Spatial Patterns and Size Distributions of Cities

Wen-Tai Hsu Tomoya Mori Tony E. Smith*†

September 2014

Abstract

City size distributions are known to be well approximated by power laws across many countries. By far the most popular explanation for such power-law regularities is in terms of random growth processes, where power laws arise asymptotically from the assumption of *iid* growth rates among all cities within a given country. But this assumption has additional consequences. Since all subsets of cities have the same statistical properties, each subset must exhibit essentially the same power law. Moreover, this *common power law* (CPL) property must hold regardless of the spatial relations among cities. Using data from the US, this paper shows first that spatial partitions of cities based on geographical proximity are significantly more consistent with the CPL property than are random partitions. It is then shown that this significance becomes even stronger when proximity among cities is measured in terms of trade linkages rather than simple geographical distance. These results provide compelling evidence that spatial relations between cities do indeed matter for city-size distributions. Further analysis shows that these results hinge on the natural "spacing out" property of city patterns in which larger cities tend to be widely spaced apart with smaller cities organized around them.

JEL Classifications : C49, R12
Keywords : city size distributions, power law, Zipf's law, random growth, inter-city space, geography, Voronoi partition, economic region, central place theory

---

*Hsu: School of Economics, Singapore Management University. Email: wentaihsu@smu.edu.sg. Mori: Institute of Economic Research, Kyoto University and Research Institute of Economy, Trade and Industry (RIETI) of Japan. Email: mori@kier.kyoto-u.ac.jp. Smith: Department of Electrical and Systems Engineering, University of Pennsylvania. Email: tesmith@seas.upenn.edu.

†For their helpful comments, we thank the seminar participants at Academia Sinica, Chinese Univ. of Hong Kong, Kobe Univ., Kyoto Univ., Singapore Management Univ., Tohoku Univ., Tokyo Univ., the 2013 Annual Meeting of the Urban Economics Association in Atlanta, III Workshop on Urban Economics in Barcelona, the 2014 Spring Meeting of the Japanese Economic Association in Kyoto. We are particularly grateful to Esteban Rossi-Hansberg for his detailed and insightful comments that helped to sharpen the focus of this work and to improve the paper in many other ways. This research is conducted as part of the project, *the formation of economic regions and its mechanism: theory and evidence*, undertaken at the Research Institute of Economy, Trade and Industry, and has been partially supported by the Grant in Aid for Research (No. 25285074 and the Global COE Program "Raising Market Quality") of the MEXT, Japan.

1 Introduction

City size distributions are known to be well approximated by power laws across a wide range of countries. The most popular approach to explaining this regularity at present is in terms of simple random growth processes (as in Gabaix [22]).[1] Such processes have been successfully incorporated in general equilibrium models that match actual city size distributions well (e.g., Duranton [17] and Rossi-Hanseberg and Wright [45]). But even in these more complex versions, power laws for city size distributions arise fundamentally from the underlying assumption of common *iid* growth rates for all cities, which is well known to have additional consequences. For if cities exhibit common *iid* growth rates, then *all* (sufficiently large) subsets of these cities must exhibit power laws with the same exponent. In particular, this *common power law* (CPL) property must hold regardless of the particular spatial relations that exist among cities. So these random growth models suggest that spatial relations among cities do *not* influence the distribution of city sizes.

However, there is a growing literature showing that space does indeed play a crucial role in shaping the economic landscape we observe. At the global scale, there is a long tradition in the international trade literature focusing on how trade frictions induced by inter-country distances (among other factors) influence trade flows between countries.[2] At the urban scale, there has been a long tradition in the urban economics literature focusing on how within-city spatial structure influences a variety of urban phenomena, including both housing and land markets.[3] Finally at the regional level, there is a small emerging literature more closely related to the present analysis that focuses on how spatial separation influences trade between cities and city growth, e.g., Donaldson [16], Duranton, Morrow and Turner [18], Hering and Poncet [30] , Michaels [40], Redding and Sturm [44], Fajgelbaum and Redding [20].

Taken together, these many research efforts suggest that the distribution of city sizes may indeed be influenced by the spatial relations among these cities. To study this question, we begin by postulating that the spatial organization and sizes of cities are linked by the *spacing-out property* that *larger cities tend to be widely spaced apart, with smaller cities*

[1] It is well documented that power laws are good descriptors of city size distributions, especially in their upper tails; see Rosenfeld et al. [46] and Ioannides and Skouras [34]. In particular, the random growth processes proposed by Gabaix only imply power laws for the upper tails of their steady-state city size distributions (see the discussion in Section 2.1). See Gabaix [23] for a survey on the extensive empirical literature on city size distributions, as well as Eeckhout [19] for similar processes that generate log-normal city size distributions.

[2] Such inter-country distances are indeed one of the most fundamental explanatory variables in all gravity-type regression models. See Anderson and van Wincoop [6] for a survey of this extensive literature.

[3] See Anas, Arnott, and Small [5] for a survey of this substantial body of literature. More recent developments can be found in Lucas and Rossi-Hansberg [38] and Ahlfedlt, Redding, Sturm, and Wolf [1].

grouped around these centers. For city landscapes that do exhibit this property, one might expect to find similar size relations among the cities in each spatial grouping. This in turn suggests that the CPL property above may indeed be stronger for such groupings than for arbitrary groupings of cities. Given this line of reasoning, our main objectives are to develop explicit tests of these hypotheses. Our first set of tests provide evidence that consistency with the CPL property is significantly higher for even simple groupings of *nearby* cities (without regard to the spacing-out property) than for arbitrary groupings of these cities. Our second set of tests provide independent evidence for the spacing-out property itself, without regard to the CPL property. Finally we combine certain aspects of these two lines of investigation by replacing groupings of nearby cities in the CPL tests with appropriately defined "economic regions" that are closer in spirit to our postulated spacing-out property. Our test results here confirm that consistency with the CPL property is even higher for these economic regions than for groupings of cities based on simple proximity relations as above.

With this brief overview, we now consider each of these testing procedures in more detail. The data used for all tests is taken from the US in 2007. In particular, *cities* are here defined to be *Core Based Statistical Areas* (CBSAs) [see Figure 7(a)].[4] Using this data, our first set of tests focus on spatial groupings of cities *without* regard to major cities themselves. The question is whether groupings of *nearby* cities are more comparable in terms of CPL than are arbitrary groupings of cities. For each number of possible groupings, K , this is accomplished by selecting K cities at random and identifying the subsets of cities closest to each of these K cities. Formally these subsets constitute a *Voronoi K-partition* in which cities are spatially grouped in the sense that all cities in the same *Voronoi region* (or cell) are closest to a common city. Power laws for the cities in each cell are then estimated by log regressions of size against rank. As detailed in Section 2.2 below, it is convenient to replace both log(rank) and log(size) by their smoother "upper average" versions which facilitate comparisons of the upper tail structures of such distributions. In this context, the level of agreement between power laws for each cell of cities is essentially determined by comparing the similarity of slopes between these log regressions (as detailed in Section 2.3 below). Finally, by simulating random K-partitions of cities and carrying out the same regression procedure, one can test whether there is significantly stronger agreement among the power-laws of these K Voronoi regions than would be expected if they were simply cells in a random K-partition of cities. With respect to these tests, our main result is to show that our US data indeed exhibit stronger agreement with CPL among these Voronoi regions over a broad range of K values.

[4] See the US Office of Management and Budget [61] for the definition of CBSA.

This initial set of tests involve only a minimal concept of "space" and make no assertions about the spacing-out property itself. But further analysis of the test results shows that a key difference between Voronoi partitions and arbitrary partitions relates to the placement of largest cities among their cells. In particular, those Voronoi partitions exhibiting the strongest agreement with the CPL property tend to be those in which the largest cities appear in *different* cells, and are thus associated with the groupings of nearby cities. In this sense, the present results can be said to establish an indirect link between the spacing-out property and CPL property.

Our second set of tests pursue this line of reasoning further by asking whether this separation property continues to be present in *all* Voronoi partitions versus random partitions. If so, then this provides compelling evidence for the spacing-out property itself, without regard to CPL. Such relationships are easily testable for, say, the r largest cities by simply counting the number of cells containing at least one of these cities in a given Voronoi K-partition, and comparing such counts with those of randomly generated K-partitions. By simulating many such comparisons, one can then determine whether these r cities are distributed over a significantly larger number of Voronoi cells than random cells. Our results show that there is indeed a significant difference.

But by their nature, these tests focus more on the separation between large cities than on the clustering of smaller cities around them. Thus, to test this latter part of the spacing-out property, we construct Voronoi partitions with reference cities that correspond to the K largest cities rather than K randomly chosen cities. For these *largest-city Voronoi K-partitions*, we then calculate the distance of each city to its reference city, and designate the sum of these distances across all cities as the *total distance measure* for this K-partition. If smaller cities are indeed clustered around the largest cities, then one would expect total distances for these largest-city Voronoi K-partitions to be significantly smaller than for similar Voronoi K-partitions with randomly selected reference cities. Our tests show that this is indeed the case. Moreover, by using an alternative measure, *total population-weighted distance*, in which the distance of a city to the reference city is weighted by the city population, the results become even stronger. These results together with the spacing out of the largest cities explain why Voronoi partitions tend to exhibit higher consistency with the CPL property than their random counterparts. In particular, Voronoi cells containing the largest cities tend also to contain substantial portions of their corresponding city clusters.

Finally, as mentioned above, we combine some of these ideas by replacing Voronoi K-partitions in the CPL test with a set of "economic regions" based explicitly on a commodity-flow interpretation of the spacing-out property. Essentially, each such *economic region* consists of a large city together with other smaller cities for which the commodity inflows from this city are larger than from any other city other than from themselves.[5] This construction essentially mirrors the spacing-out property with distance replaced by trade flows. The CPL property for the economic regions generated by the K largest cities is then tested against random K-partitions for which the numbers of cities in each cell are the same as those in each economic region. The results of these tests confirm that over a considerable range of partition sizes, K, the CPL property is even more significant for these economic regions (relative to their random counterparts) than for the simple groupings of nearby cities above.

In relating these results to the existing literature, we note first that surprisingly few empirical studies have examined the spacing patterns of cities in a systematic way, let alone attempted to identify specific properties of such patterns.[6] Thus the main contributions of the present paper are to document the spacing-out property for city locations, and to examine its relation to city-size distributions in terms of the CPL property.[7] In reference to the absence of such a relationship as implied by the *iid* growth-rate assumption in random growth models, our first set of test results show that even modestly spatial groupings of "nearby cities" exhibit significantly stronger CPL properties than arbitrary groupings. This by itself would seem to provide compelling evidence that spatial relations among cities do indeed matter. The more refined results in terms of economic regions only serve to strengthen this conclusion. In this regard, the present paper is closely related to a series of recent studies that document possible deviations from the assumption of *iid* growth rates, including Desmet and Rappaport [13], Black and Henderson [9], Holmes and Lee [29], Michaels, Rauch, and Redding [41], and Redding and Sturm [44]. One example particularly relevant to our paper is the empirical study by Redding and Sturm [44] documenting the effect of the post World War II German separation on the growth of cities near the border. Their results suggest that there was indeed a certain degree of dependence between the growth rates of *nearby* cities in this region.

Second, beyond the validity of random growth hypotheses, our results have important implications for a broader class of city-systems models. In particular, a number of structural models have recently been developed to provide quantitative assessments of the determinants of city size, including Desmet and Rossi-Hansberg [14], Behrens et al. [8], Allen and

[5]The trade flow data used here is based on the 2007 Commodity Flows Survey.

[6]The most closely related work in this respect appears to be that of Dobkins and Ioannides [15] and Ioannides and Overman [33] who find mixed results regarding the role of space (distance among cities) in influencing various city phenomena, such as size, growth, and emergence of cities.

[7]Giesen and Südekum [26] also use regional level data to examine city size distributions. However, their focus is on testing whether Gibrat's law holds in each subset of cities in Germany, and they do not test CPL per se.

Arkolakis [4], and Redding and Sturm [44]. But, these models have mostly attributed the size of cities to various city-specific factors. For example, Desmet and Rossi-Hansberg [14] postulate that city size is an implicit function of three city-specific factors ("efficiency", "amenities", and "frictions"). So all variations in city size are by construction absorbed into these factors. In particular, there is no explicit notion of inter-city space in their model. Rather cities are linked only by a free-mobility condition under general equilibrium. Thus any possible mechanisms underlying the CPL and spacing-out properties identified by our results are necessarily left unexplained by these models.

Third, this paper is also closely related to Behrens et al. [8] who formulate a spatial economic model with trade costs between cities, and estimate this model using US data. In this modeling context, regression results based on their estimates suggest that "spatial friction" (in terms of the cost of trade between cities) does not significantly influence the distribution of city sizes. However, it should be stressed that our present testing framework is independent of any specific economic modeling assumptions. Thus the results obtained here suggest that quantitative significance of spatial relations between cities on city-size distributions remains to be captured by current structural models.[8]

Finally, our results suggest certain directions for extending current theories of city systems. In particular, while these findings raise questions about the *iid* assumption, they should not to be taken as rejection of the random growth approach itself. Indeed, there may be ways to relax the *iid* assumption by allowing certain spatial dependencies among city growth rates that continue to yield power laws in the upper tail.[9] Also, the original results of Gabaix [22, Proposition 2] suggest that it may be possible to "regionalize" such *iid* assumptions in a manner more consistent with our findings. The current theoretical models which most closely account for our results are those based on the *central place theory*, dating back to the original work of Christaller [11]. The central tenets of this theory assert that the heterogeneity of goods/industries together with the spatial extent of markets give rise to hierarchies of cities, and thus to a diversity of city sizes. Along these lines, the recent model by Hsu [31] based on micro-economic behavior exhibits the CPL property among economic regions. However, his firm-entry model is highly stylized, and more realistic

[8]Note also that there is no clear correspondence between "spatial frictions" and "spatial patterns of cites". In fact it is possible to have spatial economic models in which the spatial pattern of cities is entirely independent of spatial frictions in terms of (positive) transport costs between cities (e.g., Hsu [31]). So direct comparisons between the effects of spatial frictions and spatial patterns of cities on the distribution of city sizes are at best problematic. Also see a more detailed discussion in the conclusion on several other recent structural model that account for city size differences.

[9]For example, the recent Markovian approaches to Kesten processes by Saporta [47] and Ghosh et al. [25] might offer possible methods for allowing spatial dependence between growth rates in a random growth process.

general equilibrium models are desirable. We will discuss these issues in more detail in the concluding remarks.

The rest of the paper is organized as follows. Section 2 introduces an estimation strategy for CPL and define the goodness of fit of such an estimation. Section 3 conducts CPL tests by comparing the goodness of fit under Voronoi partitions of cities with that under random partitions. Section 4 examines the spacing-out property. Section 5 constructs economic regions and conducts CPL tests by comparing the goodness of fit under economic regions with that under random partitions. Section 6 concludes.

