



Munich Personal RePEc Archive

City Shapes' Contribution to Why Donald Trump Won

Dascher, Kristof

Regensburg University, Department of Business, Economics,
Information Systems Management, and Real Estate

2020

Online at <https://mpra.ub.uni-muenchen.de/99290/>
MPRA Paper No. 99290, posted 30 Mar 2020 09:15 UTC

City Shapes' Contribution to Why Donald Trump Won

– This Version: March 15th, 2020 –

Kristof Dascher¹

Abstract: This paper identifies (unobservable) centrists and decentrists. Centrists support, whereas decentrists oppose, taxing carbon. The paper divides into two parts. Its theory derives estimators of centrists and decentrists; its empirical part provides estimates for U.S. metros and takes them to the 2008 and 2016 U.S. presidential elections. The paper finds that Donald Trump's shift away from the consensus on global warming has gained him 280,000 votes he else would not have enjoyed, in cities where decentrists were strong. The paper concludes that sprawling (compact) cities are less (more) likely to embrace carbon taxation, and provides a new rationale for globally advocating compact urban planning.

Keywords: Compact City, Global Warming, Carbon Taxation, 2016 US Presidential Election, Donald Trump, Hillary Clinton

JEL-Classifications: Q54, R52, D72

Kristof Dascher
Regensburg University
Department of Business, Economics, Information Systems Management and Real Estate
93 040 Regensburg
Germany
Email: kristof.dascher@wiwi.uni-regensburg.de

¹I am grateful to Stephen Wilson (U.S. Census Bureau) for providing me with detailed data on population by rings across U.S. metropolitan areas. Helpful comments by Christina Hanke, Alexander Haupt, Sebastian Kessing, Yarema Okhrin, and Danja Sitz. Any remaining errors are mine. Declarations of interest: none.

1 Introduction

Surely many different factors have contributed to why Donald Trump won the 2016 U.S. presidential election. While consulting the 2016-2018 archives of renowned American newspapers provides broad background and in-depth review, the literature on the 2016 U.S. presidential election appears very sparse to date. Yet understanding this election is not just important because it obviously *drives* U.S. climate policy and land use. It also is important because it *reflects* U.S. climate policy and land use. *One* of the factors underlying the election outcome, so this paper argues, is U.S. cities' shapes. These shapes shape the effects that policies towards carbon have on the electorate.²

Everything else equal, voters must have been eager to reject taxing carbon in metro areas shaped like Detroit, where an increase in commuting costs likely dominates any concomitant rise in real estate value. And they must have been willing to support taxing carbon in metro areas like New York, where rising real estate values may exceed commuting cost increases. Taxing carbon would have been a decision that befitted Hillary Clinton (who appeared concerned about global warming) but was unlikely under Donald Trump (who declared climate change a hoax and, lest voters overlook this, professed an affinity to coal). To rephrase, and at some risk of oversimplification, "sprawling" cities should have voted for Donald Trump, whereas "compact" cities should have voted for Hillary Clinton.

This is an urban landscape based explanation that relies neither on city size nor on density. Instead it builds on the independent notion of city *shape*. City shape predicts the extent to which resident-landlord-voters divide into *centrists* and *decentrists*. Centrists own their average property near the city center. Centrists welcome an increase in carbon taxes (because that makes central properties more attractive). Decentrists' properties on average lie close to the periphery. Decentrists oppose any increase in a carbon tax (since their properties become less attractive as commutes get more expensive). We provide estimators of either group's relative size, and then use sample estimates as novel regressors in a panel regression of the metro area voter share Clinton took, over and beyond size and density.

The idea that climate change denial may appeal to a fraction of the U.S. electorate for

²And since these shapes are bound to change only little over time, they may shape future elections, too.

geographical reasons is not new. Florida (2016) has raised the possibility that sprawl may have contributed to Hillary Clinton’s defeat. And Holian/Kahn (2015) have pointed to suburban living’s vulnerability to carbon cost increases. According to Holian/Kahn, suburban living – meaning bigger cars and larger homes – entails a higher cost of complying with higher taxation of carbon or stricter carbon regulation. Our “centrist” and “decentrist” shares give quantitative meaning to notions of “sprawl” (Florida (2016)) or “suburban living” (Holian/Kahn (2014)). They are reflections of city shape (Dascher (2019)), and are politically relevant. And they may even be estimated, by LP techniques.

Centrists and decentrists reveal themselves by where they own, not by where they live, and public data on landlords’ individual housing portfolios are rarely ever available. By all means, true numbers of centrists and decentrists are unobservable. As this paper and its companion paper (Dascher (2019)) argue, this is where we should exploit the information embodied in the city’s observable spatial structure. A careful look at the distribution of population across city rings will reveal bounds on the true numbers of centrists and decentrists. It is true that there are many ways in which landlords’ various properties throughout the city may combine (i.e. in which landlords are matched with their tenants). And yet – unobservable – centrists and decentrists must still be consistent with the – observable – distribution of housing across the city’s rings.

Let us briefly illustrate this simple principle. In a city with a large share of population or housing near the periphery, say, there cannot be many centrists. There simply are not many central properties that could induce their owners to behave centrist. Conversely, in a city with a large share of population or housing near the center, there cannot be many decentrists. There simply are not many peripheral properties that could motivate their owners to identify as decentrists. This simple and intuitive idea informs the entire paper. We are unable to compute true centrist or decentrist numbers. But we are able *to bound them*. We will derive lower (and upper) bounds on unobservable centrists and decentrists from the observable spatial distribution of the city’s population.

These bounds will have five properties. First, they turn out to be formulas (and simple ones). They are general functions of the spatial distribution of housing. Second, these

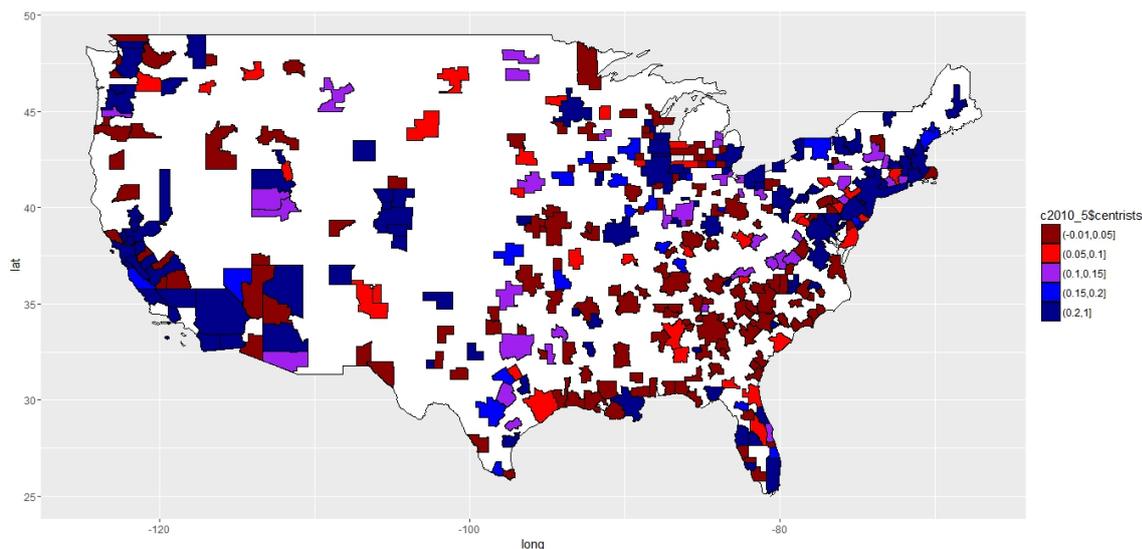


Figure 1: Minimum Centrist Share in U.S. metro areas in 2010.

bounds will never overestimate their corresponding true – centrist or decentrist – number. They are lower bounds. They provide us with conservative estimates of centrists and decentrists. Third, these bounds will never underestimate centrists’ or decentrists’ true numbers by “too much”. They are greatest lower bounds. They provide us with efficient (not unnecessarily small) estimates only. Fourth, these bounds mirror the city’s physical shape (Dascher (2019)). Finally, and fifth, these bounds must matter to controversial urban issues beyond carbon taxation, too (e.g. such as decentralization).

Bounds on centrists and decentrists matter in the following sense: Should the share of minimum centrists exceed one half of the electorate, then centrists (and their centrist agenda) can be inferred to prevail. Vice versa, if the share of minimum decentrists exceeds one half of the electorate, then decentrists’ cause wins. And even if neither share attains 0.5, our concepts need not be mute. A given interest group’s likelihood to prevail should still be increasing in its, and decreasing in the opposing group’s, minimum share. We compute minimum centrists and decentrists for all metro areas, and then employ sample estimates as extra regressors in various specifications of the Democratic vote share.

