficiently small. If r^* equals neither 0 nor $\tilde{r}/2$ then it must satisfy

$$f(r^*) = f(\tilde{r} - r^*). \tag{6}$$

- (iii) (Landlord Beneficiaries): If the average of distance weighted ring differences is non-negative then it is true that

$$\sigma s / \tilde{r} \leq l^*. \tag{7}$$

Proposition 3's Part (iii) is its most central part. Part (iii)'s equation (7) points to σ 's informational content even in societies in which landlords do not act collectively. The greater the city silhouette's skew, the more confident can we be of landlord beneficiaries' contribution to the overall support for the carbon tax. Phrased yet differently, while Proposition 2 illustrates how greater skew increases an existing landlord majority's desire for the carbon tax, Proposition 3 illustrates how greater skew strengthens landlord beneficiaries' political clout.

So far the housing stock's size and spatial distribution have been inherited from the past. Real cities are shaped not just by the forces of competition between profit maximizing developers, but also by their individual mix of landscapes that have hosted them and of past zoning that has shaped them. In this context Proposition 4 briefly turns to building height restrictions, such as floor-area-ratios (FAR). FAR are biased against the city's center because that is where buildings typically want to be taller. Suppose our housing shape function F is indexed by δ such that a greater δ raises F at any r , i.e. $F_\delta(r, \delta) \geq 0$, such that a greater δ captures the effect of lifting the height constraint marginally.¹⁰

Proposition 4 (Building Height Restrictions and the Carbon Tax)

- (i) A history of tighter regulation of building heights reduces both the city's skew, σ , and the greatest lower bound on landlord beneficiaries' number, l^* , and hence makes us less confident of landlord support for the carbon tax.
- (ii) A history of tighter regulation of building heights also raises the greatest lower bound on opponent landlords' number, l^{**} , and hence has us more confident of landlord opposition against the carbon tax.

Tighter zoning in the past, or a smaller δ , translates into a smaller skew today (Proposition 4, Part (i)). When combined with Proposition 3 this implies that countries with a history of tighter building height restrictions are less likely to introduce the carbon tax. Finally, building height restrictions also raise the lower bound on the number of landlords in opposition to the carbon tax, denoted l^{**} (Part (ii)).

¹⁰We pursue a somewhat "macroeconomic" approach, neglecting the effect of changes in δ on the city boundary here. Brueckner/Bertaud (2005), Borck (2014) and section 6 below describe how lifting the height constraint also affects \tilde{r} .

5 Federal Silhouette and Federal Skew

We return to our main line of investigation, following up on the interests of the landlord class. While initially we departed from a representative city with a given wage any urban system displays large variation in city sizes and wages (e.g. Nitsch (2004), Giesen/Südekum/ Zimmermann (2010)). By adding what probably is the simplest possible layer of wage determination we now allow wages and hence city sizes to differ across space. A word on notation: Capital letters refer to the corresponding system quantities. Now, let index j refer to cities $1, \dots, J$, so that any city-specific variables can be indexed by it. Assume any city's CBD to be that city's central commuting node. That is, workers commute to the CBD in the morning, are picked up there and ferried out to a factory at the urban fringe, only to be brought back to the CBD in the evening.¹¹

Thus, factories do not pay rents but instead incur commuting costs $t\tilde{r}$ when transporting its single worker to and fro its site. With fixed factor proportions, city j firms' unit cost function is $\omega_j + q_{j\circ} - g_j$, where ω_j is the local wage, $q_{j\circ} = t\tilde{r}_j$ refers to the cost of worker transport, and g_j is some city-wide productive amenity. Homogeneous output is tradable across cities at no cost and sold at price p , and individuals are perfectly mobile across cities. Essentially this is a Rosen/Roback-type (1982) extension if everywhere (i) perfectly competitive factories make zero profit and (ii) tenant utilities $\omega_j - t\tilde{r}_j = u$ are the same. We marry intracity equilibrium with inter-city equilibrium, as summarized by the $2J$ equations

$$\omega_j = u + t\tilde{r}_j \quad \text{and} \quad p = \omega_j + q_{j\circ} - g_j, \quad (8)$$

reflecting household indifference and firm indifference, respectively. Jointly these equations define local wages ω_j , costs-of-living or center rents $q_{j\circ} = t\tilde{r}_j$ and tenant utility $u = W_M/M$. Cities can be distinguished by their unique endowments of the public good g_j . (Any additions to this public good made possible by the carbon tax's revenue are identical in every city, and thus may be neglected below.)

Cities with greater amenities g_j are larger because in equilibrium such cities permit firms to pay higher wages. Hence these cities must confront their residents with higher costs-of-living $q_{j\circ}$. (This could also be visualized by making use of Rosen-Roback's famous two-loci diagram.) We rearrange city indices such that 1 denotes the smallest, and J the largest, city, and let \bar{n} denote the number of rings in city J , i.e. $\bar{n} = n_J$. We briefly turn back to our initial setup with discrete ring width. To adapt our notation to the necessities of addressing an entire urban system, l_{ji} , m_{ji} , and s_{ji} denote landlords, tenants, and residential properties in city j 's ring i , respectively, where $s_{ji} = l_{ji} + m_{ji}$. For any city $j \neq J$, landlords and tenants in those empty rings $n_j + 1, \dots, \bar{n}$, are zero by definition. Further, $\sum_{i=1}^{\bar{n}} s_{ji} = S_j$, or S_j , is city j 's population (irrespective of ring number), $\sum_{j=1}^J s_{ji} = S_i$, or S_i , is ring i 's population (irrespective of city membership), $\sum_{i=1}^{\bar{n}} \sum_{j=1}^J m_{ji} = M$ is the overall number of tenants, and $\sum_{i=1}^{\bar{n}} S_i = \sum_{j=1}^J S_j = S$ the federation's fixed population total.

¹¹Alternatively we might think of all production collapsed into a single point, at the CBD, with firms paying the CBD's competitive rent.