2 Methods for Analyzing Common Power Laws

Before developing these test results, it is convenient to begin with a number of methodological tools that will be used throughout. First we briefly consider the explicit class of stochastic growth models known as *Kesten processes*. These will provide us with a way of simulating processes with *known* asymptotic power laws that can be used to test our methods. Next we introduce an *upper-averaging* method for estimating power-law exponents that is particularly useful for our present purposes. Finally we develop the *categorical regression* framework that will be used to compare the degrees of the similarities among power laws across subsets of cities.

2.1 Kesten Processes

As first introduced into stochastic urban growth theory by Gabaix [22], Kesten processes provide a simple class of stochastic growth models that exhibit asymptotic power laws under fairly weak conditions. For a given a collection of *cities*, $i = 1, \ldots, n$, if S_{it} denotes the *size* (population) of city i in time period t, then it is hypothesized that the city sizes evolve over time according to a stochastic difference equation of the form

$$S_{i,t+1} = \gamma_{it} S_{it} + e_{it}, \quad i = 1, \ldots, n\,; \quad t = 1, 2, \ldots \tag{1}$$

where $(\gamma_{it} : i = 1, \ldots, n,\ t = 1, 2, \ldots)$ is a sequence of independently and identically distributed (*iid*) nonnegative *growth multipliers* and $(e_{it} : i = 1, \ldots, n,\ t = 1, 2, \ldots)$ is a sequence of small nonnegative *growth increments*. If in addition, it is assumed that these two sequences are mutually independent, then (1) is said to define a *Kesten process*.[10] Note

[10]These processes were first introduced by Kesten [35] as multivariate (matrix-valued) processes. More accessible treatments of the univariate case can be found in Vervaat [54] and Goldie [28].

that the individual processes for each city are essentially independent copies of one another, and hence must exhibit the same asymptotic behavior. In particular, it can be shown that under very general conditions there exists a limiting random variable, S, such that each city process *converges in distribution* to S, i.e.,

$$\lim_{t\to\infty} S_{it} =_d S, \quad i = 1, \ldots, n \tag{2}$$

(where $=_d$ denotes equality in distribution). More importantly, if γ denotes a representative growth multiplier, and if there exists a positive exponent, κ, for which $E(\gamma^\kappa) = 1$, then under very weak additional conditions, it can be shown that S satisfies an *asymptotic power law with exponent* κ, i.e., that there exists a positive constant, c, such that,

$$\lim_{s\to\infty} s^\kappa \Pr(S > s) = c\,, \tag{3}$$

which is more conveniently written as

$$\Pr(S > s) \;\approx\; c\, s^{-\kappa}, \quad s \to \infty\,. \tag{4}$$

So the city sizes in (1) can eventually be treated as independent random samples from a distribution with this property. In our simulations of such processes, we shall assume that growth multipliers, γ, are log normally distributed, and in particular that $\ln(\gamma) \sim N(\mu, 1)$. Here it can be shown (see Gabaix [22]) that the desired exponent, κ, is given by

$$\kappa = -2\mu \tag{5}$$

for $\mu < 0$. In addition, we assume that the small growth increments, e, are uniformly distributed on $[0, 0.01]$.

2.2 Upper Average Smoothing of Rank-Size Distributions

If a given set of cities is postulated to exhibit an asymptotic power law as in (4), and if cities are ranked by size as $s_1 \geq s_2 \geq \cdots \geq s_n$ so that i denotes the relevant *rank* of city i, then a natural estimate of $\Pr(S > s_i)$ is given by the ratio (i/n). So by (4) we obtain the following approximation,

$$\begin{aligned} i/n \approx \Pr(S > s_i) \approx c s_i^{-\kappa} \quad &\Rightarrow \quad \ln(i) \approx \ln(cn) - \kappa \ln(s_i) \\ &\Rightarrow \quad \ln(s_i) \approx b - (1/\kappa)\ln(i) \end{aligned} \tag{6}$$

where $b = \ln(cn)/\kappa$. This motivates the standard log regression procedure for estimating κ based on "rank-size" data, $[\ln(i), \ln(s_i)],\ i = 1, \ldots, n$. But as observed by many authors (e.g., Gabaix and Ibragimov [24] and Nishiyama et al. [42]) this log regression tends to underestimate the true value of κ. Several approaches have been proposed for correcting this bias, including the "1/2" rule of Gabaix and Ibragimov [24] and the "trimming rule" of Nishiyama et al. [42].

However our present objectives are somewhat different. Here we are primarily interested in comparing similarities between the *upper-tail* properties of city size distributions for different subsets of cities within a country. With this in mind, we start by smoothing the usual rank-size data in a manner that emphasizes the upper tails of this data. In particular, we transform the data $[\ln(i), \ln(s_i)],\ i = 1, \ldots, n$, by taking *upper averages* to obtain new data pairs, *upper log rank*, ULR_i, and *upper log size*, ULS_i, as defined respectively by

$$ULR_i = \frac{1}{i}\sum_{j=1}^{i} \ln(j)\,, \tag{7}$$

$$ULS_i = \frac{1}{i}\sum_{j=1}^{i} \ln(s_j)\,. \tag{8}$$

These upper averages smooth the data in a manner that emphasizes the largest cities. The theoretical and practical relevance of this transformation can be illustrated by the two plots in Figure 1. In Figure 1(a) we have plotted the rank-size data, $[\ln(i), \ln(s_i)]$, for the US as blue circles, and have superimposed the corresponding upper-average data, $[ULR_i, ULS_i]$, as a red curve (with points connected by lines for visual clarity).

As is typical for such rank-size plots, log linearity is most evident in the upper tail (largest cities) where the power law starts to emerge. In contrast, there is little indication of such a power law in the lower tail (smallest cities) where values decrease dramatically. So when $\ln(s_i)$ is regressed against $\ln(i)$, it should be clear that the regression line is "pulled down" by these lower-tail values, and becomes too steep. As seen from expression (6), this should result in an *underestimation* of κ. In contrast, the upper-average plot is not only smoother, but is also shifted toward the upper tail where the power law is most evident. This is reflected by the corresponding regression results, in which the US rank-size data yields an estimated slope of $\hat{\theta} = -1.219$, with corresponding power exponent, $\hat{\kappa} = 1/\hat{\theta} = 0.821$, while the US upper-average data yields the "flatter" slope estimate, $\hat{\theta} = -1.059$, with power exponent, $\hat{\kappa} = 0.944$.

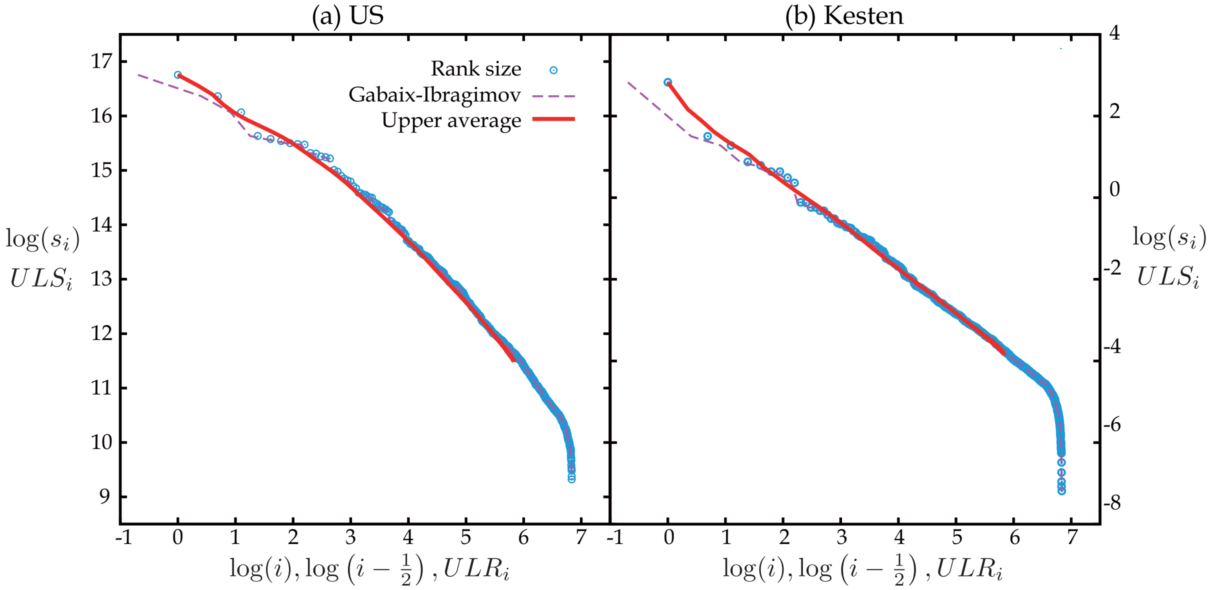

Figure 1: City size distribution from Kesten process

As an additional comparison, we also include results for the "1/2" rule by Gabaix and Ibragimov under which log rank, $\log(i)$, is replaced by $\log(i - \frac{1}{2})$ in the rank-size regression, thus weighting larger cities more heavily in a manner analogous to our upper-average approach.[11] But since this weighting scheme is somewhat less extreme than upper-averaging, the corresponding regression results yield a slope estimate, $\hat{\theta} = -1.200$, with power exponent, $\hat{\kappa} = 0.833$, larger than the standard estimate under the rank-size regression but smaller than that under the upper-average regression. The Gabaix-Ibragimov data, $[\log(s_i), \log(i - \frac{1}{2})]$, is shown by the dashed curve in Figure 1(a) (again with points connected by lines as in the upper-average data). From the plot, it is rather obvious that the upper-average data is most successful in picking up the "power-law content" from the entire distribution.

But since the "true" exponent for the US is not known, this comparison leaves much to be desired. What is needed is an example in which the true exponent is actually *known*. In this way, the relative accuracy of these methods can be compared in a more meaningful way. To do so, we have simulated a Kesten process that roughly approximates the US case. In particular, we set $n = 930$ (as in our US data) and use (4) to construct a Kesten process with power-law exponent, $\kappa = -1/\theta = 1/1.059 \simeq 0.944$, based on the upper-average

[11] It is to be noted that while the Gabaix-Ibragimov approach is often adopted to estimate the power-law exponent of the city size distribution (e.g., Behrens et al. [8]), this approach assumes that the city sizes follow an exact Pareto distribution. But actual city size distributions at the national level are often more similar to those obtained from simulated Kesten processes as in Figure 1(b) (see also Rossi-Hansberg and Wright [45]).

estimate above.[12] Starting from uniform city sizes, $S_{i0} = 1$ for all $i = 1, \ldots, n$, the steady state was approximated by iterating this process until the mean city sizes converged with respect the criterion:

$$|\overline{S}_t - \overline{S}_{t-1}| < 0.0001 \times \overline{S}_t \tag{9}$$

where $\overline{S}_t \equiv \frac{1}{n}\sum_{i=1}^{n} S_{it}$.[13] The resulting (scaled) output, $[\ln(i), \ln(s_i)]$, is shown by the blue circles in Figure 1(b). Again the transformed upper-average data, $[ULR_i, ULS_i]$, is plotted in red. Here the rank-size regression again underestimates κ with an estimated value of $\hat{\kappa} = 0.841$ $[= -1/\hat{\theta} = 1/1.189]$. The Gabaix-Ibragimov regression underestimates less, with an estimated value of $\hat{\kappa} = 0.853$, and again the upper-average regression comes closest to the true value with an estimated value of $\hat{\kappa} = 0.909$.

To gauge the robustness of this particular result, steady states were obtained for 1000 replications of the present Kesten process, and regressions were run for each replication using the rank-size (R-S), Gabaix-Ibragimov (G-I) and upper-average (U-A) approaches. In comparison to R-S/G-I estimates, the U-A estimates were closer to the true value ($\kappa = 0.944$) in all but 31/43 of the 1000 cases. The average absolute errors over the 1000 simulations for the R-S, G-I and U-A estimates were 0.1324, 0.1207 and 0.0427, respectively, i.e., the relative estimate errors for U-A versus R-S and versus G-I are 0.0427/0.1324 = 0.322 and 0.0427/0.1207 = 0.354, respectively. These results suggest that this upper-average procedure does tend to yield more reliable estimates.[14]

Finally it is worth noting that if one is interested in the upper-tail properties of a distribution, then it would seem that an obvious approach is simply to truncate the lower tail. For example, a visual inspection of Figure 1(a) suggests that the distribution for US cities could best be truncated by removing all ranks above say 600 ($ULR_{600} \approx 6.40$). But for arbitrary subsets of cities (such as those considered throughout the present paper), the systematic identification of "optimal" truncation points is not at all obvious.[15] In this light, the present upper-average approach provides a reasonably robust procedure for approximating the upper-tail structure of arbitrary city-size distributions without the need

[12] Using expresssion (5), the growth multipliers, γ , were simulated by taking independent draws of $\ln(\gamma)$ from the normal distribution, $N(\mu, 1)$ with $\mu = -\kappa/2 = -0.472$.

[13] While condition (9) is only a necessary condition for a true steady-state, this approximation appears to work reasonably well for our present purposes. Among the 1000 simulations generated below, the minimum and the maximum numbers of iterations required to achieve condition (9) were 1002 and 19,821, respectively (with an average of 3356 iterations).

[14] It should also be noted that the basic results do not change for alternative values of $\kappa < 1.0$ [i.e., for those κ values where the power-law approximation, eq.(4), to the upper tail makes sense].

[15] However, if one hypothesizes that city size data is exactly Pareto distributed, then a reasonable optimal-truncation approach has been proposed by Clauset et al. [12] in terms of maximum likelihood estimation. See the introduction to Appendix B for further discussion of this approach.

to specify truncation points.[16,17]

2.3 A Categorical Regression Framework

As stated in the introduction, our main objective in this paper is to compare the values of estimated power-law exponents for different subregions. So the smoothing achieved by upper averaging has the additional advantage of sharpening these comparisons from statistical perspectives.

It should be clear that many different summary statistics can in principle be used for measuring the similarity between sets of slopes. But a particularly convenient approach for our present purposes is based on categorical regression. To begin with, if for any given set of regions, $j = 1, \ldots, m$, we consider the null hypothesis that the slopes for these regions are identical, then under this hypothesis, the upper-average plots should differ only by their intercepts and not their slopes. So their common slope can be estimated by a simple categorical regression with regional fixed effects. To formalize this model, observe first that if each region j contains n_j cities, then for each city-region pair $(ij : i = 1, \ldots, n_j,\ j = 1, \ldots, m)$ one can use (7) and (8) to define the appropriate upper-average rank and size variables as follows:

$$ULR_{ij} = \frac{1}{i}\sum_{h=1}^{i} \ln(h) \equiv ULR_i \,, \tag{10}$$

$$ULS_{ij} = \frac{1}{i}\sum_{h=1}^{i} \ln(s_{hj}) \,, \tag{11}$$

where the identity, $ULR_{ij} \equiv ULR_i$, follows from the fact that this quantity is the same for all relevant j (i.e., all j with $n_j \geq i$). Finally, if we let region 1 denote the choice of a "reference" region and for each other region, $j = 2, \ldots, m$, define the indicator variable, D_j, by $D_j(h) = 1$ if $h = j$ and zero otherwise, then the desired *categorical regression model*

[16] By the same reasoning, this upper-average approach may also be useful for cross-country comparisons.