Fig. (1) displays the minimum share of centrists across metropolitan areas for 2010.³ The figure’s choropleth map (which best is viewed on screen) gives a flavor of why city

³Our empirical analysis makes use of a U.S. Census data set on “population by distance from the city center” for 2000 and 2010 and all U.S. metro areas, see Wilson (2012).

shapes may have contributed to why Donald Trump won. Metropolitan areas on the West coast as well as the northern part of the East coast exhibit consistently higher minimum centrist shares than interior areas do; and interior metro areas (sometimes derided as “flyover country”) often exhibit minimum centrist shares smaller than those in coastal metros. Fig. (2) also shows the minimum share of decentrists. While there are interior metros where decentrists are not very strong, decentrists are rarely strong along the two ocean coasts. Both maps coincide remarkably well with the well-known electoral map of the 2016 presidential election.

Republican and Democratic party platforms began to strongly differ only as recently as 2016. In 2016 the Republican party stated that “We oppose any carbon tax” (“Republican Platform 2016”, p. 20); while the Democratic party (“Democratic Party Platform”, p. 27) believed that “. . . America must be running entirely on clean energy by mid-century. We will take bold steps to slash carbon pollution.” But in the 2008 presidential election the two party platforms were almost indistinguishable. According to the *New York Times* (August 1st, 2016),

“[t]he divide between the two parties over the issue [of climate change in the fall of 2016, the author] is the widest it has been in the decades since it emerged as a public policy matter. That is all the more remarkable given that during the 2008 election, the Democratic and Republican positions on climate change were almost identical.”

This suggests how to go beyond the map’s (purely cross-sectional) variation. We may test our city shape explanation of why Donald Trump won from a “difference in differences” perspective, too, when comparing the 2008 and 2016 U.S. presidential elections. When party platforms were almost indistinguishable (back in 2008), minimum centrist and decentrist shares should have played no role. Yet when party platforms began to recognizably differ on global warming (in 2016), metro areas with large minimum centrist (decentrist) shares should have voted different from those with a zero such share. We will find that these expectations at least in part are borne out in our panel data. Ultimately, a metro area’s higher lower bound on *decentrists* – though not a smaller lower bound on

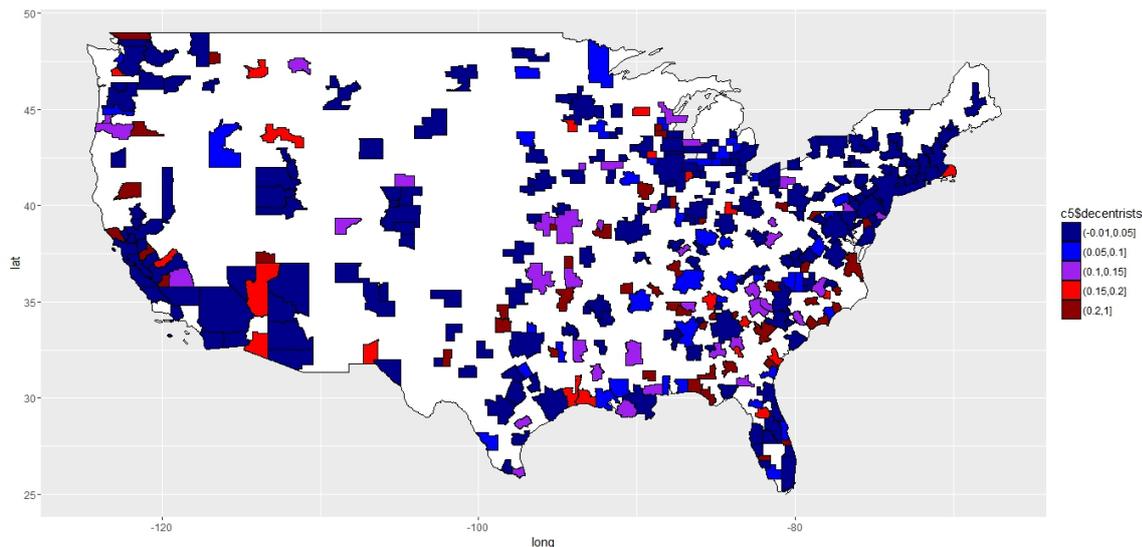


Figure 2: Minimum Decentrist Share in U.S. metro areas in 2010.

centrists – has reduced Clinton’s local take of the vote. And Trump’s declaring climate change irrelevant may – wittingly or not – have cost Clinton close to 280,000 votes.

The paper has seven sections. Sections 2 and 3 set out the basic framework for analyzing centrists and decentrists. Section 4 provides a solution to minimizing centrists for a given spatial city structure. Section 5 complements this with an analysis of minimizing decentrists. Section 6 applies our centrist/decentrist distinction to the 2008 and 2016 U.S. presidential elections. Sections 2 through 5 provide a rigorous theoretical foundation for the empirical analysis. They should not be seen as auxiliary only. They provide results that can stand on their own. Section 7 concludes.

2 Landlord-Tenant Matching

Monocentric City. A closed and monocentric city (as pioneered by Wheaton (1973), Pines/Sadka (1986) and Brueckner (1987)) juts \tilde{r} units of distance out from the CBD (with \tilde{r} determined shortly). Commuting costs for a resident living at distance r from the CBD are tr . Ricardian rent q follows $q(r) = t(\tilde{r} - r)$. The city’s overall population is s , and the urban wage is w . Residents consume one unit of housing. Housing is built by profit maximizing investors. One unit of capital k combined with one unit of land yields $h(k)$ units of housing, where $h' > 0$ and $h'' < 0$ (again, Brueckner (1987)).

Housing. If p is the price of capital, investors choose k so as to satisfy the $q(r)h_k(k) = p$ necessary for maximum profit. The optimal capital depends on rent q and price p , and so can be written as $k(t(\tilde{r} - r), p)$. Let $h(r)$ be shorthand for the building height obtained for this optimal capital choice. Then the city boundary \tilde{r} is determined by the condition that the housing market clear,

$$\int_0^{\tilde{r}} a(r)h(r) dr = s, \quad (1)$$

where $a(r)$ is land available in a ring of unit width r units of distance away from the CBD. Ratio $a(r)h(r)/s$, also written $f(r)$, indicates the share of the population commuting from within that ring to the CBD. Correspondingly, $F(r)$ denotes the share of households commuting r or less.⁴ Now divide the city into $i = 1, \dots, n$ concentric rings of equal width (n even) around the CBD, with n large enough to justify treating rent, building height, commuting times etc. as identical across ring i 's plots. Housing in ring i is app. $f(r_i)s$. We set $f(r_i)s = b_i$, to conform with the LP notation introduced shortly.⁵

Ownership. Traditional urban modeling has residents own urban housing jointly or treats landlords as absentee. Yet we want to avoid both the traditional “common ownership” or “absentee landlord” setup, lest we assume away the important centrist/decentrist-contest that is at the heart of this paper. We replace either assumption by dividing urban residents in two resident classes, resident landlords and tenants. Each landlord owns one unit of housing (an “apartment”) that he resides in himself as well as another apartment that he rents out. These two apartments, to be sure, do not need to locate in the same ring.⁶

Realistically, information on any given landlord’s two individual properties must be treated as private. And so we cannot say whether this landlord is a centrist or a decentrist. But (unknown) match matrix $X = (x_{ij})$ collects the frequencies with which the various possible matches between landlords and tenants occur, with row i (column j) indicating the landlord’s (tenant’s) location. Centrists (decentrists) are those landlords whose average property is closer to (further away from) the center than half the distance from the CBD

⁴We assume a is continuous in r . As h is (differentiable and hence) continuous in r , so is f .

⁵We will also refer to $f(r_i)$ or $f(r_i)s$ as the city’s *shape*, following terminology introduced in Arnott/Stiglitz (1981).

⁶Surely there are many other, often more complex, ways to introduce (i) resident landlords with their (ii) tenants into the city.

to the city boundary, $\tilde{r}/2$. Hence centrists are those for whom

$$(r_i + r_j)/2 < \tilde{r}/2 \tag{2}$$

or, equivalently, $i + j - 1 < n$.⁷ An analogous condition applies to decentrists.⁸

Now, a carbon tax raises the cost of commuting a mile, from t to $t' > t$. Consider a landlord who lives in a property in ring i (and hence has commuting cost tr_i), and rents out another to a tenant in ring j (and so receives $t(\tilde{r} - r_j)$). Her (net) income becomes $w + 2t(\tilde{r}/2 - (r_i + r_j)/2)$. Whether or not she will welcome the increase in t will clearly depend on whether or not her average property distance from the center $(r_i + r_j)/2$ is less than $\tilde{r}/2$, i.e. on *whether or not she is a centrist*. And so our centrist/decentrist distinction just coincides with the distinction between those who vote for Clinton and those who vote for Trump.

Matching. The previous inequality suggests that centrists (decentrists) are to be associated with entries of X that are located strictly above (below) its counter diagonal, i.e. the diagonal that stretches from X 's bottom left corner to its top right one. Moreover, being a centrist (or decentrist) does not depend on which apartment is the owner-occupied one, i or j . We may conveniently suggest that landlords always occupy the ring that is closer to the city center. And so with $i \leq j$, X becomes upper triangular. Now, to capture the overall number of households inhabiting ring i we need to sum over all of X 's entries in *both*, row i and column i . The resulting sum must equal ring i 's available stock of apartments, b_i . And so ring i 's housing constraint reads $\sum_{j=1}^n (x_{ij} + x_{ji}) = b_i$.