Federal landlord class welfare is the sum of landlords' total incomes minus landlords' aggregate travel costs, or $W_L = -t \sum_{j=1}^J \sum_{i=1}^{\bar{n}} l_{ji} r_i + t \sum_{j=1}^J \sum_{i=1}^{\bar{n}} m_{ji} (\tilde{r}_j - r_i) + \sum_{j=1}^J L_j \omega_j$. In earlier welfare analysis (in (3)) we could neglect wages. Yet here wages must enter welfare comparisons, being endogenous now. Likewise, earlier we had assumed landlords to be resident, not absentee. Now we recognize this distinction as irrelevant. Given that unit distance commuting cost t are uniform, and only ever change uniformly across cities, landlords' proximity to their tenants is irrelevant. Nor does wage variation create any incentive to relocate. E.g., a landlord in city 7's ring 3 considering to trade houses with some tenant of his in city 2's ring 3 would raise his wage by $\omega_2 - \omega_7$, if depress his rental earnings by $t(\tilde{r}_7 - \tilde{r}_2)$. By the first set of equations in (8), nothing is to be gained from this. Now, Definitions 2 (below) collect the natural federal analogues of Definitions 1' city concepts, such as federal density (Part (i)), federal density profile (Part (ii)), federal silhouette (Part (iii)), and federal skewness (Part (iv)). We note that the federal silhouette coincides with the national commuting distribution, and that federal skewness reduces to representative city skewness if (and only if) $\theta = 0.5$. Finally, Definitions 2 also introduce the average commuting distance $\bar{\rho}$. This is one possible indicator of GHG emissions (Part (v)).¹²

Definitions 2 (Federal Density, Profile, Silhouette, Skewness, Emissions)

- (i) *Federal Density ... is all cities' ring i housing divided by ring i areas, $S_i/A_i = D_i$.*
- (ii) *Federal Profile ... maps distance into density, $\{(r_1, D_1), \dots, (r_n, D_n)\} = D(r)$.*
- (iii) *Fed. Silhouette ... maps distance into ring housing, $\{(r_1, S_{.1}), \dots, (r_n, S_{.n})\} = S_i(r)$.*
- (iv) *Federal Skewness ... is $\sum_{i=1}^{\bar{n}} S_i (\theta \bar{r} - r_i)/S = \bar{\sigma}$.*
- (v) *Federal Emissions per capita ... are $(\sum_{i=1}^{\bar{n}} S_i r_i)/S = \bar{\rho}$.*

Employing $l_{ji} + m_{ji} = s_{ji}$, exploiting that $\omega_j - t\tilde{r}_j$ is constant across cities, and collecting terms, we can simplify W_L considerably (Proposition 5, Part (i)). We find that the city skew's role in our previous analysis now is assumed by the federal skew. Remarkably, landlord welfare continues to be invariant w.r.t. how landlords and tenants are assigned to cities or city rings. We do not need to consult ring-specific, city specific or even ring-and-city specific tenant shares here. Rather, we may compute federal landlord welfare W_L from aggregate figures once the share of tenants in the overall population θ , the average city width $\bar{r} = \sum_{j=1}^J \tilde{r}_j S_j / S$ and the federation's mean wage $\bar{\omega} = \sum_{j=1}^J \omega_j S_j / S$ and the federal skew $\bar{\sigma}$ are known. (An example of this follows shortly.)

Let a differentiable federal housing shape function $\Phi(r)$, with $\Phi(\tilde{r}_j) = S$, summarize all available housing between the system's many CBDs and housing located r units of distance out, irrespective of city location. We may approximate the number of dwellings in all one-unit wide rings i , S_i , by $\Phi'(r_i) \equiv \phi(r_i)$. Proposition 5's Part (ii) supplies the expression for federal landlord class welfare as ring width gets ever smaller.

Proposition 5 (Commuting Distribution Skew and Political Economy)

- (i) (*"Wide Rings"*): *Let cities be divided into rings of identical width Δr . Then federal landlord class welfare is $W_L = tS\bar{\sigma} + (1 - \theta)S\bar{\omega}$.*

¹²Average emissions $\bar{\rho}$ also depend on the modal split, and hence aggregate travelling distances ATC/t are only a first step towards assessing GHG emissions.

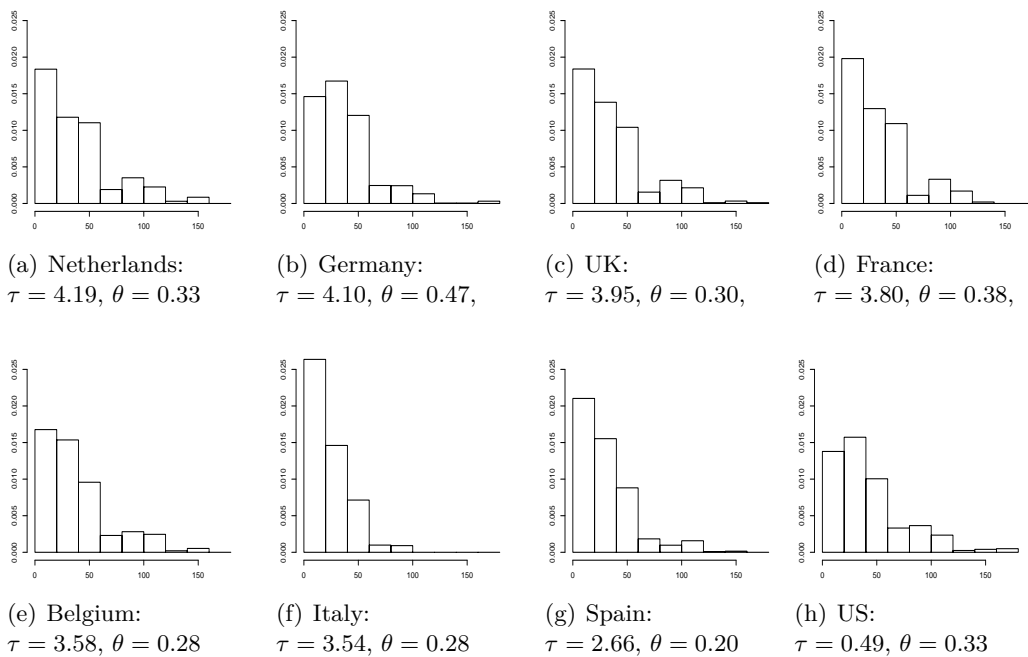


Figure 3: Country Commuting Distributions

(ii) (“Thin Rings”): Let ring width become arbitrarily small. Federal landlord welfare approaches

$$W_L = t \int_0^{\tilde{r}^J} \phi(r) (\theta \bar{r} - r) dr + S(1 - \theta) \bar{\omega} = tS(\theta \bar{r} - \bar{\rho}) + S(1 - \theta) \bar{\omega}, \quad (9)$$

where $\bar{\rho}$ is $\int_0^{\tilde{r}^J} (\phi(r)/S)r dr$, the federation’s mean commuting distance.

Proposition’s Part (ii) also insinuates that the landlord class welfare’s short run (i.e. fixed wage) response can be decomposed into the national commuting distribution’s skew plus a (typically negative) term representing the joint influence of mean city width and overall tenant share, i.e.

$$\bar{\sigma} = (\bar{r}/2 - \bar{\rho}) + \bar{r}(\theta - (1/2)). \quad (10)$$

Equation (10) generalizes the decomposition $\sigma = \tilde{r}/2 - \rho$ identified earlier (section 3). We briefly explore the uses of decomposition (10) using available micro data on the distribution of commuting lengths. The sample data underlying the following diagrams come from the 2010 European Survey of Working Conditions for European countries (ESWC). (Appendix B provides more details.) Seven European histograms for commuting time (truncated at 180 minutes) are shown in Figure 3’s panels (a) through (g), with countries sorted by the gasoline tax they charge. Since the underlying data have been collected as part of the same survey they should by and large be comparable. The eighth panel, panel (h), merely reproduces the US commuting distribution already shown in Figure 1’s panel (a).