[17] As a robustness check on our results for U-A data, we carried out all tests in Sections 3 and 5 below using both G-I and R-S data as well. In this regard, the Voronoi results in Section 3 appear to be quite robust, and the basic conclusions remain the same for all three data sets. However, there are some differences for the economic-region results in Section 5. In particular, the R-S data fails to capture any significant CPL properties of economic-region partitions for $K > 2$. (As noted above, this may be due to the overemphasis of R-S data on lower tail properties, which magnifies the bias of log linear regression estimates.) But differences between G-I and U-A are far less dramatic. While G-I places more emphasis on similarities between the mid ranges of city-size distributions than does U-A, the basic conclusions regarding CPL properties are the same for both data sets. See Appendix B for the details.

takes the form,

$$ULS_{ij} = \alpha + \theta\, ULR_i + \sum_{h=2}^{m} \beta_h D_j(h) + \varepsilon_{ij}\,. \tag{12}$$

Here it should be emphasized that this regression framework is used only to provide a convenient least-squares framework for gauging how well the given regional data agrees with the null hypothesis of a common slope. Since we are not concerned with the distribution properties of coefficient estimators in this nonparametric setting, there is no need to make assumptions about the residuals, ε_{ij}. Hence, letting $n = \sum_{j=1}^{m} n_j$, we simply adopt the *Root Mean Squared Errors* (*RMSE*) statistic,

$$RMSE = \sqrt{\frac{1}{n}\sum_{i,j}(ULS_{ij} - \widehat{ULS}_{ij})^2} \tag{13}$$

for this regression as an appropriate measure of goodness-of-fit,[18] and employ this statistic to construct a series of nonparametric tests (as detailed in Section 3 below).

To gauge how well this categorical regression procedure works using the U-A data, $[ULR_{ij}, ULS_{ij}]$ in (10) and (11) rather than the associated R-S data, $[\ln(i), \ln(s_{ij})]$, or G-I data, $[\ln(i-\frac{1}{2}), \ln(s_{ij})]$, we can employ simulated steady-state realizations from the Kesten process above. In particular, if these realized cities are randomly partitioned into a given number of subsets, then within the framework of Kesten processes, these subsets can be viewed as random samples of different sizes from the same statistical population of cities. This implies that their asymptotic power laws must be the same, and thus that the CPL property must in fact be *true* for these subsets.

To evaluate how well the CPL property is being captured by these three possible categorical regression approaches, we first simulate 1000 separate steady-state realizations of the Kesten process under $\kappa = 0.944$. For each realization, we then generate 1000 random 4-partitions of these 930 cities into disjoint subsets (subregions), $j = 1,2,3,4$, of fixed sizes (n_1, n_2, n_3, n_4). The specific subset sizes chosen for this analysis were $(n_1, n_2, n_3, n_4) = (182, 254, 261, 233)$ [which correspond to the four subregions shown in Figure 3(a) of Section 3 below].

By applying the three categorical regression procedures to each of these random partitions, one can compare how accurately each procedure captures the CPL property in terms of its mean estimate of $\hat{\kappa}$ of the common power exponent, $\kappa = 0.944$, (for this particular

[18] While similar measures could also be used here which reflect actual error magnitudes (such as mean absolute errors), *RMSE* is by far the most commonly used measure of model accuracy in nonparametric modeling. For recent illustrative applications in economics, see for example McMillen and Redfearn [39], Kitamura et al. [36], and Ait-Sahalia and Duarte [2].

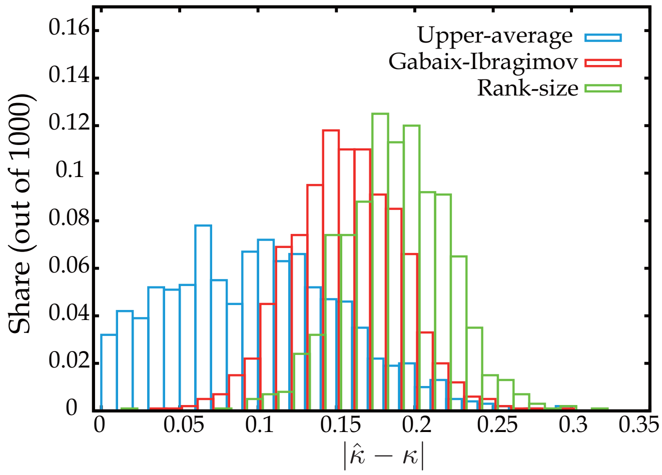

Figure 2: Comparison of categorical regression bias

steady-state realization and partition size). By determining the resulting *bias* of the mean estimate of $\hat{\kappa}$ from the true value κ for each of the 1000 steady-state realizations, one can then obtain frequency distributions of these values across the steady-state realizations for each procedure. The bias distributions for the U-A and R-S/G-I procedures are compared in Figure 2, where it is seen that the mean bias for the R-S, G-I, and U-A procedures are respectively 0.1891, 0.1564, and 0.0999. So the mean bias of the R-S/G-I procedure is about 90%/50% higher than that of the U-A procedure.[19] These results suggest that the upper-average approach continues to exhibit the best performance in this categorical regression setting.[20]

3 Voronoi Regions and the Common Power Law

Our first set of tests compare the CPL properties of random groupings of nearby cities versus purely random groupings. Here such random groupings of nearby cities are modeled as the cells of a *Voronoi K-partition* in which K reference cities are selected at random, and each cell, or *Voronoi region*, is defined by the set of cities closest to each reference city.[21] Let the number of cities in each cell, $j = 1, \ldots, K$, be denoted by n_j and the vector,

[19]Note that in principle one could also compare the overall fit of these three procedures in terms of their *mean RMSE* values. But since the underlying data sets are modified by these methods themselves, such *RMSE* values are not fully comparable. So while such a comparison again strongly favors the U-A procedure, these results are not reported here.

[20]Again, the basic results remain the same for alternative choices of partition sizes.

[21]Here "closeness" is defined in terms of travel distance between cities (CBSAs). More precisely, we use the shortest travel distances between the court houses of those counties contained in each CBSA.

$n(K) = (n_j : \ j = 1, \ldots, K)$, be designated as the *size* of the given partition. Then only random partitions of the same size will be comparable with this partition. So this size vector, $n(K)$, defines the key parameters governing the tests to be constructed. As one illustration of these parameters, Figure 3(a) displays an example of Voronoi K-partition with $K = 4$ and with $n(4) = (n_1, n_2, n_3, n_4) = (182, 254, 261, 233)$.

In this context, our basic *null hypothesis*, H_0, is that *the level of agreement of Voronoi partitions with the CPL property is statistically indistinguishable from that of similarly sized random partitions*. As in Section 2.3, this level of agreement is measured in terms of *RMSE* for the corresponding categorical regressions in expression (12) above.

For the Voronoi partition in Figure 3(a), the upper-average plots for these four Voronoi regions are shown in Figure 4(a), where the colors of each plot correspond to the partition colors in Figure 3(a).

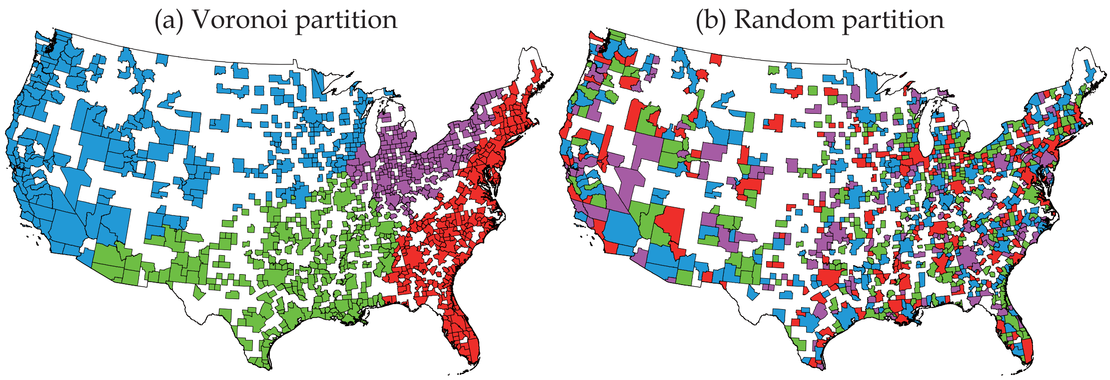

Figure 3: An example of Voronoi 4-partition

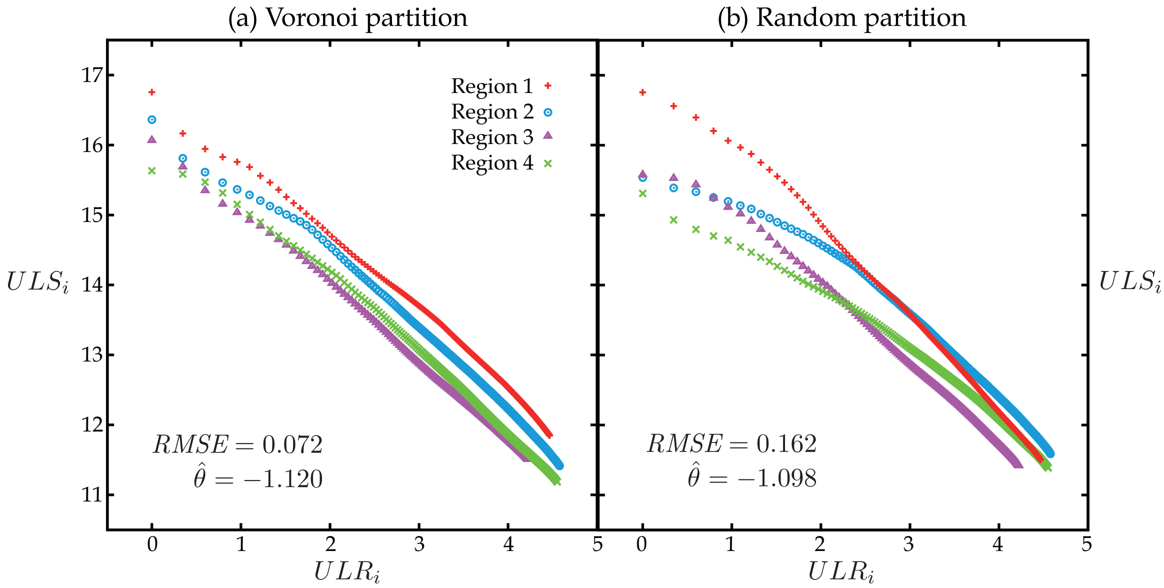

Figure 4: Upper-average distributions in Voronoi and random partitions

So to test the null hypothesis, H_0, for a given level of K, we proceed in two stages. First we generate $M = 1000$ random Voronoi K-partitions, $v = 1, \ldots, M$. Associated with each partition, v, is a given size vector, $n_v(K)$. So to estimate the distribution of *RMSE* for random partitions of size $n_v(K)$ under H_0, we generate 1000 random partitions of size $n_v(K)$ and calculate *RMSE* for each. For the Voronoi 4-partition shown in Figure 3(a), an example of random partition of the same size is shown in Figure 3(b), with corresponding upper-average plots shown in Figure 4(b). As can be seen from this figure, the upper-average plots differ from the Voronoi partition case mainly in the extreme upper tail, corresponding to the largest cities. In particular, the four largest cities (New York, Los Angeles, Chicago and Dallas) are contained in separate cells in the Voronoi partition shown in Figure 3(a), while for the random partition shown in Figure 3(b), all the four cities are contained in a single cell (the red region in the figure). As we shall see below, the locations of these largest cities within a given partition play a crucial role in determining its agreement with the CPL property.

If the *RMSE* level for partition v is denoted by $RMSE_v$, and if the number of *RMSE* values smaller than $RMSE_v$ is denoted by M_v, then the *p-value* for a one-sided nonparametric test of H_0 for partition v is given by[22]

$$p_v = \frac{M_v}{N} \ , \quad v = 1, \ldots, M. \tag{14}$$

[22]To be more precise, p_v, estimates the probability of achieving an *RMSE* level as low as $RMSE_v$ *if* it were true that partition v was in fact a typical random partition of size $n_v(K)$.

For the Voronoi partition, v, with upper-average plots in Figure 4(a), the *RMSE* value is $RMSE_v = 0.072$, and for the same sized random partition in Figure 4(b), the value is $RMSE = 0.162$. So this random partition exhibits less agreement with CPL than does partition v. In fact, for this particular Voronoi partition, *none* of the *RMSE* values for the corresponding 1000 random partitions fell below 0.072. So $p_v = 0$ for this extreme case.

The results of these tests of H_0 over the range of values, $K = 2, \ldots, 20$, are shown in Figure 5. Here the possible K values for Voronoi partitions are on the horizontal axis, and the possible p-values for these tests are on the vertical axis. To interpret these results, let us focus on the vertical slice at $K = 4$ in Figure 5. Recall that there are 1000 p-values for $K = 4$, one for each of the Voronoi 4-partitions generated. Among this population of p-values, the median value (on the red curve) is about 0.30, indicating that 50% of these p-values are below 0.30. But if Voronoi partitions were indistinguishable from random partitions as hypothesized under H_0, then one would expect that only 30% of these p-values would be at or below 0.30. More generally one can see from the corresponding percentile points at the 5%, 10%, 50%, 90% and 95% levels that this distribution of p-values is *uniformly below* what would be expected under H_0. The most interesting case of course involves p-values at or below 0.05 or 0.10. In this case 8.5% are below 0.05 and 17.5% are below 0.10. Thus in both cases, there are 75% more "significant" results than would be expected under H_0. So for the case of $K = 4$, there is substantial evidence suggesting that these Voronoi regions are exhibiting power laws that are more similar to one another than would be expected for random regions of comparable sizes. Figure 5 shows that the results become even more significant for larger values of K.