Linear Program. Summing over all centrist-related entries in X gives $\sum_{i=1}^{n-1} \sum_{j=1}^{n-i} x_{ij}$, the true, yet unknown, number of centrists, l^c . Contrast this with the smallest number of centrists conceivable, \underline{l}^c . That latter number bounds the true number of centrists l^c from below. To identify \underline{l}^c , we minimize the number of centrists given ring housing constraints and the non-negativity requirements $x_{ij} \geq 0$. This translates into the following linear

⁷This follows from assuming that residents in ring i commute distance $(i - 0.5)\tilde{r}/n$.

⁸Note that even as decentrists have properties closer to the city extremes, "extremists" probably is not a better term. – Jacobs (1961) and Breheny (2007) also use the term "decentrists", though with a very different meaning. For Jacobs, decentrists are those early 20th century urban and regional planners such as Lewis Mumford, Clarence Stein, Henry Wright and Catherine Bauer, who advocated "thinning out large cities" by dispersing their "enterprises and populations into smaller, separated cities or, better yet, towns" (p. 19).

program

$$\begin{aligned} \min_{x_{ij}} \quad & \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} x_{ij} \quad \text{s.t.} \quad \sum_{j=1}^n (x_{ij} + x_{ji}) = b_i \quad (i = 1, \dots, n) \\ & x_{ij} \geq 0 \quad (i, j = 1, \dots, n), \end{aligned} \quad (3)$$

analysis of which is the focus of the next two sections.

3 The Minimum Share of Centrists, in Two Specific Cities

We run two eight-ring city examples on how to solve the linear program (3) next. These are examples to offer some intuition on how a feasible, and even optimal, solution to linear program (3) plays out. But in fact they are much more than just examples. They motivate a trial solution that later will generalize to any given city.

Example City 1. Our first city has “city shape” $b = (38, 36, 30, 10, 12, 8, 4, 2)$. To this city, matrix X_1 in (4), in highlighting eight non-zero entries, suggests one basic feasible solution.⁹ We briefly illustrate feasibility. Adding up all entries in row 1 and column 1, for instance, gives $20 + 18 = 38$ or b_1 , while adding up all entries in row 7 (consisting of zeros only) and column 7 gives just $0 + 4$ or b_7 . Our feasible solution here displays one feature that we might expect of an optimal solution, notably that (4) assigns the maximum possible weight to entries on the counterdiagonal (in red on screen). This forces centrists’ numbers down as best as we can. We get $x_{18} = \min\{b_1, b_8\} = 2$. Similarly, $x_{27} = 4$, $x_{36} = 8$ and $x_{45} = 10$.

Put differently, whenever possible we allocate a peripheral apartment in some given outer ring j , $5 \leq j \leq 8$, to a proprietor who owns her other, second apartment in corresponding inner ring $9 - j$. This must be a necessary property of a centrist-minimizing allocation. (Suppose that X_1 violated this property, i.e. suppose $x_{18} = 1 < 2 = \min\{38, 2\}$. Since there are no apartments, anywhere, capable of successfully turning a landlord in ring 1 – someone who would otherwise be a centrist – into a decentrist, an opportunity to reduce centrists would have been irrevocably wasted.) At the same time, of course, not all

⁹Here, as well as in all other match matrices below, entries with no explicit number attached equal zero.

apartments in a given peripheral ring j may be assignable to a landlord in corresponding ring $n - j + 1$. In ring $j = 5$, for example, only 10 out of 12 apartments are.

There are $(b_1 - b_8) = 36$ apartments in ring 1 still waiting to be allocated, as are $(b_2 - b_7) = 32$ apartments in ring 2 and $(b_3 - b_6) = 22$ apartments in ring 3. We apportion these remainders to landlords owning both their properties within the same ring. Since any match on the main diagonal accounts for two apartments, we set $x_{11} = (b_1 - b_8)/2 = 18$, $x_{22} = (b_2 - b_7)/2 = 16$ and $x_{33} = (b_3 - b_6)/2 = 11$ (all blue on screen). Note that $x_{44} = 0$, given that $x_{45} = 10$ already and that row 4 and column 4 must add up to $b_4 = 10$. It remains to balance housing in ring 5, by setting x_{55} to 1 (brown on screen). – Now, invoking the simplex algorithm would reveal that the solution set out in (4) above not just is feasible but also: optimal.¹⁰ Instead of going through these details here, we offer a systematic treatment below (in the following section).

$$X_1 = \begin{pmatrix} \mathbf{18} & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{2} \\ & \mathbf{16} & 0 & 0 & 0 & 0 & \mathbf{4} & 0 \\ & & \mathbf{11} & 0 & 0 & \mathbf{8} & 0 & 0 \\ & & & 0 & \mathbf{10} & 0 & 0 & 0 \\ & & & & \mathbf{1} & 0 & 0 & 0 \\ & 0 & & & & 0 & 0 & 0 \\ & & & & & & 0 & 0 \\ & & & & & & & 0 \end{pmatrix} \quad (4)$$

We conclude that the trial number of centrists suggested by (4) also is the minimum number of centrists given the specific city shape b in hand. Adding up these centrists is simple enough. We merely need to collect the few non-zero entries found above the counterdiagonal. These are conveniently located on the upper half of the main diagonal (blue on screen). This gives $\sum_{i=1}^3 (b_1 - b_{9-i})/2$ or 45 minimum centrists. Minimum centrists' share in city population becomes $45/140$. Computing minimum centrists provides valuable information here. It is not possible for the true number of centrists to fall short of 45. But it is quite possible – if not utterly likely – for the true number of centrists to surpass 45. Of course, the latter likely occurs should true matches deviate from one of the optimal solutions.

¹⁰It is not, however, a unique optimal solution. For example, letting any landlord trade apartments with her or his tenant would generate another optimal solution.

Example City 2. Our second example city exhibits housing stocks described by “city shape” $b = (38, 14, 30, 10, 12, 8, 26, 2)$. We take an important step towards generalization by introducing the concept of ring difference δ_i now, where $\delta_i = b_i - b_{n+1-i}$ is the number of apartments in “leading” ring i minus that in “lagging” or “antagonist” ring $n + 1 - i$. It is defined for $1 \leq i \leq 4$. In our second example city, δ_i is positive for i equal to 1 or 3 (since there we have a “surplus”) and it is negative if i equals 2 or 4 (because then there is a “deficit”). Contrast this with our first example city, where all first three ring differences are positive.

True to our strategy of emphasizing the counterdiagonal, feasible solution X_2 in (5) assigns as many apartments as possible in lagging rings to owners in corresponding leading rings. And because we have a surplus in rings 1 and 3, for these rings this works just fine. All apartments in rings 8 and 6 can be assigned to landlords living in rings 1 and 3, respectively. And while this works less well for apartments in lagging rings 5 and 7, remaining apartments are not always lost on us. Ring 2’s deficit (of $-(b_2 - b_7) = 12$), for instance, we may “save up for”, or “post to”, the next best successive ring boasting a surplus. In our example, this is ring 3 (where $b_3 - b_6 = 22$). The 12 apartments reflecting ring 2’s deficit can valuably be employed to offset the better part of ring 3’s surplus.

And so we set entry x_{37} in X_2 to $b_7 - b_2$, or 12 (green). Intuitively, the 12 ring 7-apartments not assignable to ring 2-landlords now are assigned to landlords in ring 3, to at least turn those off centrism. Note that the same is not possible to do with the ring deficit arising in ring 4. There simply are no later rings. – Everything else parallels our discussion of the first example. We balance the first three rings’ housing constraints by setting $x_{11} = (b_1 - b_8)/2 = 18$, $x_{22} = 0$ and $x_{33} = (b_3 - (b_6 + (b_7 - b_2)))/2 = 5$. Again, moreover, the basic feasible solution, set out in (5), also is the optimal one. Minimum centrists are found to sum to 23, if only to see their share in the overall total attain a mere 23/140.

$$X_2 = \begin{pmatrix} \mathbf{18} & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{2} \\ & 0 & 0 & 0 & 0 & 0 & \mathbf{14} & 0 \\ & & \mathbf{5} & 0 & 0 & \mathbf{8} & \mathbf{12} & 0 \\ & & & 0 & \mathbf{10} & 0 & 0 & 0 \\ & & & & \mathbf{1} & 0 & 0 & 0 \\ & 0 & & & & 0 & 0 & 0 \\ & & & & & & 0 & 0 \\ & & & & & & & 0 \end{pmatrix} \quad (5)$$

Review. What can be learned from these two examples? We have seen that in both cities minimum centrists may be written as the cumulative sum of the first three ring differences, $\sum_{i=1}^3 (b_1 - b_{9-i})/2$. This is true even as δ_2 is positive in the first example city while negative in the second. But why does it make sense to include δ_2 in either example? The answer is this: On the one hand, including $\delta_2/2$ in the cumulative sum *when positive* acknowledges the fact that $(b_2 - b_7)/2$ landlords in ring 2 cannot be turned away from centrism. On the other hand, including $\delta_2/2$ in the cumulative sum *when negative* acknowledges the fact that $(b_7 - b_2)/2$ landlords in ring 3 can (be turned off centrism).