Suppose Figure 3’s European countries share the same mean city width \bar{r} . Then differences in landlord incentives come down to (i) differences in the commuting distribution’s skew and (ii) differences in the national tenant share. By visual inspection, Figure 3 appears

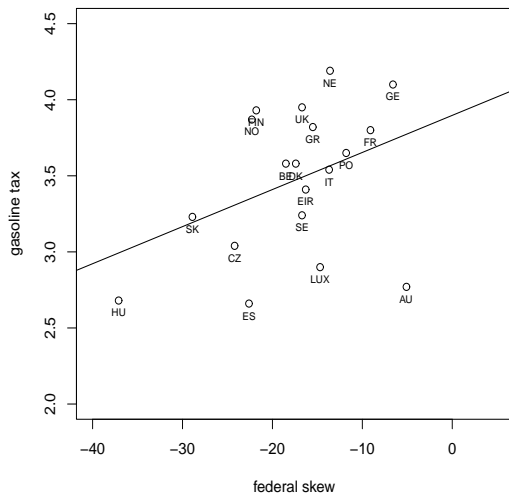


Figure 4: Federal Silhouette Skew and Gasoline Tax

to confirm the idea that countries with a higher gasoline tax boast a stronger skew, a larger tenant share, or even both. In that sense our model explains why the Netherlands, Germany, the UK, and France (panels in the Figure’s top row) exhibit higher gasoline taxes than Belgium, Italy, Spain and the US (bottom row). In fact, combining skew and tenants may even help explain why: the Netherlands have a higher tax than the UK (greater skew, more tenants), why France has a larger tax than Belgium (greater skew, more tenants), and why Italy (home to many of the world’s most famous compact cities such as Siena or Perugia) has a greater tax than Spain (greater skew, more tenants). Given its high tax on gasoline, Germany’s federal skew seems surprisingly small. Nonetheless even Germany may fit into our mold once we acknowledge its extraordinarily large share of tenants.

Of course \bar{r} is not the same across countries. Here we may try to roughly estimate European countries’ average city width $t\bar{r}$, by averaging over regional maximum commuting lengths (as briefly explained in Appendix B, Part (iv)). Appendix B’s table has the resulting data on federal skewness $\bar{\sigma}$ underlying Figure 4. There we show a scatter plot of gasoline tax t against the skew of the federal commuting distribution $\bar{\sigma}$, again for those European countries that feature in both Knittel’s (2012) Table 1 and ESWC 2010. The regression line indicates a positive correlation between these two variables. Similarly, skewness may also be important when comparing European countries with the US, rather than with each other – as discussed in the context of Figure 1.¹³

6 Tenant Support for the Carbon Tax

This section at last allows for malleable housing, and hence assumes a long run perspective. Let $q(r, t)$ be a complete list of rents $(q_1(r, t), \dots, q_J(r, t))$. We let $a(r)$ capture

¹³A rigorous empirical analysis is postponed to a future paper (also see the discussion in section 7).

all land available at distance r from all cities' CBDs, $a(r)$. And we let $g(q(r, t))$ capture construction on each unit of this land, as arising from profit maximizing developers' decisions (Brueckner (1987)). Then aggregate available housing in all rings r units of distance away from cities' CBDs and one unit wide is approximately equal to the product $\phi(r, t) = a(r)g(q(r, t))$. This product effectively explains the federation's silhouette.

Proposition 6 assesses the various effects of a marginal increase in the commuting cost parameter t . First, and as before, tenant utility u must fall (Part (i)). Intuitively, utility cannot but fall because if it were to rise (or only to remain at its initial level) costs-of-living $t\tilde{r}_j$ would have to fall throughout the urban system, implying that aggregate housing supply would necessarily fall short of aggregate demand. Imparting this information to the first two sets of equations in (8) implies that wages fall and costs-of-living rise (Proposition 6, Part (ii)). Rent in city j is $q_j(r) = t(\tilde{r}_j - r)$, given intra-city spatial equilibrium. Differentiating rent with respect to t and equating that derivative with zero implies the following cutoff \hat{r}_j :

$$\hat{r}_j = \tilde{r}_j + t \frac{d\tilde{r}_j}{dt} = \frac{d(t\tilde{r}_j)}{dt} = \frac{dq_{j0}}{dt} = \hat{r} \quad (11)$$

This cutoff \hat{r}_j is identical for all cities, and hence simply is \hat{r} . Now, for distances beyond \hat{r} rents fall and housing contracts, while for distances below \hat{r} rents and housing supply increase. Moreover, initially small cities grow, while initially large cities contract (Part (iii)).¹⁴ Consequently emissions respond to the carbon tax, too. Fundamentally, the urban system becomes greener, by emitting less GHG (Part (iv)).

If individuals' concern also is with global GHG emissions, or $S \sum_{l=1}^K \bar{\rho}_l$, where l is the country index and K is the total number of (equally sized) countries in the world, then we might add a disutility term $-v(S \sum_{l=1}^K \bar{\rho}_l)$, with $v', v'' > 0$. Now introducing the carbon tax holds out the promise of reducing federal GHG emissions. For given GHG emissions elsewhere, now not only landlords will vote for the carbon tax. A share of society's tenants will vote for it, too (Part (v)). Emission mitigation provides an important additional source of voter support for the carbon tax. This is especially true if not all landlords espouse the tax.

Proposition 6 (Carbon Tax, Greener Cities, Tenant Carbon Tax Support)

- (i) *(Tenant Welfare):* Tenant class welfare W_M is decreasing in t .
- (ii) *(Local Prices):* Wages ω_j are decreasing, while costs-of-living $t\tilde{r}_j$ are increasing, in t .
- (iii) *(Federal Compactification):* Cities initially smaller than \hat{r} expand, while cities larger than \hat{r} initially contract. And building height is increasing (decreasing) in t at all inhabited distances short of (beyond) \hat{r} .
- (iv) *(Urban Greenness):* Emissions $\bar{\rho}$ are decreasing in t .
- (v) *(Tenant Support for the Carbon Tax):* Suppose individuals worry about climate change, and that this worry is uniformly distributed across them. Now the carbon tax attracts a fraction of climate change averse tenants, too.