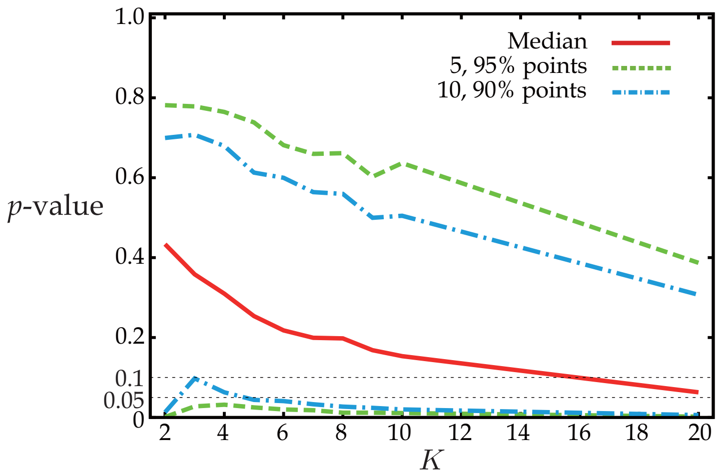

Figure 5: Comparison of *RMSE*s between Voronoi and random partitions

One key distinction between Voronoi and random partitions contributing to these results is that the largest cities tend to be more separated by Voronoi partitions than random partitions. This separation property will be established more formally in Section 4.1 below. But for the present, the relation between CPL properties and separation of large cities can be illustrated by focusing on the single significance level, $\alpha = 0.05$, in Figure 5. If for each K we denote the set of (simulated) Voronoi K-partitions that are significant at this α level by $V_\alpha^K = \{v : p_v < \alpha\}$, then we can measure the degree of large-city separation in these partitions as follows. For each partition, $v \in V_\alpha^K$, let $N_r^K(v)$ denote the number of cells in partition v containing at least one of the top r cities [so that $N_2^K(v)$ is the number of cells in v containing either New York or Los Angeles]. Finally, if $\overline{N}_r^K(\alpha)$ denotes the average of these values over V_α^K [so that $1 \leq \overline{N}_r^K(\alpha) \leq r$], then the fraction, $\overline{N}_r^K(\alpha)/r$, can be viewed as measuring the *degree of separation* of the top r cities in V_α^K. These degrees of separation are plotted over a range of K values for $r = 2, 3, 4$ in Figure 6. So at $K = 4$, for example, the degree of separation for $r = 2$ is seen to be 1.0, indicating that *every* partition significant at the $\alpha = 0.05$ level (i.e., in V_α^4) places New York and Los Angeles in different cells. Similarly, the degree of separation for $r = 4$, namely $0.78 \approx 3/4$, indicates that the top four cities are typically split among three of the four cells in these partitions. What is most important for our present purposes is that these degrees of separation for $r = 4$ exhibit a sharp increase from $K = 3$ to $K = 4$, and continue to increase for larger K. This echoes the decreasing contour for $\alpha = 0.05$ in Figure 5, and shows that the most significant Voronoi partitions with respect to CPL (at this α level) are indeed those achieving greater separation, i.e., with these four cities almost always completely separated.[23]

[23]For completeness, it should also be noted that for $K = 2, 3$ the stronger significance of these partitions again depends largely on the patterns of separation between the top four cities. For $K = 2$, it can be verified by closer examination of the partitions in V_α^2 [and can also be seen roughly from Figure 1(a)] that separations in which New York is in one cell and (Los Angeles, Chicago, Dallas) are in the other will tend to yield very similar upper-average curves for a considerable range of different Voronoi 2-partitions.

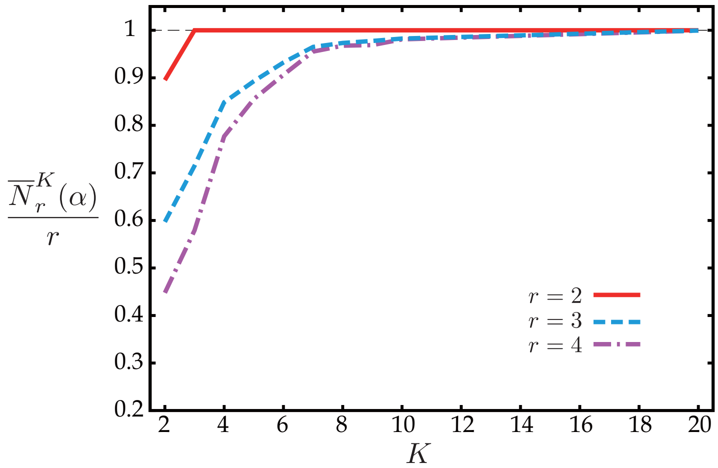

Figure 6: Voronoi Separation of Major Cities

One can gain further insight here by considering the full range of city sizes. In particular, since small cities are not only more numerous but also more ubiquitous, they tend to be evenly distributed across cells in both Voronoi partitions and random partitions. Moreover, since they exhibit less variation in size, one can expect differences in size distributions across cells to be most sensitive to the placement of the largest cities. Finally, since power laws focus on the upper tails of these distributions (as reflected by our upper-averaging procedure), one can expect that the more even spread of large cities across Voronoi-partition cells will lead to more similar power laws than for random partitions.

This leads naturally to the question of *why* the largest cities should be more evenly spread among the cells of Voronoi partitions. Here the most compelling reason seems to be that these cities are in fact more widely separated in space, i.e., are consistent with the first tenet of the *spacing-out property*. If so, then given the relative ubiquity of possible reference cities throughout the US, Voronoi partitions would then seem more likely to separate these largest cities than would random partitions of the same size. Such relations are most evident for the four largest cities (New York, Los Angeles, Chicago and Dallas) where Voronoi separation was evident in Figure 6 and where spatial separation is equally evident in Figure 7. A more detailed analysis of these relations is given in the next section, where tests are developed for both tenets of the spacing-out property.

4 The Spacing-Out Property of Cities

While Figure 6 suggests that those Voronoi partitions most consistent with the CPL property tend to separate the largest cities of a country, there remains the question of whether such separation is exhibited by all Voronoi partitions. If so, then as suggested above, this would provide strong evidence for the first tenet of the spacing-out property. In Section 4.1 below we develop a testing procedure that confirms the presence of such separation quite independently from any considerations of the CPL property.

In addition, we show the spacing-out property also asserts that smaller cities tend to be clustered around these larger centers. In Section 4.2 below we show that Voronoi partitions generated by the largest cities do indeed exhibit significantly stronger accessibility to the smaller cities in their cells than do Voronoi partitions generated by random cities. These results thus provide further support for the spacing-out property itself.

4.1 Spatial Separation of the Largest Cities

Let U denote the relevant set of cities for a given country (so that $|U| = 930$ for the case of US). For any given number, r, of the largest cities in U, and for any partition, v, of U, let $N_r(v)$ denote the number of partition cells of v containing at least one of these r cities. If there is indeed substantial spacing between the largest cities in U, then we would expect $N_r(v)$ to be larger for Voronoi partitions than for random partitions of the same size. For given values of r and K, we start by simulating M $(= 1000)$ Voronoi K-partitions, $v = 1, \ldots, M$, as in Section 3, and summarize the above counts, $N_r(v)$, by the *Voronoi count vector*,

$$N_r = [N_r(v) \, : \, v = 1, \ldots, M] \, . \tag{15}$$

For each of these Voronoi K-partitions, v, we again simulate M $(= 1000)$ random K-partitions, $\omega = 1, \ldots, M$, of the same size, $n_v(K)$. But rather than conducting separate tests for each Voronoi partition, v, as in Section 3, we now construct a summary test using appropriate mean values as follows.

First we write the random partitions for v as ordered pairs (v, ω), $\omega = 1, \ldots, M$, to indicate their size-dependency on v. In a manner paralleling $N_r(v)$, we then let $N_r(v, \omega)$ denote the number of cells in random partition (v, ω) that contain at least one of the r largest cities in U. In these terms the count vectors,

$$N_r(\omega) = [N_r(v, \omega) \, : \, v = 1, \ldots, M] \, , \quad \omega = 1, \ldots, M \tag{16}$$

can each be regarded as random-partition versions of the Voronoi count vector in (15), where each component, $N_r(v, \omega)$, of $N_r(\omega)$ is based on a random partition of the same size as Voronoi partition, v. In this setting, our basic null hypothesis is essentially that the Voronoi count vector, N_r is drawn from the same population as its random-partition versions in (16). But for operational simplicity, we focus only on the associated *mean-counts*, defined for (15) and (16), respectively, by

$$\overline{N}_r = \frac{1}{M}\sum_{v=1}^{M} N_r(v) \tag{17}$$

and

$$\overline{N}_r(\omega) = \frac{1}{M}\sum_{v=1}^{M} N_r(v,\omega) \ , \quad \omega = 1,\ldots,M\,. \tag{18}$$

In these terms, our explicit *null hypothesis*, H_0, is that *the Voronoi mean-count, $\overline{N}_r$, is drawn from the same population as its associated random mean-counts, $\overline{N}_r(\omega)$, $\omega = 1,\ldots,M$.*[24] If for the given set of simulated random partitions above, we now let M_0 denote the number of random mean-counts, $\overline{N}_r(\omega)$, larger than $\overline{N}_r$, then the *p-value*, p_0, for a one-sided test of H_0 is given [in a manner similar to (14)] by

$$p_0 = \frac{M_0}{M}\,. \tag{19}$$

The p-values, p_0, for such tests using the US data are given in Table 1 for the selected numbers of largest cities, $r = 2,\ldots,10$, and partition sizes, $K = 2,\ldots,10,20$. Here significance levels, $p_0 \leq 0.01$, $p_0 \leq 0.05$, $p_0 \leq 0.10$, are denoted respectively by ●, ◐ and ○, with blanks denoting no significance.

	K									
r	2	3	4	5	6	7	8	9	10	20
2	●	●	●	●	●	●	●	●	●	●
3	●	●	●	●	●	●	●	●	●	●
4	●	●	●	●	●	●	●	●	●	●
5		◐	●	●	●	●	◐	○		
6										
7		◐	●	●	●	●	●	●	●	
8				○	○					
9									◐	
10									○	

Table 1: The result of spatial separation test (US)

So for example, the symbol ● in cell $(2,2)$ signifies that the mean number of cells in Voronoi 2-partitions containing at least one of the 2 largest cities in the US is significantly

[24] As an alternative view of this hypothesis, observe that if one lets $N_r(v,0) = N_r(v)$ and considers the matrix of counts $C_r = [N_r(v,\omega)\ :\ v = 1,\ldots,M, \omega = 0,1,\ldots,M]$, then H_0 amounts to the hypothesis that the mean of the first column of C_r is drawn from the same population as the means of the remaining columns.

greater (at the 0.01 level) than would be expected if this were a random 2-partition. Moreover, since for $r = 2$, this significance level persists for all partition sizes up to $K = 20$, it is evident that for Voronoi 2-partitions these two largest cities (New York and Los Angeles) are almost never in the same cell of any such partition. This is hardly surprising, since New York and Los Angeles are on opposite coasts. So the key point here is that random partitions are not sensitive to "opposite coasts", while Voronoi partitions most certainly are. More generally this same degree of maximal significance is seen to persist up to the first four largest cities (New York, Los Angeles, Chicago, and Dallas), which we have already seen are spaced widely apart within the US. But when the fifth largest city (Philadelphia) is included, its close proximity to New York makes such separation less likely. Moreover, since the sixth largest city (Houston) is also close to Dallas, the significance of Voronoi separation now disappears altogether. What is more interesting is the apparent resurgence of significance when the seventh largest city (Miami) is included. Here again it is evident from the map in Figure 7 that Miami is about as far away from the six largest cities as is physically possible within continental US.

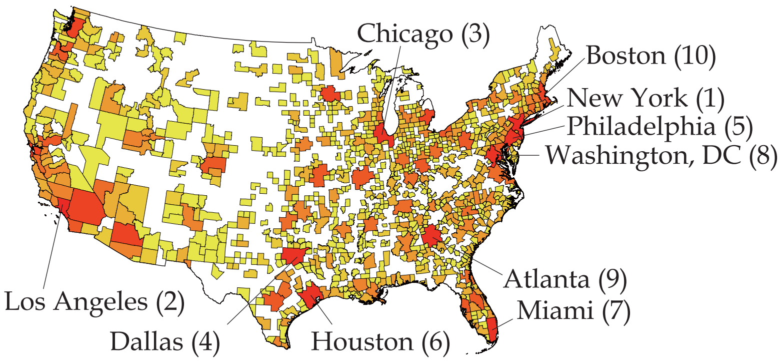

Figure 7: Locations of Cities

So again, this separation effect is strongly captured by our testing procedure. In summary, these results do indeed support the spacing-out property of largest cities within the US, and in particular, they echo the strong separation of the four largest cities seen in the tests of Section 3 above. Notice also that the spacing between these largest four cities is somewhat more uniform than the spacing between smaller cities. This is in part explained by the the tendency of smaller cities to cluster around larger cities, as we examine further in the next section.

4.2 Concentration of Smaller Cities Around the Largest Cities

We now focus on the spatial distribution of smaller cities associated with that of the largest cities studied in the previous section. For this purpose, we designate the (unique) Voronoi K-partition generated by the K largest cites as the *largest-city Voronoi K-partition*. Our objective is then to test whether these K largest cities are significantly more *accessible* to all other cities in their cells than are the corresponding reference cities in random Voronoi K-partitions (as in Section 3 above).

To formalize these concepts, we first identify the sets of cities in each partition cell. For any Voronoi K-partition, let the set of all cities in each cell, $i = 1, \ldots, K$, be denoted by U_i $(\subset U)$, and let $u_i \in U_i$ denote the *reference city* in this cell. If the *distance* from u_i to any city $u \in U_i$ is denoted by $d(u_i, u)$,[25] then the *total distance* of all cities in U to their reference cities in a given Voronoi K-partition is then given by

$$D_K \equiv \sum_{i=1}^{K} \sum_{u \in U_i} d(u_i, u) \; . \tag{20}$$

With these definitions, if the largest K cities do indeed serve as cluster centers for those smaller cities around them, then one should expect to observe values of D_K for largest-city Voronoi K-partitions that are smaller than the corresponding values, say $\widetilde{D}_K$, for similarly sized random Voronoi K-partitions. To test this assertion for a given value of $K = 2, 3, \ldots$, the appropriate *null hypothesis*, H_0, is simply that *D_K and $\widetilde{D}_K$ come from the same statistical population*. By using the 1000 samples of random Voronoi K-partitions as in the previous section, we can then compute the appropriate p-value for a one-sided the test of H_0 for this value of K.

Alternatively, it may be more appropriate to use accessibilities to city populations by weighting each distance, $d(u_i, u)$, in eq. (20) by the *population size*, s_u, of city $u \in U_i$. Note however that since $d(u_i, u_i) = 0$ for each reference city, u_i, the populations of the K largest cities will automatically be excluded from the largest-city Voronoi K-partition. But for random K-partitions, where these largest cities are generally not the reference cities, these largest populations will often be included in total population-weighted distances. Thus in order to focus on comparisons of accessibility to populations in *smaller* cities, it is appropriate to exclude the K largest city populations from all such comparisons.[26] To do

[25] Recall that our measure of distance, $d(u, u')$, between cities u and u' was defined in footnote 21 above. Note in particular that this (set) distance implies that the distance from any city to itself is zero, i.e., that $d(u, u) = 0$.