We must also wonder about why $\sum_{i=1}^3 (b_1 - b_{9-i})/2$ excludes $\delta_4/2$. In particular, why is negative $\delta_4/2$ not included in the second city's cumulative sum when negative $\delta_2/2$ is? Following our previous intuition, there is no need to “save” ring 5 apartments for later because there are no later surpluses to “swipe away”. The only remaining ring that could possibly feature a centrist landlord is ring 4. Yet here δ_4 's negative sign indicates that the planner can already afford each landlord in ring 4 a ring 5-apartment that successfully counters that landlord's initial impulse to “go centrist”. And with no further centrists to collect in the fourth ring, our cumulative sum should: stop short of it.

Tentative Ideas. Two ideas emerge from this: (i) Minimum centrists can be represented as a cumulative sum of successive ring differences. (ii) Successive ring differences should enter that cumulative sum if they are positive. And they should even enter the cumulative sum if they are negative, as long as they can help “wipe out” subsequent positive ones. Negative ring differences should be included if and only if they are followed by positive ones at least equal in size. I.e., the cumulative sum should include successive ring differences as long as this helps raise the cumulative sum. Equivalently, to *minimize* centrists we must *maximize* the cumulative sum of ring differences. We will return to this equivalence in a

moment, when generalizing our examples (in the next section).

4 The Minimum Share of Centrists, Anywhere

Primal vs. Dual Program. We allow for any $n \times 1$ vector of ring housing stocks $b = (b_1, \dots, b_n)$ now, except for ruling out any b_i to equal zero. We then put the corresponding linear program (3) into standard form. We first stack all n columns of X into one long $(n^2 \times 1)$ vector x . This gives $x' = (x_{11}, \dots, x_{1n}, \dots, x_{n1}, \dots, x_{nn})$. To address the objective function in (3) in matrix notation, let c_i equal an $n \times 1$ vector consisting of ones only except for the last i entries, which are zero instead. For example, c_3 is a list of $n - 3$ ones followed by three zeros, i.e. $c'_3 = (1, \dots, 1, 0, 0, 0)$. Define an $n^2 \times 1$ list of weights c by setting $c' = (c'_1, \dots, c'_n)$. Then our objective function $\sum_{i=1}^{n-1} \sum_{j=1}^{n-i} x_{ij}$ can be cast as the product $c'x$.

Next, let τ_i denote an $n \times 1$ vector featuring 2 in its i -th row and 1 in all other rows. For example, $\tau'_2 = (1, 2, 1, \dots, 1)$. Moreover, let J_i denote what becomes of the $n \times n$ identity matrix once row i has been replaced with τ'_i . Then the coefficient matrix A is $A = (J_1, \dots, J_n)$; it is of dimensions $n \times n^2$. The tableau in Table (1) illustrates A in its bottom left part. This table also indicates our specific vector of objective function weights c (in its first row) as well as the vector of ring housing stocks b (last column).¹¹

1	1	1	1	...	1	0	...	1	0	0	0	...	0	0	
2	1	1	1	...	1	1	...	1	0	0	0	...	0	0	b_1
0	1	0	0	...	0	0	...	0	1	0	0	...	0	0	b_2
0	0	1	0	...	0	0	...	0	0	1	0	...	0	0	b_3
			\vdots									\vdots			\vdots
0	0	0	0	...	1	0	...	0	0	0	0	...	1	0	b_{n-1}
0	0	0	0	...	0	1	...	1	1	1	1	...	1	2	b_n

Table 1: Matrix A , objective function weights c and housing stocks b

With this extra notation in hand, linear program (3) may equivalently be stated as $\min_x c'x$ subject to $Ax = b$ and $x \geq 0$, where the equality constraints may also be read off Table (1)'s rows. This program's dual is $\max_y y'b$ such that $y'A \leq c'$, where y is the dual's $(n \times 1)$

¹¹As inspection of A makes clear, ours is not a transportation problem (e.g., as defined in Hadley (1963)).

vector of unknowns, $y' = (y_1, \dots, y_n)$. Table (1) also indicates the dual's constraints; these can be read off its columns. For instance, the constraint complementary to x_{11} being strictly positive simply is $2y_1 \leq c_{11} = 1$ (see first column in Table (1)).

Rather than immediately analyze the general case, we focus on a seemingly special case first. This case allows us to best connect with the principles that emerge from our discussion of the two example cities (section 3). To address this special case, let us introduce the partial cumulative sum $\Delta(i) = \sum_{j=1}^i \delta_j/2$. This sum cumulates successive ring differences δ_j up to ring i , where of course $i \leq n/2$. And let index i^* be the index that maximizes this cumulative sum, i.e.

$$i^* = \arg \max_i \sum_{j=1}^i (b_j - b_{n+1-j})/2. \quad (6)$$

Our point of departure on the way to the fully general solution is a city for which (i) $\Delta(i^*) > 0$ and (ii) all ring differences δ_i are negative except when $i = i^*$, when $\delta_{i^*} > 0$.

Trial Solution. We set out a basic feasible solution to the primal problem next. Table (2) shows X in tabular form and may be a useful reference as we go along. Again, entries of X never addressed are zero. Moreover, also note the formal resemblance between Table (2) on the one hand and matrices X_1 and X_2 on the other. Now, we begin by considering the elements on the counterdiagonal of match matrix X . Here we set (red on screen)

$$x_{i,n+1-i} = \min \{b_i, b_{n+1-i}\} \quad (i = 1, \dots, n/2). \quad (7)$$

Given our sign assumptions regarding the δ_i , this entails setting all entries $x_{1,n}$ “up” to $x_{i^*-1,n+2-i^*}$, and again from $x_{i^*+1,n-i^*}$ to $x_{n/2,n/2+1}$, equal to the leading ring's stock, b_i . Only $x_{i^*,n+1-i^*}$ becomes the lagging ring's stock, b_{n+1-i^*} . Note how this assignment makes as many owners of property in leading rings (voters who otherwise likely are centrists) as possible disavow centrism.

Moreover, set (green on screen)

$$x_{i^*,n+1-i} = (b_{n+1-i} - b_i) \quad (i = 1, \dots, i^* - 1). \quad (8)$$

Note that the expressions on the r.h.s. represent ring deficits. Deficits originating in rings prior to i^* are posted to leading ring i^* , as the earliest next ring offering up an excess.

Ro./Co.	1			i^*			$n/2 + 1$		$n - i^*$				n
1	0												$x_{1,n}$
		\ddots											0
			0									$x_{i^*-1,n+2-i^*}$	0
i^*				x_{i^*,i^*}						$x_{i^*,n+1-i^*}$	$x_{i^*,n+2-i^*}$	\dots	$x_{i^*,n}$
					0				$x_{i^*+1,n-i^*}$				
						\ddots			\ddots				
$n/2$						0	$x_{n/2,n/2+1}$						
$n/2 + 1$							$x_{n/2+1,n/2+1}$						
										\ddots			
$n - i^*$									$x_{n-i^*,n-i^*}$				
										0			
											0		
												\ddots	
n													0

Table 2: Non-Zero Elements in Basic Feasible Solution

“Apartment savings” originating in rings up to i^* then are matched up with apartments in ring i^* . This generalizes how we proceeded earlier when setting x_{37} equal to 12 in example city 2.

Next, let (blue on screen)

$$x_{i^*i^*} = \left(b_{i^*} - (b_{n+1-i^*} + \sum_{k=1}^{i^*-1} (b_{n+1-k} - b_k)) \right) / 2, \quad (9)$$

or $\Delta(i^*)$. At first sight nothing seems to preclude $x_{i^*i^*}$ from being strictly negative, in contradiction to primal variables’ non-negativity constraints. However, recall that i^* maximizes the cumulative sum of ring differences. And so $\sum_{j=1}^{i^*} \delta_j / 2 \geq 0$, i.e. a non-negative number. And note that this latter number just coincides with the r.h.s. of (9). Put yet differently, ring excess δ_{i^*} is more than sufficient to offset the ring deficits δ_k associated with, and inherited from, all the rings prior to i^* . And so $x_{i^*i^*}$ really is non-negative.

At last we set (brown on screen)

$$x_{n+1-i,n+1-i} = (b_{n+1-i} - b_i) / 2 \quad (i = i^* + 1, \dots, n/2). \quad (10)$$

Ring deficits originating in rings following i^* are relegated to main diagonal elements below the counterdiagonal, to the desirable effect of contributing nothing to the number

of centrists. Note how equations (7), (8), (9) and (10) set out a feasible solution of the primal.

Complementary Slackness. We invoke complementary slackness between the primal and the dual. For $i = 1, \dots, n/2$, entries on the counterdiagonal $x_{i,n+1-i}$ are strictly positive (see (7)), as is the main diagonal element $x_{i^*i^*}$ (see (9)). By complementary slackness, the corresponding constraints of the dual – read off the corresponding columns of Table (1) – must be met with equality, and so

$$y_i = -y_{n+1-i} \quad (i = 1, \dots, n/2) \quad \text{and} \quad y_{i^*} = 1/2. \quad (11)$$

These equations specify the weights on ring housing stocks b_i in the dual's objective.