For completeness, let the landlord class be free in its choice of carbon tax now. Where

¹⁴Formally, the cutoff \hat{r} need not be smaller than \tilde{r}_j . Equation (11) tells us that for growing cities ($d\tilde{r}_j/dt > 0$) cutoff \hat{r}_j is further out than the initial urban boundary \tilde{r}_j , and the opposite is true for contracting cities.

landlord class welfare is $St(\theta\bar{r} - \bar{\rho}) + (1 - \theta)S\bar{\omega} - (1 - \theta)v(S\sum_{l=1}^K \bar{\rho}_l)$, setting welfare's derivative with respect to t equal to zero defines the government's optimal choice. Now consider the landlord-weighted average of both sides of the first equation in (11). This is $d\bar{r}/dt = (\hat{r} - \bar{r})/t$. Substituting the expression on the r.h.s. for $d\bar{r}/dt$ in the first order condition, employing $\hat{r} = dq_{j\circ}/dt = -d\omega_j/dt$ and simplifying the resulting equation yields

$$- \left((1 - \theta)v' + t \right) \frac{d\bar{\rho}}{dt} + (1 - 2\theta) \frac{d\bar{\omega}}{dt} - \bar{\rho} = 0. \quad (12)$$

The optimal carbon tax strikes the balance between these natural, model-induced benefits and costs.¹⁵ On the one hand, the condition's first term represents both the environmental and reduced-commuting-distances marginal gains, recalling $d\bar{\rho}/dt < 0$ (Proposition 6, Part (iv)). On the other hand, the second and third term summon the greater marginal cost induced by a loss in wages and more expensive commuting, given that $d\bar{\omega}/dt < 0$ (Part (ii)).

7 Discussion

The paper generalizes its basic model into various directions, but further modifications may of course be thought of. This section provides a short overview over possible extensions.

Polycentric Cities: Throughout we have assumed that cities are monocentric. Yet real cities are polycentric. Following Brueckner (2011), such cities might also be framed as unions of smaller monocentric ones. Then of course it is the skewness of each of these smaller cities that matters, rather than the skewness of their union. Besides, note that this paper may even help explain why cities become polycentric. Recall from Proposition 1 that landlord class welfare equals $st\sigma$ in the basic model. Now consider a slightly modified representative city with (i) a negative skew (e.g., as in Figure 2's panels (c), (d), and (g)) and (ii) a ring road circling along the city boundary and costless to travel.

Sprawl: Suppose landlords pushed a city's shops from their initial location at the CBD to the city's "other end", somewhere along the ring road. This amounts to replacing each distance r_i in our expression for landlord welfare $w_l = \sum_{i=1}^n ((\tilde{r}/2) - r_i)s_i$ by $\tilde{r} - r_i$. As the new welfare w'_l the landlord class obtain the negative of the initial one, $w'_l = -w_l$. So cities with negative skew (landlord class welfare) initially "suddenly" achieve positive skew (landlord class welfare), "simply" by turning the initial pattern of commutes on its head. This approach may help explain why some cities encourage large shopping center developments at their gates, or decentralization of tertiary employment more generally, while others do not (Glaeser/Kahn (2004), Wheaton (2004)).

Owner-Occupiers: One might argue that owner-occupiers deserve a separate treatment in the model. Owner-occupiers are likely to be against the carbon tax (even though subletting part of their property may help some of them offset the extra commuting costs). Typically

¹⁵Except for the analysis of a variation in θ , comparative statics on this condition are not straightforward, however. And while the analysis of a change in θ is straightforward, this analysis not only must rely on the traditional ad-hoc assumption that $W_{tt} < 0$ but also only produces an ambiguous sign, given that $W_{t\theta}$ may either be positive or negative.

owning a flat in a condominium comes along with greater transaction costs than owning a semi-detached house. We expect housing tenure to decrease with density. Thus in a typical city owner-occupiers may coincide with the peripheral housing stock. Then to assess the interests of the landlord class proper we may refer to the skewness of a modified silhouette, being the silhouette that is obtained once the distribution of commuting distances is truncated by owner-occupiers' commuting distances.

Tax Refund: In the basic model, a tax of one Euro generates tax revenues equal to the aggregate travelling distance, or $s\rho$. If these revenues were channeled back lump sum, then the landlord class's welfare change would become a weighted average of silhouette skew and tax revenues, equal to $s((\tilde{r}/2) - \rho) + (s\rho)/2$ or $s(\tilde{r} - \rho)/2$. At the same time, the tenant class' welfare change becomes $s(\rho - \tilde{r})/2$, or just the opposite of the landlord class' welfare change. Refunding the carbon tax makes landlords and tenants pursue diametrically opposed policies, and weakens the link between urban form and carbon tax.

Non-Linear Travel Costs: In our basic model, commuting costs are linear. Let us replace tr_i by $h(r_i)$ and $t\tilde{r}$ by $h(\tilde{r})$, with h some increasing function of t . Then our derivation of w_l in (3) goes through virtually unchanged. Landlord welfare in the non-linear transport cost case becomes $w_l = \sum_{i=1}^n (h(\tilde{r})/2 - h(r_i))s_i$. So while corresponding ring population figures s_i and s_{n+1-i} this far receive equal weight in the expression for landlord welfare, with non-linear commuting cost this would no longer be the case. We suspect the principles of subsequent analysis to remain unchanged, if not more difficult to expound.

Open Cities: Many cities embark on climate policies of their own (Millard-Ball (2012)), and some of their policies have a direct impact on commuting also, such as implementing bus and bicycle lanes. These cities are open, instead of closed, as they must take into account the effect their policies have on their mobile population. College towns are frequently thought to be particularly "green". From this paper's perspective, this is not because these towns' inhabitants are particularly "progressive" but because these towns are full of tenants. A high local share of tenants tempts landlords to drive rents up, by raising the costs of commuting.

Empirical Analysis: Testing this paper's silhouette skew/climate policy-nexus seems desirable. Yet the silhouette skew's covariates will require us to go a great deal further than we went here, as will mutual causation. For example, strong silhouette skews may expose buildings more to accelerating winds (Roaf/Crichton/Nichol (2009)), a society's carbon tax choice is likely to feed back into its silhouettes, etc. At the same time, exploiting the strong existing variation in local commuting policies (again: Millard-Ball (2012)) offers one way out. Variation in local climate policies, silhouettes, and tenant-landlord splits call for testing this paper's model in a sample of open cities.

8 Conclusions

The paper's premise is that the existing variation in greenhouse gas mitigation policies is not well explained by pointing to inter-country differences in income, in climate change exposure, or in carbon resource endowments. We offer an alternative explanation of why

climate policies vary. In the paper's model, a country's carbon taxation increases in its city silhouettes' skew. Subsequent extensions to this theme do not turn over this silhouette-skewness/climate-policy-nexus. Put differently, the interests of the landlord class manifest themselves in the country's urban form. From an environmental perspective, this observation may further our understanding of climate policies' determinants. From an urban perspective, this observation may further our understanding of whether a city's form informs us about its inhabitants' preferences – a theme that has traditionally been urban historians' domain.