[26] As will become clear below, this convention has the additional advantage of yielding a *conservative* test of clustering around the largest cities. In particular, the inclusion of largest-city populations must necessarily

so, if we now denote the set of K largest cities in U by U^K, then for any given largest-city Voronoi K-partition, the appropriate modification of D_K above is now taken to be the *total population-weighted distance* as defined by,

$$D_K^* \equiv \sum_{i=1}^{K} \sum_{u \in U_i - U^K} s_u d(u_i, u) \,. \tag{21}$$

If the total population-weighted distance for a random K-partition is similarly denoted by $\widetilde{D}_K^*$, then the appropriate *null hypothesis*, H_0^*, for this alternative test is now that D_K^* *and* $\widetilde{D}_K^*$ *come from the same statistical population.*

Since the largest-city Voronoi K-partition is unique for each K, the hypotheses, H_0 and H_0^*, are tested by simulating 1000 random Voronoi K-partitions and calculating appropriate p-values (for one-sided tests) as the share of associated total distance values, $\widetilde{D}_K < D_K$, under H_0, and the share of total population-weighted distance values, $\widetilde{D}_K^* < D_K^*$, under H_0^*, respectively. The results of these tests are plotted in Figure 8 for $K = 1, \ldots, 20$. Turning first to H_0 (plotted in red), the significance results for $K = 3$ and 4 reflect the strong tendency in Voronoi separation for $r = 3$ and 4 in Table 1. Note that high p-value at $K = 2$ is expected. For since the largest two cities (New York and Los Angeles) are located on opposite coasts, random pairs of reference cities will almost always have better overall access to cities than these two. The subsequent rise in p-values at $K = 5$ and 6 echoes the spatial separation results for $r = 5$ and 6 in Table 1. In particular, given the respective closeness of Philadelphia to New York and Houston to Dallas, the addition of each of these reference cities yields only a small increase in overall accessibility relative to randomly chosen reference cities. Similarly, the improvement in accessibility when Miami is added ($K = 7$), and deterioration when Washington, D.C. is added ($K = 8$) also reflect the cases of $r = 7$ and 8 in Table 1. But overall, there is a discernible tendency of cities to exhibit more clustering around the largest cities than around randomly selected reference cities.

This tendency is much more dramatic when population accessibilities are compared. As shown by the blue curve in Figure 8, these results are *uniformly more significant* than for the case of simple inter-city distances. Indeed, except for the "bi-coastal" case ($K = 2$) and the "Houston next to Dallas" case ($K = 6$), these results are all strongly significant ($p \ll .05$). Thus, the single most important conclusion here is that relative to randomly selected reference cities, *the largest cities in the US tend to exhibit significantly better access to their surrounding city populations.*

In relation to the results in Section 3 above, the spatial relations between smaller and

increase the total population-weighted distances for almost all random K-partitions.

larger cities studied here show that the city subsets around the three or four largest cities are roughly comparable to one another, each consisting of similarly sized cities. This in part suggests why Voronoi partitions tend to exhibit higher consistency with the CPL property than their random counterparts. In particular, those Voronoi cells containing the largest cities tend also to contain substantial portions of their corresponding city clusters.

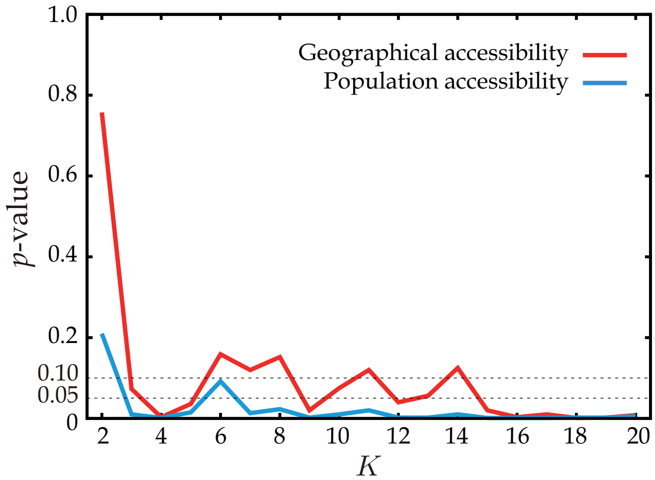

Figure 8: Result of total-accessibility test

5 Economic Regions and the Common Power Law

Our final objective is to determine whether the CPL property is stronger when comparing more economically meaningful regions. As mentioned in the Introduction, we here replace simple distance proximities by commodity flow dependencies. Such dependencies are based on the Commodity Flow Survey (CFS) for 2007. This data identifies total shipments between 111 regions in the continental US, as defined by the CFS. In particular, 64 of these regions are CFS-defined metropolitan areas, and the remaining 47 regions are either states that do not overlap these metropolitan areas or "remainder of the state" regions including those part of states outside the metro areas.[27] Each CFS metropolitan area is either an individual CBSA, or a Combined Statistical Area (CSA) consisting of multiple CBSAs.[28]

[27] The 47 regions correspond to the continental states excluding Rhode Island as it is completely contained in a CFS-defined metro area, Boston-Worcester-Manchester.

[28] However, there is one case in which a single CBSA (Washington-Arlington-Alexandria) has been divided into two CFS metropolitan areas [designated, respectively, as the Washington-Arlington-Alexandria CBSA and the Washington-Baltimore-Northern Virginia CSA (Virginia part)]. In order to reconcile this CFS data with the set of CBSAs defining "cities" in the present paper, we have thus aggregated these two CFS areas into a single Washington-Arlington-Alexandria metropolitan area (consisting of three CBSAs, Washington-Arlington-Alexandria, Winchester and Culpeper). For the com-

We start in Section 5.1 below by constructing an operational definition of *economic regions* in terms of these commodity flow dependencies. In Section 5.2, we then test the significance of the CPL property for these economic regions against comparable sets of random partitions. In Section 5.3 these CPL test results are shown to be even stronger than comparable results for the Voronoi partitions in Section 3. Finally in Section 5.4, we develop an alternative method for comparing differences in upper-average distributions between economic regions and between their corresponding random partitions. In particular, we construct a new measure of similarity of between upper-average distributions in terms of the order-consistency properties of their ULS_i levels across ranks, i. Here it is shown that this measure can in many cases provide even sharper comparisons between power laws across regions.

5.1 Economic Regions

If R denotes the set of all CFS regions, $i = 1, \ldots, 111$, we first identify each region, $i \in R$, with its associated set of cities as follows. Let the set of all cities, U, be partitioned into cells, $\{U_i : i \in R\}$, so that $u \in U_i$ if and only if region i accounts for the largest population share of city u. In the analysis to follow we refer to U_i as the set of cities for region i. For convenience we then order regions in terms of their largest cities, so that by again letting s_u denote the size of city u it follows that regions, $i, j \in R$, will satisfy $i < j$ if and only if $\max_{u \in U_i} s_u > \max_{u \in U_j} s_u$. Thus the first K regions will generally be associated with the K largest cities in U.[29] For each K $(= 1, 2, \ldots)$, the desired sets of K economic regions then correspond essentially to the largest-city regions together with their associated economic hinterlands.

These ideas can be made more precise terms of commodity-flow dependencies as follows. If for any regions, $i, j \in R$, we let f_{ij} denote the *commodity flow* (in dollar value) from region i to region j, then the (*commodity*) *flow dependency*, $\lambda_{ij} \in [0, 1)$, of region j on region i is taken to be the fraction of the total commodity-inflow to j that comes from i, i.e.,

$$\lambda_{ij} \equiv \frac{f_{ij}}{\sum_{k \in R} f_{kj}} \quad , \tag{22}$$

plete list of the CFS metropolitan areas, refer to the website of the US Department of Transportation: http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/2007/ metropolitan_areas/index.html.

[29] In particular, the largest cities in the first $K = 20$ regions (used in the analysis below) match the largest 20 cities in U, with the two exceptions of Los Angeles-Long Beach-Santa Ana (second largest) and Riverside-San Bernardino-Ontario (14th largest) that belong to the same CFS region.

where in particular, λ_{jj} is designated as the *self-flow dependency* of region j. For any given set of K "central" regions, one can then generate appropriate economic hinterlands by simply assigning every other region in R to its largest supplier among these K regions. But the definition of "central" regions themselves is more subtle. Here it might seem natural to simply choose the first K regions, i.e., with the largest cities. But this ignores the relative flow dependencies among these regions. For example, while Philadelphia is the fifth largest city, it exhibits a strong flow dependency on New York $(\lambda_{1,5} = 0.119)$. This suggests that in central-region systems with $K \geq 5$, it might be more appropriate to treat Philadelphia as part of the New York hinterland. More generally, the notion of "centrality" itself appears to involve a tradeoff between flow dependencies and largest-city sizes. To make this tradeoff explicit, we now parameterize possible collections of K central regions in terms of the maximum allowable flow dependency between any pair of central regions, designated as their *threshold-dependency level*, $\lambda \in (0, 1)$. For any given values of λ and K, we then define the appropriate set of *central regions*, $R_{\lambda,K}$, to be the first K regions, $j \in R$, with no flow dependencies on larger regions that are higher than either λ or their own self-flow dependency, λ_{jj}.[30,31] To be more precise, if we now let $R_j^* = \{i \in R : i < j\}$ denote the set of regions with larger maximum city size than region, j, then membership in $R_{\lambda,K} \equiv \{j_m : m = 1, \ldots, K\} \subset R$ is defined by $j_1 = 1$ and for all $m = 2, \ldots, K$ by

$$j_m = \arg\min \left\{ j > j_{m-1} : \max_{i \in R_j^*} \lambda_{ij} \leq \min\{\lambda, \lambda_{jj}\} \right\} . \tag{23}$$

In essense, central regions, $R_{\lambda,K}$, constitute the set of K largest regions exhibiting no mutual flow dependencies stronger than λ. Note however that parameters, K and λ, are by no means independent. In particular, for sufficiently small values of λ, only $K = 1$ is possible, i.e., the entire country is in the economic hinterland of New York. However,

[30] Note that one could in principle require the first condition to hold for *all* other regions rather than simply larger regions. However, there are exceptions where smaller regions are the largest suppliers of larger regions, especially when the smaller region is a major transshipment point (such as a port) or a border region. The most important instance for our purposes is Houston ($j = 6$), which is a major supplier of Dallas ($j = 4$). For example at the $\lambda = 0.05$ level, it can be seen from Table 3 in the Appendix that $\lambda_{4,6} = 0.036 < \lambda < 0.055 = \lambda_{6,4}$, which would exclude Dallas as a central region for this level of λ. So to avoid such exceptional cases, we apply this condition only to larger regions. But it should also be noted here that such difficulties are in part due to the fact that CFS data does not distinguish transshipment points from origin and destination points, thus tending to overestimate outflows originating at transshipment points.

[31] Note also that while the second condition is reasonable, it is actually only binding for one CFS region. In particular, San Diego-Carlsbad-San Marcos imports 36.4% from Los Angeles-Long Beach-Riverside, while its domestic supply (self-flow dependency) accounts for only 29.9%. But since San Diego-Carlsbad-San Marcos hardly constitutes an economic center comparable to Los Angeles-Long Beach-Riverside, this creates no problem for the present analysis.

for first 20 regions in R considered in the present analysis, all relevant numbers of central regions, $2 \leq K \leq 20$, are possible for threshold-dependency levels, $\lambda \geq 0.05$. Finally, it should be clear that even in this most relevant range, the set of central regions, $R_{\lambda,K}$, can be quite different from the first K regions in R. These differences are of course most dramatic for small λ. In the case of $\lambda = 0.05$, for example, four of the ten largest regions (Philadelphia, Miami, Washington DC, and Boston) are all excluded by their strong flow dependencies on New York. Additional details and examples can be found in the Appendix, where all flow dependencies among the first 20 regions in R are depicted in Table 3.

Given this definition of central regions, $R_{\lambda,K}$, we can now define the associated system of *economic regions*, $E_{\lambda,K}$, as follows. For each central region, $j \in R_{\lambda,K}$, let the corresponding economic region, E_j, consist of all regions in R for which region j is the largest supplier, i.e.,

$$E_j \equiv \left\{ i \in R \; : \; j = \arg\max_{r \in R_{\lambda,K}} \lambda_{ri} \right\} . \tag{24}$$

This automatically generates a K-partition of U under threshold-dependency level, λ,

$$E_{\lambda,K} \equiv \left\{ E_j : \; j \in R_{\lambda,K} \right\}, \tag{25}$$

which we now designate as the *economic-region K-partition* for λ.

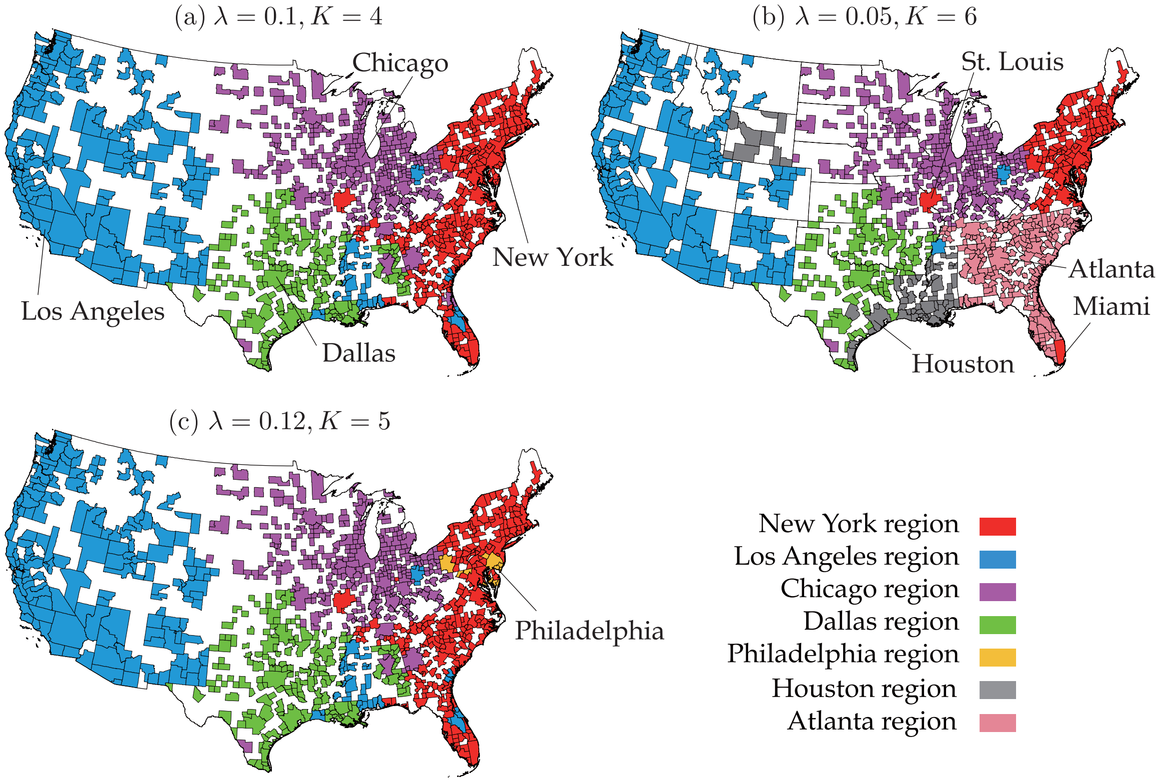

Figure 9: Economic regions for selected values of λ and K

These economic-region partitions are illustrated by the examples in Figure 9 for selected combinations of λ and K. Here each colored cell represents the geographical coverage of a single city in U. Those cities of the same color all belong to a single economic region. In particular, panel (a) shows the economic-region 4-partition, $E_{0.1,4}$, that happens to be the same for all $\lambda \in [0.037, 1.0]$. Here the corresponding central regions, $R_{\lambda,4}$, consist of the four largest cities (New York, Los Angeles, Chicago and Dallas). Notice the strong resemblance between these four economic regions and Voronoi 4-partition in Figure 3(a).