For $i = 1, \dots, i^* - 1$, entries $x_{i^*,n+1-i}$ are strictly positive, too (see (8)). Again, by complementary slackness, corresponding constraint inequalities in the dual become binding. And so, according to Table (1), $y_{i^*} = -y_{n+1-i}$. Combining this with $y_{n+1-i} = -y_i$ and the fact that $y_{i^*} = 1/2$ (see (11)) gives the first set of equations in (12). At last we make use of equations (10). For $i = i^* + 1, \dots, n/2$, constraint (in)equalities translate into $y_i = 0$. Joint with the first set of equations in (11), this in turn implies the second set of equations in (12):

$$y_i = 1/2 \quad (i = 1, \dots, i^* - 1) \quad \text{and} \quad y_i = 0 \quad (i = i^* + 1, \dots, n - i^*). \quad (12)$$

Table (3) collects the full solution to equations (11) and (12), denoted \bar{y} and easily shown to be feasible, too.

i	1	...	i^*	$i^* + 1$...	$n - i^*$	$n - i^* + 1$...	n
\bar{y}_i	1/2	1/2	1/2	0	0	0	-1/2	-1/2	-1/2

Table 3: The dual's optimal solution

Basic Feasible Solution is Optimal. Let us now put together feasibility and complementary slackness, using standard reasoning in linear programming (Chvatal (1980), Hadley (1963), Luenberger/Ye (2016)). First, feasibility of \bar{x} and \bar{y} implies $b = A\bar{x}$ and $\bar{y}'A \leq c'$, respectively, and hence $\bar{y}'b = \bar{y}'(A\bar{x}) = (\bar{y}'A)\bar{x} \leq c'\bar{x}$. Second, complementary slackness implies $(\bar{y}'A - c')\bar{x} = 0$ or $(\bar{y}'A)\bar{x} = c'\bar{x}$. And so we may conclude that $\bar{y}'b = c'\bar{x}$.

This in turn implies that $c'\bar{x}$ equals minimum centrists, and hence that \bar{x} solves (3). Of course, if \bar{x} is optimal, then so is \bar{y} , justifying Table (3)'s title.

We compute the objective function values for primal and dual, providing a check on optimality of \bar{x} and \bar{y} as well as, of course, the desired minimum number of centrists itself. On the one hand, summing over all entries above the counter diagonal the objective function value in the primal gives $x_{i^*i^*}$ as on the r.h.s. of equation (9). But then:

$$\boxed{\underline{l}^c = \Delta(i^*) = \max_i \sum_{j=1}^i (b_j - b_{n+1-j})/2.} \quad (13)$$

On the other hand, computing the sum of ring stocks using the optimal weights in (11) and (12) yields the very same formula, i.e. $\sum_{j=1}^{i^*} (b_j - b_{n+1-j})/2$. *This formula represents the optimal value of both primal and dual. And so it also represents the minimum conceivable number of centrists.* We briefly pause to appreciate its generality: the greatest cumulative ring difference gives a universal closed form solution for minimum centrists. It provides an observer of an arbitrary given city with a prediction of centrists' minimum.

Our proof is for a city whose ring differences, with the exception of δ_{i^*} , are all negative (also see the first two rows in Table (6) in the Appendix). The Appendix shows how the proof quickly generalizes. Subsections 9.2 through 9.4 show that our results in essence remain unchanged as some, or even all, ring differences exhibit an arbitrary sign. Formula (13) remains valid throughout. This is quite straightforward since also accounting for positive ring differences (Appendix) is simpler than accounting for negative ones (this section): witness solution X_1 as opposed to X_2 (in section 3). Now, translating minimum centrist numbers in formula (13) into minimum centrists' share in all landlords, by dividing $\Delta(i^*)$ by $s/2$, gives the following variant of this result:

Proposition 1: (Greatest Cumulative Ring Difference and Centrists)

Centrists' minimum conceivable share of the landlord population, $\underline{\lambda}^c$, is given by the greatest cumulative ring difference, $\underline{\lambda}^c = \max_i \sum_{j=1}^i (b_j/s - b_{n+1-j}/s)$.

Proposition 1 extends Dascher (2019), where $\underline{\lambda}^c$ is introduced a mere *lower bound* to centrists' true number. We here add that $\underline{\lambda}^c$ even is the *greatest* lower bound (because it

is the minimum). This makes us more confident to work with $\underline{\lambda}^c$ empirically (section 6).

5 Centrists vs. Decentrists

Minimum Decentrists. We bring in decentrists now. Intuitively, where before we have used b_{n+1-i} to “swipe away” or “swamp” potential centrists in i (as best as we could), conversely we now use b_i to “swamp” decentrists in $n + 1 - i$ (as best as we can). Applying a proof similar to that in section 4 (omitted for brevity), we find that minimum decentrists correspond to: minus the least cumulative ring difference. That is, if $i^{**} = \arg \max_i \sum_{j=1}^i (-(b_j - b_{n+1-j}))/2$, then minimum decentrists \underline{l}^d are equal to

$$\boxed{\underline{l}^d = -\Delta(i^{**}) = -\min_i \sum_{j=1}^i (b_j - b_{n+1-j})/2.} \quad (14)$$

Translating this number into a share gives

Proposition 2: (Least Cumulative Ring Difference and Decentrists)

Decentrists’ minimum conceivable share of the landlord population, $\underline{\lambda}^d$, is given by minus the least cumulative ring difference, $\underline{\lambda}^d = -\min_i \sum_{j=1}^i (b_j/s - b_{n+1-j}/s)$.

Upper Bounds. We quickly turn lower bounds in Propositions 1 and 2 into corresponding upper bounds. Subtracting centrists from overall landlord population $s/2$ gives the sum of decentrists and indifferent landlords. This in turn is the sum of all elements of X strictly below or on the counter diagonal. The following linear program looks for the maximum sum of decentrists/indifferents:

$$\max_{x_{ij}} \left(s/2 - \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} x_{ij} \right) \quad \text{s.t.} \quad \sum_{j=1}^n (x_{ij} + x_{ji}) = b_i \quad (i = 1, \dots, n) \\ x_{ij} \geq 0 \quad (i, j = 1, \dots, n). \quad (15)$$

Comparing linear programs, clearly the maximizer to (15) coincides with the minimizer to (3). But this implies that $s/2 - \underline{l}^c$ is the maximum conceivable number of decentrists/indifferents. And so $s/2 - \underline{l}^c$ is an upper bound to decentrists only (Proposition, Part (ii)). A similar argument suggests that $s/2 - \underline{l}^d$, where \underline{l}^d is the minimum number of decentrists, is an upper bound to centrists (Proposition 3, Part (i)).

Proposition 3: (Upper Bounds on Centrists and Decentrists)

(i) λ^c is bounded from above by $1 - \underline{\lambda}^d$. (ii) λ^d is bounded from above by $1 - \underline{\lambda}^c$.

We adopt a random voter turnout perspective, as in Brueckner/Glazer (2008). Let the (carbon tax) policy's probability of electoral success π be an increasing (decreasing) function of the centrist share λ^c (decentrist share λ^d). More specifically, $\pi = G(\lambda^c, \lambda^d)$ with partial derivatives $G_1 > 0$ and $G_2 < 0$. This we combine with the additional (plausible yet by no means forceful) assumption that λ^c is increasing in $\underline{\lambda}^c$, and that λ^d is increasing in $\underline{\lambda}^d$. This implies $\pi = G(\underline{\lambda}^c, \underline{\lambda}^d)$, again with partial derivatives $G_1 > 0$ and $G_2 < 0$. We then consider the linearized version of

Empirical Model of Voting on a Carbon Tax:

The probability of the carbon tax proposal's electoral success π is increasing (decreasing) in centrists' (decentrists') minimum share in the electorate $\underline{\lambda}^c$ ($\underline{\lambda}^d$).

It is true that centrists and decentrists constitute only half of the electorate. Tenants, as the other half, see their real income dwindle as the urban cost of living $t\tilde{r}$ rises. At the same time, there are extra benefits to taking t to t' . Taxing urban commutes helps fight global warming (as the city structure defined in (1) gradually adapts over time) or at the very least provides a psychological benefit. Taxing urban commutes also generates tax revenue, part of which might be redistributed to the electorate.

6 US Metropolitan Areas and Presidential Election 2016

During the campaign for the 2016 U.S. presidential election, Donald Trump certainly was the candidate more prone to side with those who reject a carbon tax. At the same time, Hillary Clinton was the one more likely to raise the cost of carbon consumption.¹² Clinton appears to have been centrists' favorite candidate, Trump must have been decentrists' favorite. We compute Clinton's share in all votes cast in support of either Clinton or Trump using Dave Leip's data set on the 2016 U.S. presidential election.¹³ Table (10) (in

¹²Following the *New York Times* during the campaign suggests as much. For example, see "Climate Change Divide Bursts to the Forefront in Presidential Campaign", *New York Times* August 1st, 2016.