9 Literature

- Altemeyer-Bartscher, M., Markandya, A. and D. Rübbelke (2011), The Private Provision of International Impure Public Goods: the Case of Climate Policy. BC3 Working Paper Series 2011-09. Basque Centre for Climate Change.
- Arnott, R. and J. Stiglitz (1982) Aggregate Land Rent and Aggregate Transport Costs, *Economic Journal* 91: 331-347.
- Bento, A., S. Franco and D. Kaffine (2006) The Efficiency and Distributional Impacts of Alternative Anti-Sprawl Policies, *Journal of Urban Economics*, *Journal of Urban Economics* 59.1: 121-141W.
- Bertaud, A. (2004) The Spatial Organization of Cities, Institute of Urban and Regional Development, University of California at Berkeley.
- Bertaud, A. and J. Brueckner (2005) Analyzing building-height restrictions: Predicted Impacts and Welfare Costs, *Regional Science and Urban Economics* 35.2: 109-125.
- Borck, R. (2014) Will Skyscrapers Save the Planet? Unpublished Manuscript.
- Borck, R. and M. Wrede (2005) Political Economy of Commuting Subsidies, *Journal of Urban Economics* 57: 478-499.
- Brueckner, J. (1987) The Structure of Urban Equilibrium, in: E. Mills (ed.) *Handbook of Regional and Urban Economics* Vol. 2, North-Holland:
- Brueckner, J. (2005) Transport Subsidies, System Choice, and Urban Sprawl, *Regional Science and Urban Economics* 35.6: 715-733.
- Brueckner, J. (2011) *Lectures on Urban Economics*, MIT Press.
- Bunting, T., P. Filion and H. Priston (2002) Density Gradients in Canadian Metropolitan Regions 1971-1996, *Urban Studies* 39: 2531-2552.
- Carraro, C. and D. Siniscalco (1998) International Environmental Agreements: Incentives and Political Economy, in: *European Economic Review* 42.3-5: 561-572.
- Desmet, K. and E. Rossi-Hansberg (2012) On the Spatial Impact of Global Warming, NBER Working Paper 18546.
- Fujita, M. (1989) *Urban Economic Theory: Land Use and City Size*, Cambridge University Press.
- Gagné, C., St. Riou, and J.F. Thisse (2012) Are Compact Cities Environmentally Friendly?, *Journal of Urban Economics* 72.2: 123-136.
- Giesen, K., J. Südekum, and J. Zimmermann (2010) The Size Distribution across All Cities – Double Pareto Log-normal Strikes, *Journal of Urban Economics* 68: 129-137.
- Glaeser, E. (2011) *Triumph of the City*, Macmillan.

- Glaeser, E. and M. Kahn (2004) *Sprawl and Urban Growth*, in: J. Henderson and J.-F. Thisse, *Handbook of Regional and Urban Economics*: 2481-2527.
- Glaeser, E. and M. Kahn (2010) *The Greenness of Cities. Carbon Dioxide Emissions and Urban Development*, *Journal of Urban Economics* 67.3: 404-418.
- (IPCC) Intergovernmental Panel on Climate Change (2013) *Fifth Assessment Report*.
- Kahn, M. (2006) *Green Cities. Urban Growth and the Environment*, Washington: Brookings.
- Kim, J. and D. Brownstone (2013) *The Impact of Residential Density on Vehicle Usage and Fuel Consumption*, University of California Transportation Center Working Paper.
- Knittel, Ch. (2010) *Reducing Petroleum Consumption from Transportation*, *Journal of Economic Perspectives* 26.1: 93-118.
- de Lara et al. (2013) *Congestion Pricing and Long Term Urban Form. Application to Paris Region*, *Regional Science and Urban Economics* 43.2: 282-295.
- Larson, W., L. Feng and L. Yezer (2012) *Energy Footprint of the City: Effects of Urban Land use and Transportation Policies*, *Journal of Urban Economics* 72.2: 147-159.
- Lopate, Ph. (2004) *Waterfront. A Walk around Manhattan*, Anchor Books.
- Millard-Ball, A. (2012) *Do City Climate Plans Reduce Emissions?*, *Journal of Urban Economics* 71.3: 289-311.
- Mohring, H. (1961) *Land Values and the Measurement of Highway Benefits*, *Journal of Political Economy* 69.3: 236-249.
- Nitsch, V. (2004) *Zipf Zipped*, *Journal of Urban Economics* 57: 86-100.
- Nordhaus, W. (2013) *The Climate Casino. Risk, Uncertainty, and Economics for a Warming World*, New Haven: Yale University Press.
- Pucher, J. (1988) *Urban Travel Behavior as the Outcome of Public Policy: The Example of Modal Split in Western Europe and North America*, *Journal of the American Planning Association* 54: 509-520.
- Roaf, S., D. Crichton, and F. Nichol (2009) *Adapting Buildings and Cities for Climate Change*, Architectural Press.
- Roback, J. and Sh. Rosen (1982) *Wages, Rents, and the Quality of Life*, *Journal of Political Economy* 90.6: 1257-1278.
- Wheaton, W. (1974) *A Comparative Static Analysis of Urban Spatial Structure*, *Journal of Economic Theory* 9: 223-237.
- Wheaton, W. (2004) *Commuting, Congestion, and Employment Dispersal in Cities with Mixed Land Use*, *Journal of Urban Economics* 55: 417-438.

Appendix A

Proof of Proposition 1: (Political Economy and Urban Form)

(iii) (Landlord Class Welfare) Note first that A may be decomposed into the simple sum of two even more strongly patterned matrices:

$$A = \begin{pmatrix} -tr_1 & \dots & -tr_1 \\ \vdots & & \vdots \\ -tr_n & \dots & -tr_n \end{pmatrix} + \begin{pmatrix} t(\tilde{r} - r_1) & \dots & t(\tilde{r} - r_n) \\ \vdots & & \vdots \\ t(\tilde{r} - r_1) & \dots & t(\tilde{r} - r_n) \end{pmatrix} \quad (13)$$

where the first (commuting costs) matrix is labeled A_1 and the second (rents) matrix is referred to as A_2 . The value of this decomposition lies in representing A as the sum of two matrices that either have identical rows (in the case of A_1) or identical columns (as with A_2).

With this decomposition we may alternatively rewrite landlords' welfare (2) as

$$\begin{aligned} w_l &= \iota' (B \circ A) \iota &= \iota' (B \circ (A_1 + A_2)) \iota \\ & &= \iota' (B \circ A_1 + B \circ A_2) \iota \\ & &= \iota' (B \circ A_1) \iota + \iota' (B \circ A_2) \iota. \end{aligned} \quad (14)$$

Here the third equality makes use of the distributive law for Hadamard products and the fourth equality conforms to the standard rules of conventional matrix multiplication.