5.2 Test of the CPL Property

To test the significance of the CPL property for the economic regions, we constructed economic-region partitions, $E_{\lambda,K}$, for selected values of (λ, K), and (as in Section 3) generated 1000 random partitions of similar sizes for each case.[32] The relevant range of significant results are shown in Table 2, where each cell contains the p-value (p_0) for a

[32]The values of λ used were (i) 0.05 to 0.15 in increments of 0.01, and (ii) 0.20 to 1.00 in increments of 0.10. The values of K used were $K = 2, \ldots, 20$. Note also that while it is possible to consider values $\lambda < 0.05$ in some cases, these flow dependency thresholds are so low that the resulting economic regions tend exhibit little spatial cohesion whatsoever.

one-sided test of H_0 given that particular (λ, K) pair.[33] As in Table 1, significance levels, $p_0 \leq 0.01$, $p_0 \leq 0.05$, $p_0 \leq 0.10$, are denoted respectively by ●, ◐ and ○, with blanks denoting no significance. To interpret these results, we first note that since the four largest cities are highly independent of one another in terms of commodity flows (as shown in Table 3 of the Appendix), the economic-region partitions, $E_{\lambda,K}$, are the same for all $\lambda \geq 0.05$ when $K \leq 4$.[34] Thus the test results shown in the first three columns continue to hold for all $\lambda \geq 0.12$, as indicated in the table. In particular, the economic-region partitions for both $K = 3$ and 4 are significantly more consistent with the CPL property than random partitions regardless of mutual flow-dependency considerations.[35] (The insignificance of CPL for the $K = 2$ case will be discussed in Section 5.3 below.)

To examine these results in more detail, it is instructive to compare the upper-average distributions of economic-region partitions with representative random partitions of the same size. The upper-average distributions of $E_{\lambda,K}$ for $K = 4$ (and all $\lambda \geq 0.05$) are shown in panel (a) of Figure 10. To represent random partitions of the same size, we use the random partition with *median RMSE* value, as shown in panel (b). As in the Voronoi 4-partition example of Figure 4 in Section 3, the distinction between observed and random partitions is again seen to be most pronounced in the upper tails of the distributions. Note in particular that for the largest 100 cities (i.e., up to $ULR_i = 3.64$), the upper-average distributions for these four economic regions do not cross one another, while many such crossings occur in the corresponding random partition.[36]

---

[33] The only significant cases not shown (namely with $\lambda = 0.06, 0.07$ and $K \geq 18$), all include single-city regions for which power laws are not meaningful.

[34] To be more precise, $R_{\lambda,4}$ consists of the largest four cities for all $\lambda > 0.037$, while Dallas will be contained in Los Angeles region for $\lambda \leq 0.037$. Similarly, $R_{\lambda,3}$ consists of the largest three cities for all $\lambda > 0.029$, while Chicago will be contained in New York region for $\lambda \leq 0.029$. Refer to Table 3 in the Appendix for these threshold levels of λ.

[35] Notice that even under different values of threshold-dependency levels, say, λ and λ', the economic-region partitions are identical, i.e., $E_{\lambda,K} = E_{\lambda',K}$, if the set of central regions are identical, i.e., $R_{\lambda,K} = R_{\lambda',K}$.

[36] Such comparisons will be made more explicit in Section 5.4 below.

λ	K = 2	3	4	5	6	7	≥ 8
0.05		◐	●	◐			
0.06		◐	●	◐	●		
0.07		◐	●	◐	●		
0.08		◐	●	◐	●		
0.09		◐	●	◐	●	◐	
0.10		◐	●	◐	●	◐	
0.11		◐	●	◐	●	◐	
≥ 0.12		◐	●				

Table 2: Result of CPL test for economic regions

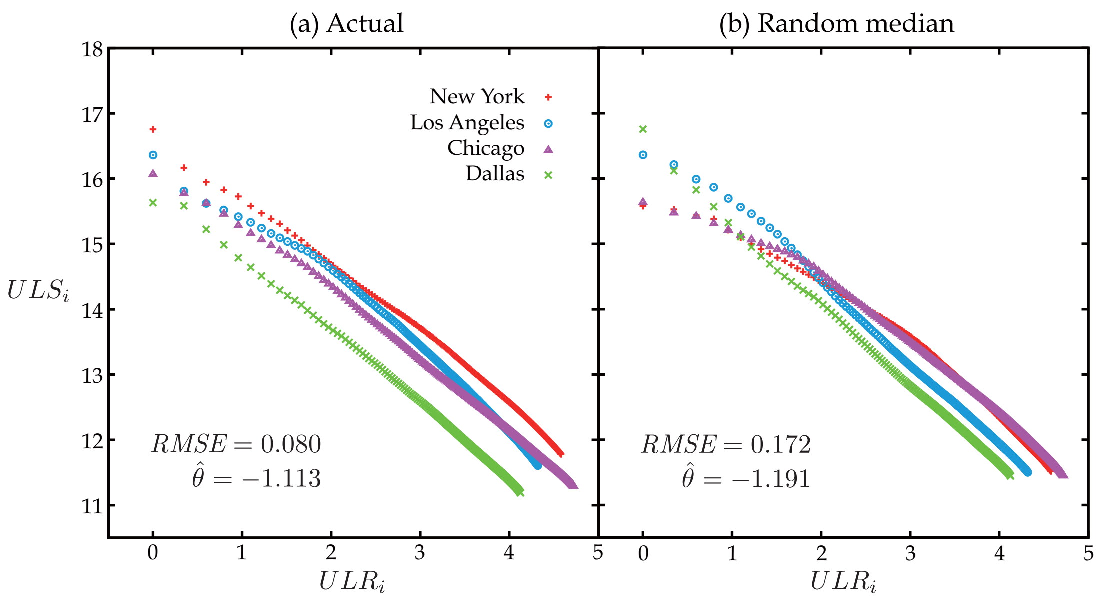

Figure 10: Upper-average distributions under $(\lambda, K) = (0.1, 4)$

For $K \geq 5$, the CPL property continues to be significant until either Philadelphia (the 5th largest) or Atlanta (the 9th largest) is added as a central region. As mentioned above, Philadelphia belongs to the hinterland of New York when $\lambda \leq 0.119$, but forms its own economic region at all higher values of λ. Similarly, Atlanta belongs to the Chicago hinterland for $\lambda \leq 0.032$, but forms its own economic region at all higher levels. The presence or absence of these two cities appear to be the major factors governing the pattern of significance levels for $K \geq 5$ in Table 2.

The "Atlanta effect" can be illustrated by the case, $(\lambda, K) = (0.05, 6)$, shown in Figure 9(b), which is relatively close to the $K = 4$ case but no longer exhibits any significant consistency with the CPL property. To understand this dramatic difference, note first that the six central cities include the four largest cities together with Houston and Atlanta.[37] Here both Philadelphia and Miami now belong to the New York region, even though Philadelphia is larger than Houston, and Miami is larger than Atlanta. Notice also that the New York region for this case is about a half its size under $K = 4$ in Figure 9(a), where the southern half is now taken by Atlanta except for the isolated city of Miami. As a consequence, there are too few small cities in the economic region of New York to sustain the CPL property with other economic regions. As seen in Figure 11, this is reflected in the upper-average distribution of the New York region, which now exhibits strong concavity in the lower tail compared to that shown in Figure 10(a) for $(\lambda, K) = (0.1, 4)$.

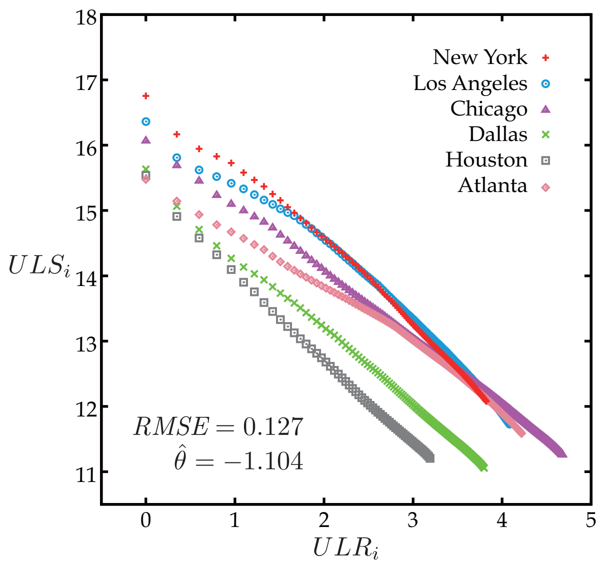

Figure 11: Upper-average distributions under $(\lambda, K) = (0.05, 6)$

Next, the "Philadelphia effect" is well illustrated by the case, $(\lambda, K) = (0.12, 5)$, in Figure 9(c). Notice that the coverage of the Philadelphia region is very limited.[38] As a consequence, the upper-average distribution of this region differs markedly from those of other regions, as seen in Figure 12. This in turn deteriorates the strength of the CPL property seen at lower levels of λ.

[37] To be more precise, the same 6-partition as depicted in Figure 9(b) is obtained for all $\lambda \in (0.037, 0.052]$. In particular, among the central cities in $R_{0.05,6}$, Dallas would belong to the hinterland of Los Angeles for $\lambda \leq 0.037$, while Miami (which is larger population size than Atlanta) belongs to the hinterland of New York for $\lambda \leq 0.052$, but would join $R_{\lambda,6}$ for $\lambda > 0.052$ in place of Atlanta.

[38] Notice also that the set of five central cities is identical for all $\lambda \geq 0.12$, since the threshold-dependency, λ, is relevant only for the selection of the central cities.

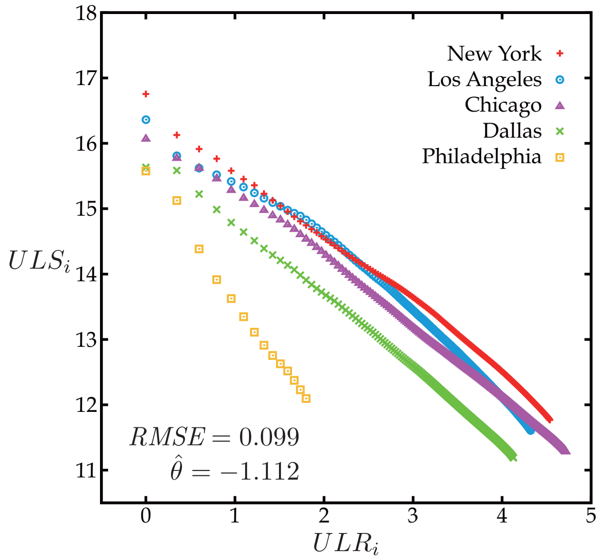

Figure 12: Upper-average distributions under $(\lambda, K) = (0.12, 5)$

Thus, while the case of $K = 4$ yields strong consistency with the CPL property across all values of λ , the above examples show that for larger values of K, the addition of economic regions with smaller central cities tends to increase the variation among upper-average distributions, leading to a deterioration of the CPL property. However, it should also be noted that this deterioration may in part be due to our partition-based definition of economic regions. In reality such regions tend to overlap, and may even form hierarchical relations. For example, rather than requiring Philadelphia to form a separate region at higher levels of commodity-flow dependency, it may be more appropriate to treat Philadelphia as a *subcenter* within the New York region. Along these lines, it has been shown by Akamatsu et al. [3] (using the same data as ours) that such hierarchical economic regions yield even stronger support for the CPL property (as discussed further in the Conclusions).

5.3 Comparison with Voronoi Partitions

It is also of interest to compare these CPL results with those obtained in Section 3 for the simpler case of Voronoi partitions. This comparison is shown in Figure 13, where the 5% and 10% bands for Voronoi partitions in Figure 5 are here reproduced (in green and blue, respectively). The significance levels (p-values) for economic-region partitions with $\lambda = 0.1$ and $K = 2, \ldots, 20$ are then superimposed (in red) on these bands (where the gray band can be ignored for the moment). Before comparing the more significant results, we begin by noting the conspicuous lack of significance for economic regions at the $K = 2$ level, in contrast to the Voronoi results. Closer examination of this case shows that while

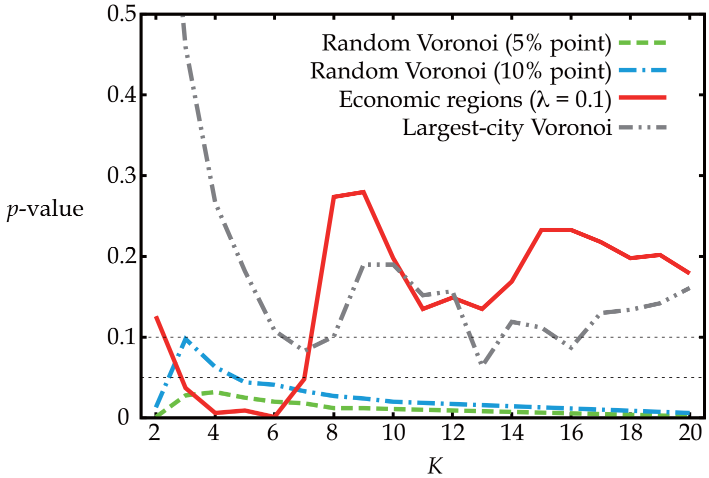

Figure 13: CPL Comparisons of Economic-Region and Voronoi Partitions

many Voronoi 2-partitions tend to split the country evenly between New York and Los Angeles (given the higher density of possible reference-city pairs on the two coasts), the economic-region 2-partition case is actually somewhat more random. In particular, those cities in the third major economic region, Chicago, are split quite randomly according to which flow dependency on New York or Los Angeles is the larger one, leading to less spatial coherence of these two regions. In contrast, the similarities between these three major regions are fully reflected at the $K = 3$ level, leading to results comparable to the top 5% of Voronoi cases. Moreover, for the most important cases of $K = 4$ through 6 (as discussed above) there is far stronger consistency with the CPL property than for comparable Voronoi cases. In particular, the p-values here are actually below the 1% band (not shown), indicating that the CPL property is more significant for these economic-region partitions than for 99% of the 1000 randomly generated Voronoi partitions.