¹³This dataset provides votes at the county level. We aggregate these data for metropolitan areas.

the Appendix) illustrates Clinton’s share in the overall vote.¹⁴ Consulting the distribution of the Democratic vote share for the election in 2016 shows that in half of all metro areas Clinton captured no more than 43% of votes cast for either herself or Trump. Of course, the metro areas in which Clinton did not do well are also those that are populated less, and so this observation is consistent with Clinton winning the overall popular vote.

The 2016 Presidential Election. Data on population, as well as population-weighted densities, by distance (in miles) from the city center are provided by the U.S. Census Bureau (Wilson (2012)) for all U.S. metropolitan areas and years 2000 and 2010. We first exploit these data for year 2010, i.e. a year predating the 2016 election by six years. These data are ideally suited to the purposes of this paper.¹⁵ We define the city boundary as the index of the last ring exhibiting a population weighted density of greater than 500.¹⁶ Given this boundary we compute bounds $\underline{\lambda}^c$ and $\underline{\lambda}^d$ for every metro area in 2010 (as outlined in Propositions 1 and 2). Table (10) sketches our bounds’ distributions. In more than three fourths of all metro areas does the minimum share of centrists fail to cross the 0.5-threshold. Nor can decentrists claim to be decisive often. Three fourths of metro areas exhibit a minimum decentrist share of 12% or less.

Figure (3) provides some extra illustration, in mapping the distance of four large metro areas’ rings from the metro area center into rings’ housing shares b/s . This figure also shows corresponding lower and upper bounds on centrist shares, i.e. $\underline{\lambda}^c$ and $1 - \underline{\lambda}^d$. Not all of these barcharts conform with intuition. Phoenix may be sprawling, and Boston may be compact at the center, yet corresponding metro areas exhibit comparable minimum centrist shares, of 45% and 44%, respectively. As the barcharts show, very different city shapes can conceal, or give rise to, very similar minimum centrist shares. We also see that commuting densities for Houston and Detroit are much less amenable to centrism, and for Detroit we may even state that at best 99% of all landlords could be centrist. We also note

¹⁴This Table also has descriptive statistics on all other variables discussed below, for both of the years (2008 and 2016) of the full data set analyzed later.

¹⁵Model and data set properties certainly do not even agree roughly. Metro areas clearly are far from monocentric (Glaeser/Kahn (2004)). Moreover, neither will tenants constitute exactly one half of the population, nor will housing ownership be evenly distributed across landlords. (Yet again, considering everyone a homevoter (as is sometimes done) seems even less appropriate.)

¹⁶However, the first ring was included even if that ring’s weighted density fell below 501. Also note that for some 30 metro areas, one or more variables are missing, and so these observations were dropped.

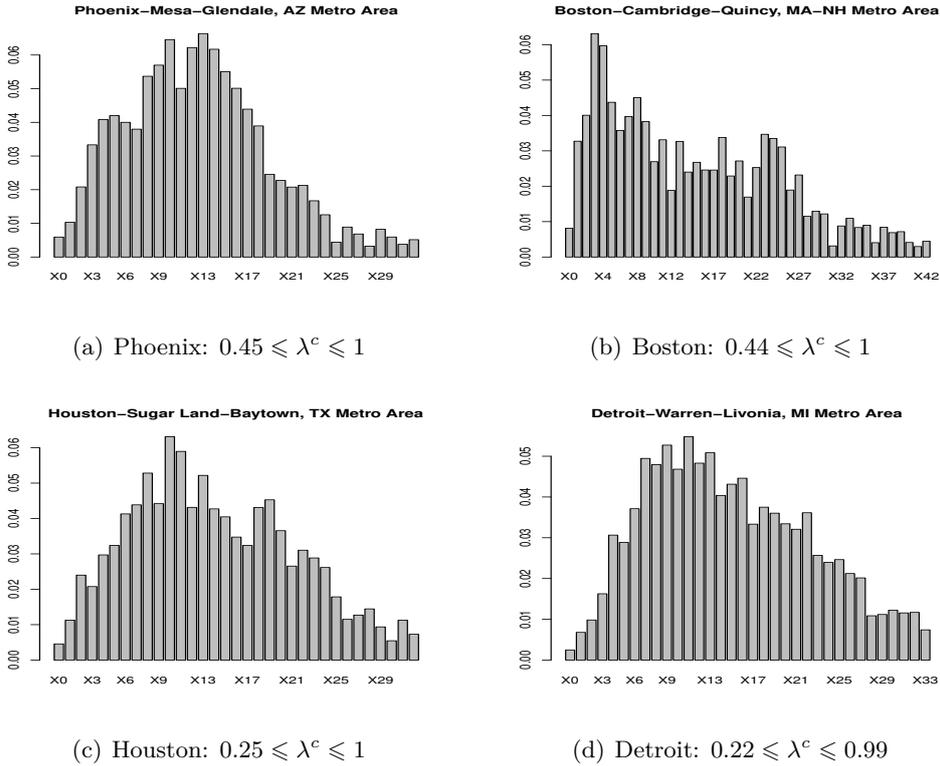


Figure 3: Population by Distance from Center and Bounding Centrists

that metropolitan areas whose commuting densities are more skewed to the right tend to house more centrists.

With the simplest possible specification (the linear probability model) in mind, we next regress Clinton’s share of votes on both our bounds (computed for 2010) as well as (i) the share of whites in metro population (in 2016), (ii) average income (2016), (iii) the share of those who completed a bachelor’s degree in metro population (2016, (iv) metro area size population (2016) and even (v) metro area average population density (available for 2010).¹⁷ We use the share of whites to capture Trump’s resonance among white voters, average income to address a lack of taste for redistribution, we use bachelor degrees to proxy for voters’ resilience against populist slogans as well as metro size and urban density to capture minorities’ greater attraction to larger, and denser, urban areas.

¹⁷These five covariates closely follow the early discussion in Florida (2016). Data on average income by metro area are from 2016 and were retrieved from the Bureau of Economic Analysis. Data on bachelor degrees in metro areas relate to those who are 25 years or older and are from the American Community Survey’s “Educational Attainment Package” for 2016. Data on metro size, metro average density and whites in total population are provided by U.S. Census Bureau.

	(1)	(2)	(3)	(4)	(5)	(6)
ShareMinimumCentrists	0.174*** (0.060)	0.156*** (0.058)	0.137** (0.058)	0.143*** (0.051)	0.142*** (0.051)	0.011 (0.048)
ShareMinimumDecentrists	-0.200*** (0.073)	-0.230*** (0.070)	-0.228*** (0.069)	-0.146** (0.061)	-0.145** (0.056)	-0.098* (0.056)
ShareWhite	-	-0.326*** (0.057)	-0.315*** (0.057)	-0.378*** (0.050)	-0.378*** (0.053)	-0.351*** (0.048)
MeanIncome	-	-	0.014*** (0.005)	-0.011** (0.005)	-0.011** (0.005)	-0.018*** (0.005)
ShareCollege	-	-	-	1.505*** (0.153)	1.501*** (0.156)	1.299*** (0.143)
SizeMetro	-	-	-	-	0.001 (0.012)	-0.032*** (0.012)
DensityMetro	-	-	-	-	-	0.006*** (0.006)
\bar{R}^2	0.09	0.17	0.18	0.37	0.37	0.48
N	338	338	338	338	338	338

Table 4: OLS Regressions (Standard errors in parentheses)

OLS regressions (1) through (6) in Table (5) explore the role of minimum centrists and decentrists for Clinton’s tally. Column (1) shows that a greater minimum share of centrists increases Clinton’s share of votes, while a greater minimum share of decentrists decreases it, and both estimates are significant (certainly at the 10 percent level; standard errors are found in parentheses). For instance, observing minimum decentrists to go up by ten percentage points permits us to roughly predict a 2.3 percentage point *drop* in Clinton’s vote share. At the same time it is true that our bounds shed light on a tiny fraction of the overall variation of Clinton’s performance across metropolitan areas only. We explain more of this variation when adding the share of whites (see column (2)). This also leaves our coefficient estimates for $\underline{\lambda}^c$ and $\underline{\lambda}^d$ largely unchanged. The same is true after also including average income as an additional regressor, in column (3).

Column (4) also controls for the share of bachelor degree recipients among those who are 25 years or older, and column (5) adds metro area population. Coefficient estimates for $\underline{\lambda}^c$ and $\underline{\lambda}^d$ continue to be significant here, too. Finally, also including average metropolitan population-weighted density in column (6) at last has the estimate for $\underline{\lambda}^c$ turn insignificant, while that for $\underline{\lambda}^d$ remains significant (at least at the 10% level).¹⁸ One possible interpreta-

¹⁸Coefficients for our controls are not our focus, but we nonetheless may briefly note that our estimates of almost all of them (the exception being metro size) conform with what we would expect: A metro area tends to vote more strongly for Clinton if it is (i) less white, (ii) poorer (less rich), (iii) more educated and (iv) denser.

tion of these results is that, incidentally, Donald Trump had nothing to lose from denying climate change. On the one hand, denying climate change did not cost Trump any votes in those metro areas where $\underline{\lambda}^c$ was larger (that variable's coefficient must be zero). On the other hand, denying climate change did give Trump additional votes in those metro areas where $\underline{\lambda}^d$ was larger (this variable's coefficient is negative).