We analyze the sum on the r.h.s. of the last equality in two steps. First consider the second term here. In it the expression $\iota' (B \circ A_2)$ is nothing but a $1 \times n$ vector exhibiting $t(\tilde{r} - r_1) \cdot \sum_i b_{i1}$ in column 1, $t(\tilde{r} - r_2) \cdot \sum_i b_{i2}$ in column 2, $t(\tilde{r} - r_3) \cdot \sum_i b_{i3}$ in column 3, etc. Yet these products in turn reduce to $t(\tilde{r} - r_1)m_1$, $t(\tilde{r} - r_2)m_2$, $t(\tilde{r} - r_3)m_3$ etc. respectively because the sum of all entries in column i of B just represents total tenants in ring i , by definition of B . We conclude that the second term on the last line of (14) is

$$\iota' (B \circ A_2) \iota = \sum_{i=1}^n t(\tilde{r} - r_i) m_i. \quad (15)$$

We can trace through a similar argument when analyzing the first term on the r.h.s. of the last equality in (14). There the expression $(B \circ A_1) \iota$ is nothing but the $n \times 1$ vector containing $(-tr_1) \cdot \sum_j b_{1j}$ in row 1, $(-tr_2) \cdot \sum_j b_{2j}$ in row 2, $(-tr_3) \cdot \sum_j b_{3j}$ in row 3 etc. These latter sums may be rewritten as $(-tr_1)l_1$, $(-tr_2)l_2$, $(-tr_3)l_3$, etc. Thus we have shown that

$$\iota' (B \circ A_1) \iota = \sum_{i=1}^n (-tr_i) l_i, \quad (16)$$

We join equations (14), (15) and (16), and simplify. This yields the first expression in equation (3). \square

Proof of Proposition 3: (Silhouette Skew and Political Economy)

(i) (Greatest Lower Bound): From the main text we recall that all entries on (above, below) A 's counterdiagonal are equal to (greater than, smaller than) zero, with A defined as in (1).

Now, if $s_1 - s_n > 0$ read on, else head for the next paragraph. It is conceivable (if unlikely) that all residents in ring n are matched up with some resident in ring 1. There are at most $s_1 - s_n$ matches for which $a_{1n} = a_{n1} = 0$. All remaining matches possess strictly positive value. So $(s_1 - s_n)/2$ is one first lower bound.

If $s_1 + s_2 - (s_n + s_{n-1}) > 0$ read on, else proceed to the third paragraph. Note that $a_{1n} = a_{n1}$, $a_{2,n-1} = a_{n-1,2}$ and $a_{2n} = a_{n2}$ all are non-positive. It is possible (if unlikely) that all residents in rings $n - 1$ and n are matched with residents in rings 1 and 2. So $(s_1 + s_2 - (s_n + s_{n-1}))/2$ is another lower bound.

Next consult $(s_1 + s_2 + s_3 - (s_n + s_{n-1} + s_{n-2}))/2$, etc. Proceeding in this way we compute cumulative ring differences whenever they are positive. We ultimately end up with a set of $n/2$ positive lower bounds at best. From this set of lower bounds we pick the one that is greatest, representing the cumulative sum of ring differences from 1 out to n^* . \square

(ii) (Greatest Lower Bound): Let $0 = r_0 < r_1 < \dots < r_{n'-1} < r_{n'} = r'$ represent a partition of $[0, r']$. We approximate the sum $\sum_{i=1}^{n'} (s_i - s_{n+1-i})$ by setting

$$\sum_{i=1}^{n'} (s_i - s_{n+1-i}) \approx \sum_{i=1}^{n'} (f(\hat{r}_i) - f(\hat{r}_{n+1-i})) \Delta r,$$

where \hat{r}_i is from the open interval (r_{i-1}, r_i) . Then we let ring width Δr converge to zero. This has this latter expression tend to $[F(r') - (s - F(\tilde{r} - r'))]$, or $[F(r) - (s - F(\tilde{r} - r))]$ after dropping the prime.

To identify the greatest from all these lower bounds we maximize $[F(r) - (s - F(\tilde{r} - r))]$ with respect to $r \in [0, \tilde{r}/2]$. This requires the optimal r' , labeled r^* , to satisfy:

$$f(r^*) = f(\tilde{r} - r^*),$$

if r^* is contained in $(0, \tilde{r}/2)$. Substituting r^* back into the maximand gives the greatest lower bound, equal to $l^*(r) = [F(r^*) - (s - F(\tilde{r} - r^*))]/2$. \square

(iii) (Landlord Beneficiaries): Let $0 = r_0 < r_1 < \dots < r_{n-1} < r_n = \tilde{r}$ represent a partition of $[0, \tilde{r}]$. We depart from the last expression in (3), an expression we approximate by the following Riemann sum:

$$t \sum_{i=1}^n f(\hat{r}_i) (\tilde{r}/2 - \hat{r}_i) \Delta r \quad (17)$$

where \hat{r}_i is from the open interval (r_{i-1}, r_i) . Since $f(r)(\tilde{r}/2 - r)$ is continuous in r , the resulting sequence of Riemann sums converges to

$$w_l = t \int_0^{\tilde{r}} f(r) (\tilde{r}/2 - r) dr = t s \sigma, \quad (18)$$

as we let Δr tend to 0. Now let us rewrite $s\sigma$ as in

$$\begin{aligned} s\sigma &= \int_0^{\tilde{r}/2} f(r) (\tilde{r}/2 - r) dr + \int_{\tilde{r}/2}^{\tilde{r}} f(r) (\tilde{r}/2 - r) dr \\ &= \dots + \int_{-\tilde{r}/2}^{-\tilde{r}} f(-r) (\tilde{r}/2 + r) dr && \text{(reflection)} \\ &= \dots + \int_{\tilde{r}/2}^0 f(-(r - \tilde{r})) (-\tilde{r}/2 + r) dr && \text{(translation)} \\ &= \dots - \int_0^{\tilde{r}/2} f(\tilde{r} - r) (\tilde{r}/2 - r) dr = \int_0^{\tilde{r}/2} (f(r) - f(\tilde{r} - r)) (\tilde{r}/2 - r) dr, \end{aligned}$$

where the second equality follows from reflecting the second term in the sum on the r.h.s. across the vertical axis, while the third equality follows from translating the second term in the sum on the r.h.s. by \tilde{r} to the right. The last equality represents $s\sigma$ as the weighted sum of those ring differences ($f(r) - f(\tilde{r} - r)$) prominent in the text. Breaking up this last expression for $s\sigma$ further gives

$$\begin{aligned} s\sigma &= (\tilde{r}/2) \int_0^{\tilde{r}/2} (f(r) - f(\tilde{r} - r)) dr - \int_0^{\tilde{r}/2} (f(r) - f(\tilde{r} - r)) r dr \\ &= (\tilde{r}/2) (F(\tilde{r}/2) - (s - F(\tilde{r}/2))) - \int_0^{\tilde{r}/2} (f(r) - f(\tilde{r} - r)) r dr \quad (19) \\ &\leq \tilde{r}l^*. \end{aligned}$$

After all, by the property of r^* being optimal,

$$(\tilde{r}/2) (F(\tilde{r}/2) - (s - F(\tilde{r}/2))) \leq (\tilde{r}/2) (F(r^*) - (s - F(r^*))) = \tilde{r}l^*.$$

So the first term in brackets on the r.h.s. of (19) falls short of $\tilde{r}l^*$. On the other hand, the second term in brackets, representing a distance-weighted average of ring differences, is positive by the Proposition's assumption. We conclude that $s\sigma \leq \tilde{r}l^*$. \square

Proof of Proposition 4 (Building Height Restrictions and Silhouette Skew):

(i) We first turn to σ . We state σ as

$$\sigma = s(\tilde{r}/2) (F(\tilde{r}/2, \delta)/s) - s \int_0^{\tilde{r}} (f(r, \delta)/s) r dr.$$

This casts F/s as the c.d.f. of r , and f/s as its p.d.f. Less stringent zoning is reflected by an increase in δ . Now, by the Proposition's assumption an increase in δ makes F increase at any point r , i.e. $F(r, \delta'') > F(r, \delta')$ for all $\delta'' > \delta'$ and $r \in (0, \tilde{r})$. Effectively we are assuming stochastic dominance on the part of F/s .