However, one may ask whether these results might not be improved by using largest-city Voronoi partitions rather than random choices. So as one final comparison, the p-values for the K largest-city Voronoi partitions, $K = 2, \ldots, 20$, are shown by the gray band in Figure 13. Surprisingly these results are never more than weakly significant, and in addition, are nowhere close to the top 10% of the randomly generated Voronoi partitions. Recall from Figure 8 however that in terms of the spacing-out property, these largest-city Voronoi partitions exhibit significantly higher population accessibility (lower total population-weighted distances) than do random Voronoi partitions. Moreover, they also tend to exhibit higher population accessibility than economic-region partitions for any given K. In particular, when largest-city Voronoi partitions share the same reference cities as

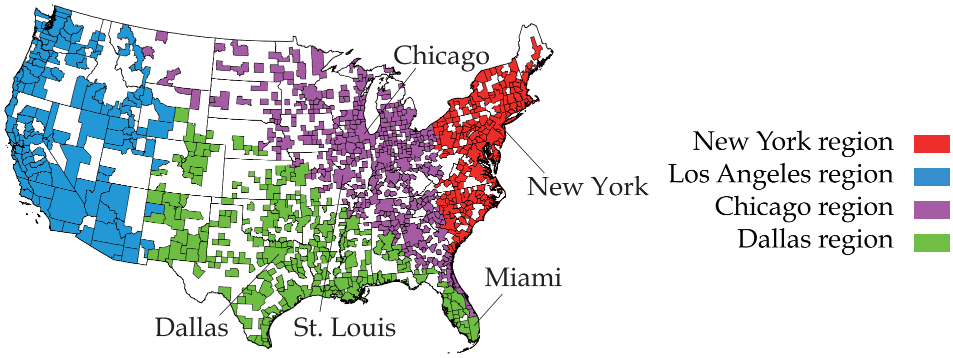

Figure 14: The largest-city Voronoi 4-partition

economic-region partitions (as for example when $\lambda = 0.1$ and $K \leq 4$) then by construction they *must* exhibit higher population accessibility.[39]

But given the stronger CPL properties of economic-region partitions, it is of interest to ask how these two types of partitions differ when their reference cities consist of the same largest cities. In general terms, since the cells of both partitions tend to include those cities closest to each largest city, the major differences are near the boundaries of each cell. A closer look at the data indicates that for those major cities relatively far from any of the K largest cities (i.e., those near the regional boundaries), their largest trade partners are often not the closest of the K largest cities. This can be illustrated for $K = 4$ by a comparison of the economic-region partition (for $\lambda = 0.1$) in Figure 9(a) with the largest-city Voronoi partition shown in Figure 14. For example, while Miami is closest to Dallas [i.e., is in the green region of Figure 14], its largest flow dependency is on New York [i.e., is in the red region of Figure 9(a)]. Similarly, St. Louis is closest to Chicago [i.e., is in the magenta region of Figure 14], but again with largest flow dependency on New York [i.e., is also in the red region of Figure 9(a)]. These observations suggest that in terms of the CPL property itself, perhaps a trade-linkage interpretation of the spacing-out property would be more appropriate than our present geographical version.

5.4 Order-Consistency of Upper-Average Distributions

Recall from the example in Figure 10 that a distinguishing feature of economic-region partitions exhibiting significant CPL properties is that their upper-average distributions appear to be more "parallel" than those of comparable random partitions. This effect can

[39] Since each non-reference city in a Voronoi partition must always be assigned to its closest reference city, it follows by definition that both distance sums in (20) and population-weighted distance sums in (21) are necessarily *minimal* with respect to the given set of reference cities.

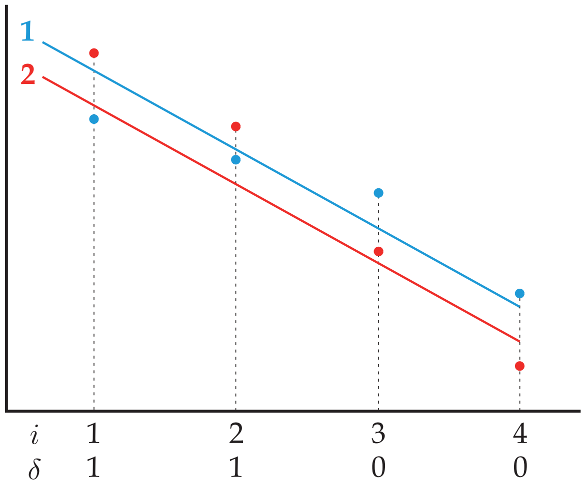

Figure 15: Categorical Regression Example

be described more formally in terms of the *order consistency* among these upper-average distributions, i.e., the order consistency among the ULS_i values at each rank value *i*. As we now show, this type of consistency is far more stable for economic-region partitions than for random partitions as shown below.

To do so, we begin by observing that if all upper-average curves for a given regional partition were perfectly parallel (and thus were exactly consistent with the CPL property) then the vertical ordering of ULS values at each rank value would necessarily agree with their common ordering predicted by the categorical regression used to test this CPL property. So the simplest way to measure overall order consistency is to compare these orderings at each rank value with the common predicted ordering. A stylized version of such comparisons is shown in Figure 15, where only two sets of "red" and "blue" cities are shown, each consisting of four cities (so that comparisons can be made at all four rank values, $i = 1, 2, 3, 4$). Here the red and blue lines correspond to the results of the categorical regression (which by construction yields that pair of parallel lines minimizing the overall sum of squared errors for both sets of cities). So the *ULS* values for blue cities are here predicted to be above those for red cities at every rank. This is seen to be true at ranks $i = 3, 4$, but not true at ranks $i = 1, 2$. So of the possible comparisons that can be made in this case, one can say that the degree of order consistency is 50%. However, for much larger examples, it is more convenient to focus on *inconsistencies*, which tend to be fewer in number and to exhibit wider relative variations. Thus in the present example, the *degree of order inconsistency* is also 50%. This is essentially the test statistic we seek to construct.

To formalize these ideas for a given number of partition cells, K, let the set of cities in each partition cell, $k = 1, \ldots, K$, be designated as the k^{th} *city set*, U_k, of *size* $|U_k|$ (in a manner paralleling Section 5.1 above). Thus in the example of Figure 15 there are two city sets of equal size, $|U_1| = 4 = |U_2|$. But more generally, these sets will be of different sizes. So for each city rank, i, the only city sets, U_k, for which this rank is meaningful are those for which $|U_k| \geq i$. If this collection of city sets is now denoted by $U_K(i) = \{k \in \{1, \ldots, K\} : |U_k| \geq i\}$, then the relevant ordering of ULS_i values involves only those city sets in $U_K(i)$. With this in mind, recall next from Section 2.3 that city sets, k, now constitute the appropriate "regions", j, in expression (12), so that ULS_{ik} denotes the relevant ULS_i value for each city set, $k \in U_K(i)$. In these terms the *i-rank*, $r_i(k)$, of city set, $k \in U_K(i)$, is given by the number of such city sets with ULS_i values no larger than that of k, i.e., by

$$r_i(k) = \left|\{\ell \in U_K(i) : ULS_{ik} \leq ULS_{i\ell}\}\right| \tag{26}$$

so that the city set with highest ULS_i value has rank one (assuming no ties). This in turn yields the desired *i-rank ordering*,

$$r_i(k) \leq r_i(\ell) \Leftrightarrow ULS_{ik} \geq ULS_{i\ell} \tag{27}$$

for all $k, \ell \in U_K(i)$ and $i = 1, \ldots, n \ (= |U|\,)$.

Given these definitions, our main objective is to compare each ordering in (27) with the common ordering generated by the categorical regression for partition, K. To do so, recall from expression (12) that if (for convenience) we now set $\beta_1 = \hat{\beta}_1 = 0$ for the reference region, $k = 1$, then the predicted values, $\widehat{ULS}_{ik}$, in this regression are given by

$$\widehat{ULS}_{ik} = \hat{\alpha} + \hat{\theta}\, ULR_i + \sum_{h=1}^{m} \hat{\beta}_h D_j(h) = (\hat{\alpha} + \hat{\theta}\, ULR_i\,) + \hat{\beta}_k\,. \tag{28}$$

Thus the ordering of these predicted values reduces to the identity,

$$\widehat{ULS}_{ik} \geq \widehat{ULS}_{i\ell} \Leftrightarrow \hat{\beta}_k \geq \hat{\beta}_\ell \tag{29}$$

which is seen to be *independent* of i. This in turn yields *regression ranks*,

$$\hat{r}(k) = |\{\ell = 1, \ldots, K : \hat{\beta}_k \leq \hat{\beta}_\ell\}| \tag{30}$$

which generate the desired *regression-rank ordering*,

$$\hat{r}(k) \leq \hat{r}(\ell) \Leftrightarrow \widehat{ULS}_{ik} \geq \widehat{ULS}_{i\ell}, \tag{31}$$

for all $k, \ell \in U_K(i)$ and $i = 1, \ldots, n$.

Finally, to determine the degree of inconsistency between this regression-rank ordering and each i-rank ordering in (27), note first that since we are primarily interested in comparisons of these orderings in the *upper-tail* (where CPL properties are most critical), it is not essential to consider all possible i-rank orderings. In particular, for any given *cut-off level*, $I \leq n$, one may choose to consider only the I top ranked cities, $i = 1, \ldots, I$. (In the analysis to follow, we set $I = 100$.) Next observe that if for each $i \leq I$ and $k \in U_K(i)$ we let $R_i(k) = \{\ell \in U_K(i) : \ell > k\}$, then the total number of distinct comparisons to be made is given by $N_K(I) = \sum_{i=1}^{I} \sum_{k \in U_K(i)} |R_i(k)|$. Moreover, observe that for any $\ell \in R_i(k)$, a disagreement occurs between the orderings of r_i and $\hat{r}$ if and only if the *signed differences*, $sgn[r_i(k) - r_i(\ell)]$ and $sgn[\hat{r}(k) - \hat{r}(\ell)]$, are not equal (including possible zeros).[40] So if we let $\delta(x, y) = 1 \Leftrightarrow sgn(x) \neq sgn(y)$, and $\delta(x, y) = 0$ otherwise, then the desired *degree of order inconsistency* is given by the following fraction of disagreements,

$$\phi_K(I) \equiv \frac{1}{N_K(I)} \sum_{i=1}^{I} \sum_{k \in U_K(i)} \sum_{\ell \in R_i(k)} \delta(r_i(k) - r_i(\ell), \hat{r}(k) - \hat{r}(\ell)) \,. \tag{32}$$

Note in particular that $\phi = 0$ if the upper-average distributions of partition cells never cross one another.

To employ $\phi_K(I)$ as a test statistic for a given K-partition, here we consider only the top 100 cities ($I = 100$) and compute $\phi_K(100)$ for this partition.[41] Our *null hypothesis, H_0,* is again that this value is not statistically distinguishable from those values, $\widetilde{\phi}_K(100)$, derived from random K-partitions of the same size (as in Section 5.2 above). To construct a one-sided test of H_0, we again sample 1000 random K-partitions of the same size, and estimate the p-value for this test by the fraction of $\widetilde{\phi}_K(100)$ values smaller than $\phi_K(100)$. The results of these tests for the relevant range of partition sizes, $K = 2, \ldots, 7$, and commodity-flow thresholds, $\lambda \in [0.05, 0.13]$, are shown in the left panel of Figure 16.[42] For purposes of comparison, the ratios of the actual $\phi_K(100)$ values to the medians, $\overline{\phi}_K(100)$, of their corresponding random ϕ-values are shown in the right panel.

[40] Recall that the *sign function* is defined by $sgn(x) = -1, 0, 1$ iff $x < 0,\ x = 0,\ x > 0$.

[41] Essentially the same results are obtained under different values of I.

[42] These tests were actually conducted for $\lambda \in [0.05, 1.00]$, but none of the curves show significant change beyond $\lambda = 0.13$.

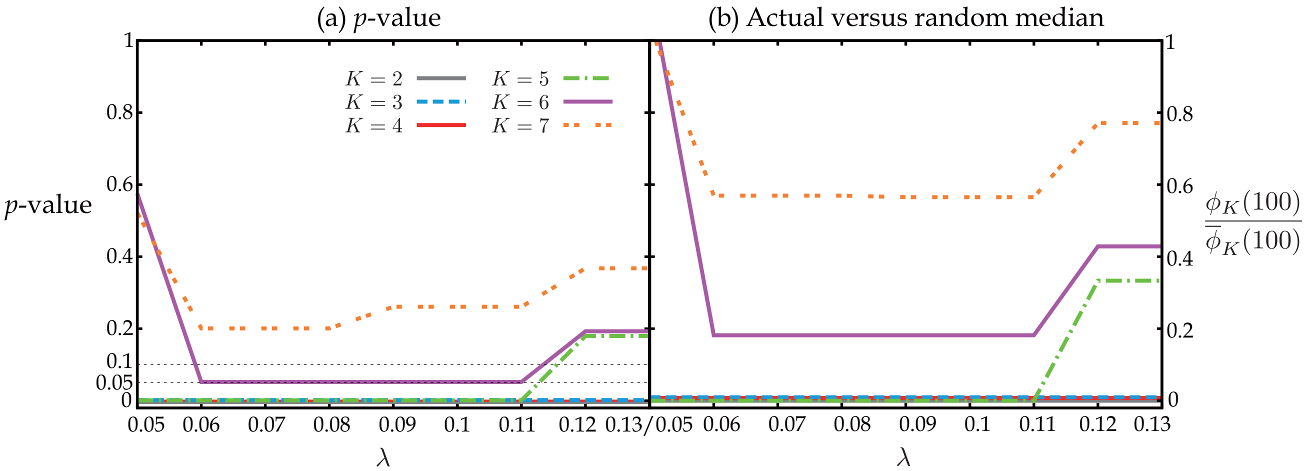

Figure 16: Consistency in the orders of upper-average distributions across partition cells

From these p-value results, it is clear that all $K = 2, \ldots, 5$ are extremely significant for $\lambda \in [0.05, 0.11]$, and that $K = 6$ is also significant (at the 0.05 level) for $\lambda \in [0.06, 0.11]$. For example, Figure 10 shows that $\phi_4(100) = 0$ for $\lambda = 0.10$,[43] where in all Figures 10 through 12, the rank value, $I = 100$, corresponds to upper log rank value, $ULR_{100} \simeq 3.64$. The sharp increase in significance for $K = 6$ at $\lambda = 0.06$ corresponds precisely to the "Atlanta effect" described in Section 5.2 above, and similarly, the sharp decrease in significance for $K = 5$ and 6 at $\lambda = 0.12$ corresponds to the "Philadelphia effect". So, aside from these special effects, it should be clear that those K-partitions of economic regions exhibiting the strongest CPL properties in Section 5.2 above also exhibit the strongest order-consistency properties with respect to their upper-average distributions. But, the result under $K = 2$ indicates that the order-consistency test captures certain aspects of similarity among upper-average distributions not captured by the CPL test in Section 5.2. In particular, although the CPL property was not significant at $K = 2$ in terms of the test in Section 5.2, the economic regions of the largest two cities still exhibit a similarity in city size distributions in terms of the order-consistency.[44]

6 Conclusions

In this paper we have examined the question of whether spatial relations among cities may influence the distribution of city sizes. Specifically, we have tested the implication of *iid*

[43] More generally, $\phi_K(100) = 0$ for $K = 2, 3, 4$ at all values $\lambda \in [0.05, 0.13]$, and $\phi_5(100) = 0$ at values $\lambda \in [0.05, 0.11]$.