Comparing 2008 and 2016 Elections. Metro areas certainly differ in more respects from one another than just in those accounted for in Table (4). Such unobservable effects may correlate with our observed explanatory variables.¹⁹ To address at least the possibility of time-constant unobservables we extend our sample to year 2008.²⁰ Regression (7) (in Table (5)) reestimates our full specification adding metro area fixed effects. Coefficient estimates for minimum centrist and decentrist shares have the expected signs, yet only that for $\underline{\lambda}^c$ is significant. We next add a time fixed effect in regression (8) to account for election idiosyncracies. This captures then FBI director James Comey's decision to reinvestigate Hillary Clinton's emails, the emergence of the "Access Hollywood Tape" or the "Christopher Steele file", etc., all just weeks ahead of the election. This latter effect is negative, and highly significant. Yet coefficients on our lower bounds turn insignificant.

As argued in the introduction, specification (8) may suffer from lumping two year-specific slopes together when really they should be allowed to differ. As long as party positions with respect to global warming were almost the same (which appears true for 2008), minimum centrist and decentrist shares should have played no role. Yet as soon as party positions started to differ (as was true in 2016), metro areas with large minimum centrist (decentrist) shares should have begun to vote different from those with a zero such share. To account for this issue, we include two interaction terms in column (9), interacting both our lower bounds with the 2016 year dummy. Consulting Table (5), we find that coefficient

¹⁹This is why we refrain from modeling random effects below.

²⁰U.S. Census Bureau data on "population by distance from the city center" are available for just one more year, 2000. So we can only include one other election year. We need to choose between 2008 and 2012, both of which could reasonably be matched with U.S. Census data from 2000. We include the data corresponding to the U.S. presidential election 2008, when the Republican and Democratic parties' platforms on climate change resembled each other more. We add 2008 data on the share of whites, the share of those with bachelor, mean income and population using exactly the same sources as for 2016. Data on weighted population density as well as the data underlying minimum centrist and decentrist shares use the U.S. Census Bureau data on 2000. Data on Democratic and Republican votes in the 2008 election once more come from Dave Leip's website.

	(7)	(8)	(9)	(10)
ShareMinimumCentrists	0.084*** (0.019)	0.012 (0.018)	0.022 (0.020)	0.004 (0.018)
ShareMinimumDecentrists	-0.043 (0.029)	0.004 (0.026)	0.089* (0.054)	0.073 (0.047)
ShareWhite	0.262*** (0.093)	0.165* (0.084)	0.174** (0.084)	0.338*** (0.074)
MeanIncome	0.002 (0.005)	0.003 (0.005)	0.002 (0.005)	0.007 (0.004)
ShareCollege	-2.014*** (0.236)	-0.449* (0.269)	-0.485* (0.268)	-0.044 (0.237)
SizeMetro	-0.144*** (0.040)	0.253*** (0.037)	0.246*** (0.037)	0.153*** (0.018)
DensityMetro	0.022 (0.015)	0.015 (0.013)	0.002 (0.001)	0.023* (0.016)
Year16	-	-0.046*** (0.005)	-0.045*** (0.007)	0.129*** (0.018)
ShareMinimumCentrists · Year16	-	-	0.023 (0.024)	0.016 (0.021)
ShareMinimumDecentrists · Year16	-	-	-0.075 (0.054)	-0.084* (0.046)
ShareWhite · Year16	-	-	-	-0.219*** (0.021)
Metro Area Fixed Effects?	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
\bar{R}^2	0.20	0.26	0.26	0.31
<i>N</i>	673	673	673	673

Table 5: Fixed Effects Regressions (Standard errors in parentheses)

estimates for lower bounds are insignificant – as we should expect. Yet our estimates for the interaction terms are insignificant also – contradicting our theory.

And then we may still have misspecified the empirical model. Arguably we should allow regression slopes not just to differ across election years for minimum centrists and decentrists. They should be allowed to differ for the share of whites, too. The Republican anti-immigration platform that had emerged by 2016 radically differed from that in 2008. Column (10) accounts for Donald Trump’s appeal to white voters with an anti-immigrant preference, by also interacting the share of whites with the 2016 year dummy. We find the following. Not only is the corresponding coefficient estimate of the interaction term strongly negative, and highly significant. Also, the negative coefficient estimate on the interaction term for minimum decentrist share turns significant now (at the 10% level). Suppose that one metro area had a minimum decentrist share in 2010 that was 10 percentage points higher than that of some other (otherwise identical) metro area. Then the Democratic vote share would have been smaller, too, by 0.84 percentage points.

Thought Experiment. We employ the column (10) result to carry out a simple thought experiment. Multiplying our coefficient estimate of -0.084 by minimum decentrist share λ_i^d in metro area i gives the drop in Clinton’s share of the vote in i caused by Trump’s disavowing global warming. Multiplying by the total vote in that metro area and summing over all metro areas with positive λ_i^d yields a predicted drop of 279,529 votes. It is in that sense that the US urban landscape has interacted with Trump’s agenda to contribute to Hillary Clinton’s defeat.²¹ Somewhat more explicitly, where in the earlier presidential election John McCain and Barack Obama appeared to agree on the necessity of combatting climate change, Donald Trump’s shifting away from this consensus and embracing the climate sceptic’s position may have secured him the crucial extra vote that put him into office.

7 Conclusions

This paper offers a centrist-decentrist perspective on the 2016 U.S. presidential election, and suggests that denying climate change has helped Donald Trump win in 2016. It also suggests that climate change may continue to divide American society. Because housing is durable, today’s city shapes will persist into the future. We must expect the rift between centrists and decentrists over future policy towards global warming to afflict future elections, too. One remedy is to compensate sprawling cities – whose homeowners are bound to lose from fighting global warming – for tolerating a carbon tax. Another is to encourage planning for more compact cities. More compact cities not necessarily emit less carbon (see Gaigne/Riou/Thisse (2012), Borck (2016)). But they are more likely to embrace taxing it.

Going beyond U.S. presidential elections, we provide novel estimators of centrists and decentrists. These estimators are highly general, and are simple to apply to any arbitrary city. We suggest that conflicts between centrists and decentrists could help explain other contested policies beyond carbon taxation, too. Examples include rationing central city land, decentralizing retail, tightening building height limits, implementing minimum lot

²¹Florida (2016) appears to suggest as much when arguing that the share of those “who drive alone”, as an indicator of sprawl, should reduce Clinton’s share of the vote.

size, or introducing road tolls. To these policy fields this paper's greatest lower bounds on centrists and decentrists are readily applicable, too.

8 Literature

- Arnott, R. and J. Stiglitz (1981) Aggregate Land Rents and Aggregate Transport Costs, *Economic Journal* 91: 331-347.
- Borck, R. (2016) Will Skyscrapers Save the Planet? Building Height Limits and Urban Greenhouse Gas Emissions. *Regional Science and Urban Economics* 58: 13-25.
- Breheny, M. (2007) Centrists, Decentrists, and Compromisers: Views on the Future of Urban Form, in: H. Priemus, K. Button and P. Nijkamp (eds.) *Land Use Planning*: 47-69.
- Brueckner, J. (1987) The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills-Model, in: E. Mills (ed.), *Handbook of Regional and Urban Economics* Vol. II: 821-845.
- Brueckner, J. and A. Glazer (2008) Urban Extremism, *Journal of Law, Economics and Organization* 24.2: 307-318.
- Chvatal, V. (1980) *Linear Programming*, Freeman and Company.
- Dascher, K. (2019) Function Follows Form, in: *Journal of Housing Economics* 44: 131-140.
- Florida, R. (2016) How America's Metro Areas Voted, in: *City Lab*, Dec. 6th.
- 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.
- Hadley, G. (1963) *Linear Programming*, Addison.
- Holian, M. and M. Kahn (2015) Household Demand for Low Carbon Public Policies: Evidence from California, *Journal of the Association of Environmental and Resource Economists* 2.2: 205-234.

- Jacobs, J. (1961) *Death and Life of Great American Cities*, Vintage Books.
- Luenberger, D. and Y. Ye (2016) *Linear and Nonlinear Programming*, Heidelberg: Springer, 4th ed.
- Pines, D. and Sadka, E. (1987) Comparative Statics Analysis of a Fully Closed City, *Journal of Urban Economics* 20: 1-20.
- Spivak, M. (1994) *Calculus*, 3rd ed., Cambridge University Press.
- Wheaton, W. (1973) A Comparative Static Analysis of Urban Spatial Structure, *Journal of Urban Economics* 9: 223-237.
- Wilson, St. et al. (2012) *Patterns of Metropolitan and Micropolitan Change: 2000 to 2010*. Report Number C2010SR-01 (U.S. Census Bureau).