On the one hand, this implies that the first term on the r.h.s. of (20) is increasing in δ . On the other hand, this implies that the second term on the r.h.s. of (20) is decreasing in δ (because with stochastic dominance the mean of r is). Combining these observations implies that σ is increasing in the “non-zoning-parameter” δ .

We next turn to l^* . We first restate l^* as $l^* = [F(r^*, \delta) - (s - F(\tilde{r} - r^*, \delta))]/2$, where r^* satisfies

$$f(r^*, \delta) = f(\tilde{r} - r^*, \delta), \quad (20)$$

from (6). By the envelope theorem, the marginal effect of δ on l^* is given by

$$\frac{\partial l^*(r^*, \delta)}{\partial \delta} = F_\delta(r^*, \delta) + F_\delta(\tilde{r} - r^*, \delta).$$

This latter derivative is strictly positive. Hence the greatest lower bound l^* is increasing in δ . \square

(ii) Consider landlords who are made strictly worse off by the carbon tax. Following an argument similar to that pursued in Proposition 3's Parts (i) and (ii) the greatest lower bound on landlord opponents against the carbon tax, denoted l^{**} , is found by maximizing $[s - F(\tilde{r} - r, \delta) - F(r, \delta)]/2$ with respect to r . Let r^{**} denote the corresponding maximizer.

Then l^{**} is given by $[s - F(\tilde{r} - r^{**}, \delta) - F(r^{**}, \delta)]/2$, where $r^{**} \in (0, \tilde{r}/2)$ satisfies

$$(f(r^{**}, \delta) = f(\tilde{r} - r^{**}, \delta). \quad (21)$$

By the envelope theorem, the marginal effect of δ on l^{**} is given by

$$\frac{\partial l^{**}(r^{**}, \delta)}{\partial \delta} = -F_\delta(\tilde{r} - r^{**}, \delta) - F_\delta(r^{**}, \delta).$$

This derivative is strictly negative. Hence the greatest lower bound on the number of landlord opponents to the carbon tax, l^{**} , is decreasing in δ . \square

Proof of Proposition 5: (Commuting Distribution Skew, Political Economy)

(i) (“Wide Rings”): W_L can be written as follows:

$$\begin{aligned} W_L &= t \sum_{j=1}^J \sum_{i=1}^{\bar{n}} m_{ji}(\tilde{r}_j - r_i) - t \sum_{j=1}^J \sum_{i=1}^{\bar{n}} l_{ji}r_i + \sum_{j=1}^J \sum_{i=1}^{\bar{n}} l_{ji}\omega_j \\ &= t \sum_{j=1}^J \sum_{i=1}^{\bar{n}} s_{ji}\tilde{r}_j - t \sum_{j=1}^J \sum_{i=1}^{\bar{n}} s_{ji}r_i + \sum_{j=1}^n L_j (\omega_j - t\tilde{r}_j). \\ &= t \sum_{j=1}^J \tilde{r}_j \sum_{i=1}^{\bar{n}} s_{ji} - t \sum_{i=1}^{\bar{n}} r_i \sum_{j=1}^J s_{ji} + (\bar{\omega} - t\bar{r}) \sum_{j=1}^n L_j \\ &= t \sum_{j=1}^J \tilde{r}_j S_j - t \sum_{i=1}^{\bar{n}} r_i S_i + L(\bar{\omega} - t\bar{r}) \\ &= t S\bar{r} - t \sum_{i=1}^{\bar{n}} r_i S_i + (1 - \theta)S(\bar{\omega} - t\bar{r}) \\ &= t \sum_{i=1}^{\bar{n}} S_i \theta\bar{r} - t \sum_{i=1}^{\bar{n}} S_i r_i + (1 - \theta)S\bar{\omega} \\ &= t \sum_{i=1}^{\bar{n}} S_i (\theta\bar{r} - r_i) + (1 - \theta)S\bar{\omega} \end{aligned} \quad (22)$$

We comment on the third equation in (22). This exploits the spatial equilibrium feature of $\omega_j - t\tilde{r}_j$ being equal to u , by the first equation in (8), and hence being independent of j . Thus it must also equal its federal mean, $\bar{\omega} - t\bar{r}$, alternatively expressed as the difference between federal mean wage and average maximum travel cost, $\bar{\omega} - t\bar{r}$. \square

(ii) (“Thin Rings”): Let $0 = r_0 < r_1 \dots r_{n-1} < r_n = \tilde{r}_J$ represent a partition of $[0, \tilde{r}_J]$. We depart from the first term of the last expression in (22) (the second term in this expression being a constant), an expression we approximate by the following Riemann sum:

$$t \sum_{i=1}^n \phi(\hat{r}_i) (\theta\bar{r} - \hat{r}_i) \Delta r \quad (23)$$

where \hat{r}_i is taken from the open interval (r_{i-1}, r_i) . Note that $\phi(r)(\theta\bar{r} - r)$ is continuous in r . Hence, if we let Δr tend to 0 the resulting sequence of Riemann sums converges to

$$W_L = t \int_0^{\tilde{r}_J} \phi(r) (\theta\bar{r} - r) dr \quad \square \quad (24)$$

Proof of Proposition 6: (Carbon Tax, Greener Cities, Tenant Support)

(i) (Tenant Welfare): Suppose first that u remains constant. Then wages ω_j or costs-of-living $t\tilde{r}_j$ do not change either, given that both these concepts are tied to their initial levels by the first two sets of equations in (8) for as long as tenant utility u is given. Yet if costs-of-living do not change rents throughout every city must fall in response to the tax (except for at the center), implying a reduction in the economy's aggregate housing supply. Alternatively, suppose that u increases in response to the tax increase. Then costs-of-living would even have to fall, again implying a fall in rents throughout the city. We conclude that tenant utility must fall. \square

(ii) (Compensating Differentials): Differentiating the two equations of (8) with respect to t gives $dq_{j\circ}/dt = -(1/2)(du/dt)$. Next, differentiating $q_j(r) = t(\tilde{r}_j - r)$ with respect to t , rearranging and inserting the previous derivative gives

$$\frac{d\tilde{r}_j}{dt} = -\frac{1}{2t} \frac{du}{dt} - \frac{\tilde{r}_j}{t}. \quad (25)$$

Taking the derivative of federal land market equilibrium $\int_0^{\tilde{r}_J} \phi(q(r, t))dr = S$, making use of Leibniz' rule, and making appropriate substitutions for dq_j/dt and $d\tilde{r}_j/dt$, yields

$$\phi(\tilde{r}_J) \left(-\frac{1}{2t} \frac{du}{dt} - \frac{\tilde{r}_J}{t} \right) + \int_0^{\tilde{r}_J} \frac{d\phi}{dq} \left(-\frac{1}{2} \frac{du}{dt} - r \right) dr = 0. \quad (26)$$

Rearranging the last equation can be used to solve for du/dt , which we already know to be strictly negative (Part (i)). Backsubstituting this solution further into $dq_{j\circ}/dt = -(1/2)(du/dt)$ yields the change in center rent, so $dq_{j\circ}/dt > 0$. Finally, the wage change is given by $d\omega_j/dt = -dq_{j\circ}/dt < 0$. \square

(iii) (Compactification): According to (11), $d\tilde{r}_j/dt > 0$ for all cities for which $\tilde{r}_j < \hat{r}$. The reverse is true for all cities for which $\tilde{r}_j > \hat{r}$. We conclude that cities initially smaller than \hat{r} grow, while cities initially larger than \hat{r} contract. Further, as discussed in the main text,

$$\frac{d\phi(t)}{dq} \frac{dq}{dt} \geq 0 \quad \text{iff } r \leq \hat{r}.$$

Every city's silhouette increases at distances short of \hat{r} , yet decreases at distances beyond \hat{r} . \square

(iv) (Urban Greenness): The marginal change in the aggregate commuting distance is

$$\begin{aligned} S \frac{d\bar{\rho}}{dt} &= \int_0^{\tilde{r}_J} \left(\frac{d\phi}{dt} \right) r dr + \phi(\tilde{r}_J) \tilde{r}_J \left(\frac{d\tilde{r}_J}{dt} \right) \\ &< \int_0^{\hat{r}} \left(\frac{d\phi}{dt} \right) \hat{r} dr + \int_{\hat{r}}^{\tilde{r}_J} \left(\frac{d\phi}{dt} \right) \hat{r} dr + \phi(\tilde{r}_J) \tilde{r}_J \left(\frac{d\tilde{r}_J}{dt} \right) \\ &< \int_0^{\hat{r}} \left(\frac{d\phi}{dt} \right) \hat{r} dr + \int_{\hat{r}}^{\tilde{r}_J} \left(\frac{d\phi}{dt} \right) \hat{r} dr + \phi(\tilde{r}_J) \hat{r} \left(\frac{d\tilde{r}_J}{dt} \right) \\ &= \hat{r} \underbrace{\left(\int_0^{\tilde{r}_J} \left(\frac{d\phi}{dt} \right) dr + \phi(\tilde{r}_J) \left(\frac{d\tilde{r}_J}{dt} \right) \right)}_0 = 0. \end{aligned} \quad (27)$$

As to the first inequality in (27), note that the cutoff \hat{r} (defined in equation (11)) puts a larger (smaller) weight on the positive (negative) integrand in the first (second) integral

than does the original r . Hence the sum of the resulting two integrals must strictly be larger. Moreover, replacing \tilde{r}_J by the smaller \hat{r} further reduces the weight of the last, negative, term. As to the second inequality in (27), note that (11) also implies that

$$(\hat{r} - \tilde{r}_J)\left(\frac{d\tilde{r}_J}{dt}\right) = t\left(\frac{d\tilde{r}_J}{dt}\right)^2 > 0.$$

But then $\tilde{r}_J(d\tilde{r}_J/dt) < \hat{r}(d\tilde{r}_J/dt)$ also. This explains the second inequality in (27). The last equality in (27), finally, exploits the fact that the derivative of aggregate housing supply with respect to t must, in the face of unyielding aggregate demand, be zero. \square

Appendix B:

The paper's small data set is:

Country	t	θ	$\bar{\rho}$	\bar{r}	$\bar{\sigma}$	$\bar{\sigma}/\bar{r}$
Spain	2.66	0.20	36.0	66.9	-22.6	-0.34
Hungary	2.68	0.10	46.2	91.2	-37.1	-0.41
Austria	2.77	0.43	33.6	66.3	-5.1	-0.08
Luxembourg	2.90	0.32	36.0	66.6	-14.7	-0.22
Czech Republic	3.04	0.21	40.2	75.9	-24.2	-0.32
Slovakia	3.23	0.10	35.2	63.5	-28.9	-0.45
Sweden	3.24	0.29	43.2	91.4	-16.7	-0.18
Ireland	3.41	0.27	42.1	95.3	-16.3	-0.17
Italy	3.54	0.28	30.6	60.4	-13.7	-0.23
Belgium	3.58	0.28	43.3	88.6	-18.5	-0.21
Denmark	3.58	0.33	42.8	77.1	-17.4	-0.23
Portugal	3.65	0.25	24.9	52.2	-11.8	-0.23
France	3.80	0.38	38.4	77.1	-9.1	-0.12
Greece	3.82	0.23	32.0	71.7	-15.5	-0.21
Norway	3.87	0.17	33.5	66.2	-22.3	-0.33
Finland	3.93	0.26	41.2	74.6	-21.8	-0.29
UK	3.95	0.30	43.3	88.4	-16.7	-0.19
Germany	4.10	0.47	42.8	77.1	-6.6	-0.09
Netherlands	4.19	0.33	45.0	94.9	-13.6	-0.14
US	0.49	0.33	47.0	–	–	–

Variables t , θ , ρ , \bar{r} , σ are retrieved, or computed, as follows: (i) t : 2010 gasoline taxes from Knittel (2012, Table 1), given in \$ per gallon. (ii) θ : 2010 US tenant share (1-homeownership rate): US Census Bureau, European countries' shares: Eurostat. (iii) European commuting data are on those 19 countries that feature both in Knittel (2012) and the European Survey on Working Conditions 2010 (ESWC) (variable q31). US commuting data are from the American Community Survey 2011 (ACS) (variable TRANTIME). Commuting times (in min.) are two way for the ESWC and one way for ACS, so US transit times must be multiplied with 2.¹⁶ Then $t\bar{\rho}$ simply is mean commuting time. (iv) Commuting data in (iii) also underlie $t\bar{r}$: We compute, for each country and each NUTS 2-region j , commuting length at the 90%-percentile as our estimate of $t\tilde{r}_j$. Then we calculate $t\bar{r}$ as the weighted average of the regional city width estimates, with sample region inhabitant numbers as weights.

¹⁶Ultimately we should consider earlier years for commuting times.