9 Appendix

9.1 Cumulative Ring Difference

We introduce some extra notation first. Consider the cumulative sum $\Delta(h)$, for some h between 1 and $n/2$. Suppose $\Delta(h)$ is preceded by some other cumulative sum $\Delta(g)$ that is greater than it, i.e. $\Delta(h) < \Delta(g)$ for $g < h$.²² Borrowing terminology established in the context of the “Rising Sun Lemma” (Spivak (1994)), then we will say that the cumulative sum up to, and including, ring difference h is “in the shadow of” the cumulative sum up to, and including, ring difference g .²³ Of course, there will be rings that are never overshadowed. Among those, ring i^* , defined in equation (6), is the one exhibiting the greatest cumulative sum, $\Delta(i^*)$. For the two numerical cities in section 3, to give an example, $i^* = 3$.

9.2 Not All Ring Differences Negative

In the main text we took the first step towards a fully general analysis. Our point of departure was the city of the type spelt out in Table (6). The table header has the ring difference index i , the second row provides ring difference δ_i 's sign, and the third row indicates whether or not the corresponding cumulative ring difference $\Delta(i)$ is overshadowed (\bullet is a suggestive shorthand) or not (\circ). As mentioned above, in this city all ring differences both prior to i^* and beyond $i^* + 1$ are negative and overshadowed.

²²We define $\Delta(0) \equiv 0$. Even the first ring may be overshadowed, by $\Delta(0)$, if $\delta_1 < 0$.

²³In our first example city, the cumulative ring difference at 4 is overshadowed (by the cumulative ring difference at 3, say), while in the second example city cumulative ring differences at 2 and 4 are (by cumulative ring differences 1 and 3, respectively, for example).

i	1	2	3	...	$i^* - 1$	i^*	$i^* + 1$	$i^* + 2$...	$n/2$
δ_i	-	-	-	-	-	+	-	-	-	-
$\Delta(i)$	•	•	•	•	•	○	•	•	•	•

Table 6: A Parametric City

Nothing of substance changes if one (or more) of those shadow differences is (are) positive, rather than negative. To see this we turn to the city set out in Table (7) below, with the second ring the one ring to have flipped its sign. We assume that everything else remains the same, and so $\Delta(2) < \Delta(0)$ while i^* keeps maximizing $\Delta(i)$.²⁴

i	1	2	3	...	$i^* - 1$	i^*	$i^* + 1$	$i^* + 2$...	$n/2$
δ_i	-	+	-	-	-	+	-	-	-	-
$\Delta(i)$	•	•	•	•	•	○	•	•	•	•

Table 7: Negative and Positive Ring Differences

We introduce the following three, i.e. not numerous, changes to the primal's solution: (i) Entry $x_{2,n-1}$ ceases to be b_2 and turns into b_{n-1} instead. (ii) Entry x_{2n} becomes $(b_2 - b_{n-1})$, replacing the zero it was before. (iii) Entry $x_{i^*,n}$ drops from ring difference $(b_n - b_1)$ to the "difference of ring differences" $(b_n - b_1) - (b_2 - b_{n-1})$. These changes maintain feasibility, as is easily checked by consulting the housing constraints of the four rings affected.

Note that x_{i^*,i^*} is not among the entries changed. This particular entry continues to equal $\Delta(i^*)/2$. Since this entry is the only one to enter the primal objective's optimal value, our formula does not change either. Note the role of ring 2 still being overshadowed here. While δ_2 is positive, it is not sufficiently so to offset the negative δ_1 that precedes it. And hence $(b_n - b_1) - (b_2 - b_{n-1})$ or $x_{i^*,n}$ indeed is strictly positive. Now let us check the implied changes for the dual. Since $x_{2,n-1}$ and $x_{i^*,n}$ continue to exceed zero, complementary constraints of the dual continue to be binding. And since x_{2n} now also exceeds zero, the corresponding dual constraint becomes binding, so that $y_2 = -y_n$. This we knew before, and so this extra equation is redundant. We conclude that formula $\Delta(i^*)/2$ continues to apply. Of course, the objective's numerical value changes.

Exploring a sign change for any other ring difference, or for additional ring differences, proceeds along similar lines. That is, formula $\Delta(i^*)/2$ continues to capture the minimum number of centrists whatever the signs of the ring differences in rings up to i^* , as long as these ring differences are overshadowed.

9.3 Not All Ring Differences Overshadowed

What (if anything) changes if one (or more) of the ring differences are not overshadowed? Let us allow for the possibility that not all ring differences prior to i^* are overshadowed, as in Table (8). Let all ring differences from 1 up to $i' - 1$ be in the shadow of ring 0, and

²⁴These assumptions are not restrictive. First, if $\Delta(0) < \Delta(2)$, we would have to consider alternating spells of ring differences in the shadow and not in the shadow. This case is considered shortly. And second, if i^* shifted due to δ_2 flipping its sign, nothing would change in the argument below as long as $2 < i^*$.

all ring differences between $i' + 1$ and $i^* - 1$ be overshadowed by i' , so that i^* is not in the shadow. One optimal feasible solution assigns $\sum_{j=1}^{i'} \delta_j/2$ to $x_{i'i'}$, and $\sum_{j=i'+1}^{i^*} \delta_j/2$ to $x_{i^*i^*}$, and zero to any other element above the counterdiagonal. The corresponding minimum number of centrists becomes the sum of these two (only non-zero) terms. But this is just our familiar $\Delta(i^*)/2$. Adding extra spells of ring differences in the shadow adds nothing of substance here.

i	1	...	$i' - 1$	i'	$i' + 1$...	$i^* - 1$	i^*	$i^* + 1$...	$n/2$
$\Delta(i)$	•	•	•	○	•	•	•	○	•	•	•

Table 8: Alternating Spells of Shadow and Light

At last we turn to the question of what happens if any ring differences following $i^* + 1$ (rather than preceding i^*) exhibit a positive sign. Recall that, by definition of i^* , ring differences beyond i^* must be overshadowed. Let one of these ring differences be positive, rather than negative, i.e. $i^* + 2$ say. Being in the shadow of i^* , the excess arising in ring difference $i^* + 2$ is “swamped” by the deficit in the previous ring difference at $i^* + 1$. Once more, there is no change in the number of minimum centrists.

9.4 Sample Statistics

	min	25%-qu	med	75%-qu	max
DemShare ₀₈	0.19	0.38	0.47	0.56	0.88
DemShare ₁₆	0.19	0.35	0.43	0.53	0.81
MinimumCentristShare ₀₈ (Δ_{08}^c)	0	0.09	0.23	0.35	1.00
MinimumCentristShare ₁₆ (Δ_{16}^c)	0	0	0.08	0.21	0.76
MinimumDecentristShare ₀₈ (Δ_{08}^d)	0	0	0	0.03	0.32
MinimumDecentristShare ₁₆ (Δ_{16}^d)	0	0	0.03	0.12	0.66
ShareBachelor ₀₈	0.08	0.13	0.16	0.18	0.32
ShareBachelor ₁₆	0.08	0.13	0.17	0.20	0.32
ShareWhite ₀₈	0.48	0.74	0.84	0.90	0.97
ShareWhite ₁₆	0.43	0.74	0.82	0.89	0.96
SizeMetro ₀₈ (millions)	0.04	0.09	0.15	0.32	3.49
SizeMetro ₁₆ (millions)	0.04	0.10	0.16	0.37	4.11
DensityMetro ₀₈ (1,000 per sq mile)	0.51	1.26	1.91	2.76	8.30
DensityMetro ₁₆ (1,000 per sq mile)	0.52	1.27	1.86	2.71	8.42
MeanIncome ₀₈ (10,000 \$)	1.91	3.36	3.98	4.66	11.50
MeanIncome ₁₆ (10,000 \$)	1.82	3.32	3.97	4.74	17.58

Table 9: Descriptive Statistics

Metro Area	MinimumCentrists λ^c	MinimumDecentrists λ^d
Houston - Sugarland - Baytown	0.095	0.009
Atlanta - Sandy Springs - Marietta	0	0.023
Phoenix - Mesa - Glendale	0.299	0.002
Riverside - San Bernardino - Ontario	0.367	0.017
Minneapolis - St. Paul - Bloomington	0.202	0.013
San Diego - Carlsbad - San Marcos	0.415	0
Tampa - St. Petersburg - Clearwater	0.368	0
St. Louis	0.232	0.009
Baltimore - Towson	0.308	0
Denver - Aurora - Broomfield	0.382	0
Pittsburgh	0.274	0.008
Portland - Vancouver - Hillsboro	0.318	0
Charlotte - Gastonia - Rock Hill	0.031	0.043
Orlando - Kissimmee - Sanford	0.076	0.043
San Antonio - New Braunfels	0.182	0
Sacramento - Arden Arcade - Roseville	0.201	0.025
Cleveland - Elyria - Mentor	0.341	0.006
Las Vegas - Paradise	0.192	0.001
Kansas City	0	0.112
Columbus	0	0.046
San Jose - Sunnyvale - Santa Clara	0.761	0
Indianapolis - Carmel	0.135	0.020
Austin - Round Rock - San Marcos	0.101	0.021
Nashville - Davidson - Nurfreesboro - Franklin	0	0.064

Table 10: Minimum Centrist/Decentrist Shares, 25 Largest Metro Areas in 2016 subsample