[44] In fact, the upper-average distributions of economic regions do not cross at the relevant ranks under $K = 2$ and 3, i.e., $\phi_2(I) = \phi_3(I) = 0$ for all $I \geq 1$.

random growth processes that a CPL must hold across arbitrary subsets of cities, regardless of their spatial relations. Using CBSA data from the US, we have shown that this CPL property is in fact much stronger for spatial groupings of *nearby cities* (as determined by Voronoi partitions) than for random groupings of cities. In addition, we conjectured that such spatial groupings are characterized by the *spacing-out property* that larger cities tend to be widely spaced, with smaller cities grouped around these centers. Our second series of results found (independently of any CPL considerations) that there is strong evidence for this spacing-out property in the US. We then combined certain aspects of these results by replacing Voronoi groupings of nearby cities with *economic regions* that reflect the spacing-out property among cities in terms of commodity-flow relations rather than simple spatial proximity. Our final set of results confirm that the CPL property is even stronger for these economic regions.

With respect to current theories of city systems, our present results appear to be best accounted for by central place theory.[45] As mentioned in the introduction, the model of Hsu [31] yields an explicit CPL property among economic regions under certain regularity conditions. Moreover, while this firm-entry model is highly stylized, it has been shown by Akamatsu et al. [3] that the basic results of this model continue to hold in the more general *new economic geography* versions of central place models *à la* Fujita et al. [21] and Tabuchi and Thisse [49]. In particular, these general equilibrium models continue to exhibit the same structural features of central place theory, namely numerous industries with heterogeneous degrees of scale economies and numerous locations from which agents choose to live and work.[46]

Finally, we suggest an empirical extension for the identification of economic regions. One key feature of central place models is the *hierarchical nesting* of market areas, and the above-mentioned central place models actually exhibit CPL property across these hierarchies. Recall that our present testing schemes involve only simple partitions of city sets when constructing economic regions. However, our results (as illustrated for example by the "Philadelphia effect") suggest that hierarchical systems of nested regions may in many cases be more appropriate. Along these lines, the paper by Akamatsu et al. [3] above proposes a *hierarchical partitioning scheme* inspired by the central place models. Here

[45] It should be noted here that Behrens, Duranton and Robert-Nicoud [7] have recently shown that talent heterogeneity among cities can lead to Zip's law (under certain additional conditions). But, since there is no notion of inter-city space in their model, this cannot be said to account for our findings.

[46] In a related work, Brackman et al. [10] have shown that the observed diversity of city sizes can to some degree be reproduced by introducing negative externalities of agglomeration into the standard multiple-region NEG model of Krugman [37]. However, such negative externalities cannot account for either the CPL or spacing-out properties. The key difference of their model from central place models is the presence of only a single industry group, which precludes any central place hierarchy.

each economic region is further partitioned into *economic subregions* based on intraregional commodity flows. In this setting it is shown (using the same data as ours) that such a nested structure exhibits strong CPL properties between economic regions and their subregions (such as between the New York and Philadelphia regions).

A Flow-Dependency Levels among the 20 Largest-city CFS Regions

This Appendix includes further details on flow dependencies among the first 20 regions in R. These dependencies are shown in Table 3, where flow dependency, λ_{ij}, corresponds to the cell in row i and column j. Here the colored cells (both blue and red) identify all flow dependencies, $\lambda_{ij} > 0.05$, which thus exclude the associated column region, j, from central region systems with $\lambda = 0.05$ that contain row region i. In particular, the red cells identify all flow dependencies, $\lambda_{ij} > 0.10$, so that these column regions, j, are also excluded from central region systems with $\lambda = 0.10$ that contain row region i.

	Rank of CFS regions (in terms of the largest city size)																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. New York	**0.497**	0.027	0.029	0.029	0.119	0.018	0.052	0.083	0.031	0.099	0.017	0.024	0.026	0.028	0.018	0.036	0.050	0.032	0.108	0.025
2. Los Angeles	0.035	**0.568**	0.020	0.037	0.039	0.014	0.038	0.043	0.029	0.022	0.014	0.092	0.151	0.068	0.028	0.364	0.020	0.028	0.000	0.062
3. Chicago	0.016	0.016	**0.410**	0.019	0.017	0.017	0.020	0.019	0.032	0.015	0.058	0.009	0.026	0.016	0.060	0.010	0.043	0.013	0.013	0.023
4. Dallas	0.009	0.013	0.009	**0.360**	0.007	0.034	0.009	0.004	0.012	0.006	0.009	0.009	0.021	0.013	0.008	0.015	0.011	0.010	0.007	0.031
5. Philadelphia	0.066	0.005	0.009	0.005	**0.360**	0.003	0.008	0.048	0.009	0.022	0.009	0.006	0.005	0.005	0.006	0.008	0.006	0.010	0.034	0.005
6. Houston	0.004	0.008	0.009	0.055	0.008	**0.603**	0.003	0.002	0.008	0.003	0.011	0.004	0.005	0.000	0.003	0.004	0.008	0.004	0.002	0.000
7. Miami	0.005	0.003	0.002	0.006	0.003	0.003	**0.436**	0.003	0.007	0.003	0.001	0.001	0.003	0.002	0.002	0.002	0.003	0.028	0.003	0.002
8. Whasington, D.C.	0.003	0.000	0.001	0.000	0.003	0.000	0.001	**0.142**	0.000	0.001	0.000	0.000	0.002	0.001	0.000	0.000	0.000	0.000	0.042	0.000
9. Atlanta	0.005	0.005	0.010	0.010	0.006	0.006	0.028	0.009	**0.369**	0.005	0.006	0.003	0.007	0.006	0.005	0.006	0.007	0.035	0.016	0.009
10. Boston	0.035	0.006	0.007	0.007	0.018	0.003	0.015	0.014	0.008	**0.442**	0.007	0.008	0.001	0.008	0.008	0.014	0.004	0.008	0.013	0.009
11. Detroit	0.006	0.006	0.016	0.012	0.007	0.005	0.005	0.004	0.011	0.010	**0.354**	0.005	0.011	0.009	0.018	0.000	0.026	0.013	0.046	0.014
12. San Francisco	0.007	0.033	0.007	0.015	0.006	0.008	0.017	0.012	0.005	0.009	0.004	**0.496**	0.028	0.015	0.006	0.028	0.004	0.008	0.003	0.019
13. Phoenix	0.002	0.013	0.005	0.008	0.001	0.002	0.003	0.003	0.005	0.002	0.002	0.011	**0.417**	0.006	0.004	0.014	0.002	0.003	0.001	0.014
14. Seattle	0.002	0.007	0.004	0.003	0.002	0.000	0.005	0.002	0.000	0.002	0.003	0.006	0.000	**0.416**	0.003	0.005	0.001	0.002	0.002	0.005
15. Minneapolis	0.007	0.006	0.018	0.008	0.009	0.003	0.006	0.007	0.000	0.006	0.006	0.006	0.012	0.007	**0.459**	0.018	0.007	0.011	0.006	0.009
16. San Diego	0.002	0.028	0.000	0.008	0.007	0.002	0.005	0.007	0.000	0.004	0.001	0.006	0.008	0.010	0.003	**0.298**	0.004	0.002	0.004	0.005
17. St. Louis	0.005	0.000	0.022	0.004	0.005	0.003	0.004	0.005	0.010	0.003	0.005	0.002	0.003	0.003	0.006	0.002	**0.384**	0.004	0.007	0.003
18. Tampa	0.003	0.002	0.002	0.002	0.003	0.001	0.020	0.006	0.004	0.003	0.001	0.001	0.001	0.001	0.002	0.001	0.002	**0.355**	0.002	0.002
19. Baltimore	0.009	0.001	0.002	0.002	0.022	0.001	0.002	0.104	0.006	0.006	0.002	0.001	0.001	0.001	0.002	0.002	0.001	0.002	**0.194**	0.000
20. Denver	0.002	0.004	0.002	0.004	0.002	0.003	0.002	0.001	0.004	0.002	0.002	0.007	0.008	0.008	0.000	0.000	0.001	0.000	0.001	**0.433**

Table 3: Flow-dependency levels among the 20 largest-city CFS regions

B CPL Test under G-I and R-S data

In this section, we redo the CPL tests in Sections 3 and 5 using both G-I and R-S data for purposes of comparison with the U-A data adopted in those sections. In particular, since G-I represents an alternative weighting scheme for the upper tails of city size distributions, tests based on G-I data serve as a natural robustness check for our U-A results.

Before doing so however, it is of interest to note that if the upper tail of the US city-size distribution is hypothesized to follow an exact Pareto distribution, then alternative testing procedures could in principle be constructed using the optimal truncation approach of Clauset et al. [12]. These authors employ maximum likelihood estimation together with Kolmogorov-Smirnov (KS) distance to determine that truncation point which yields a best fit to the upper tail. For the US case, these KS distances are plotted at each potential cutoff rank (between 2 and 930) in Figure 17. As is clear from the figure, the optimal trucation in this case is unambiguously at the 916th rank, yielding $\hat{\kappa} = 0.841$ under the G-I regression. But, since this value differs from the untruncated estimate ($\hat{\kappa} = 0.833$) by less than 1%, the optimal truncation for this US data amounts essentially to no truncation at all. Thus, the CPL tests under G-I and R-S data below will be conducted without truncation.

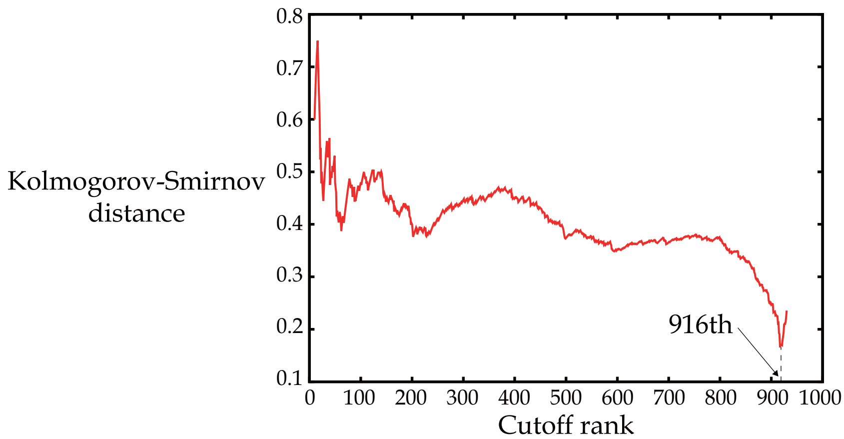

Figure 17: Optimal truncation under Pareto distribution

B.1 Voronoi versus Random Partitions

In this section, we conduct the CPL test in Section 3 using G-I and R-S data. Panels (a) and (b) in Figure 18 are parallel versions of Figure 5 using G-I and R-S data, respectively. As in Figure 5, one can see from the corresponding percentile points at the 5%, 10%, 50%, 90% and 95% levels that the distribution of p-values is uniformly below what would be expected under H_0. Thus, the basic conclusion remains the same under both the G-I and R-S data.

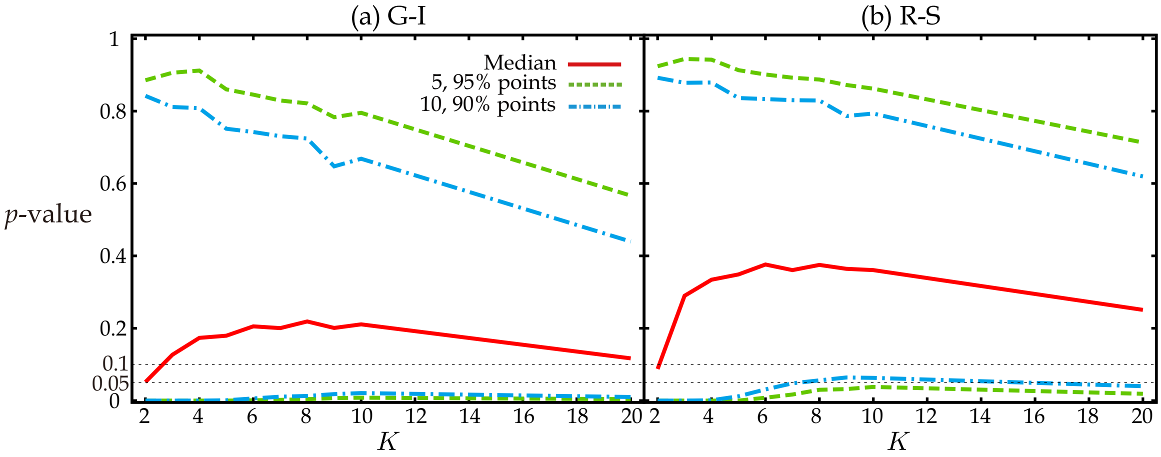

Figure 18: Summary of the CPL test

The most noticeable difference from Figure 5 under the U-A data is that for both the G-I and R-S data there are now a substantially larger number of Voronoi partitions exhibiting significant CPL property (both at the .05 and .10 levels) for small values of K. This can be explained by the fact that when the number of partition cells (K) is small, the upper tails for each partition cell under U-A are relatively more sensitive to the allocations of the largest cities among cells. Hence, the CPL property is relatively less pronounced under U-A for small K.

B.2 Economic-Regions versus Random Partitions

In a similar manner, we now repeat the CPL tests in Section 5 for both the G-I and R-S data (with parameter values $\lambda = 0.05, \ldots, 0.12$ and $K = 2, \ldots, 8$). The results paralleling Table 2 for the G-I data are shown in Table 4. Here again the basic conclusion of Section5 remains the same.

	K						
λ	2	3	4	5	6	7	8
0.05	●	○	◐				
0.06	●	○	◐		◐		○
0.07	●	○	◐		◐		○
0.08	●	○	◐		◐		○
0.09	●	○	◐		◐		
0.10	●	○	◐		◐		
0.11	●	○	◐				
0.12	●	○	◐	○		○	

Table 4: Result of the CPL test under G-I data

The most distinctive difference is the case of $K = 2$ for which the CPL property was not significant under the U-A data. Recall from the discussion in Section 5.2 that the upper tails of the city size distributions in the New York and Los Angles regions did not exhibit significant CPL due mainly to the fact that the largest cities in the third largest economic region, Chicago, have been rather randomly allocated to these two regions. But the lower tails of the city size distributions for these regions happen to be fairly similar, as can be seen from the R-S and U-A distributions of the New York and Los Angeles regions in Figure 19. Hence, given that G-I is relative more sensitive to lower tails than U-A, the CPL property is now seen to be significant (at the .01 level) under G-I. The results for $K \geq 3$ under G-I data are similar to those under U-A.

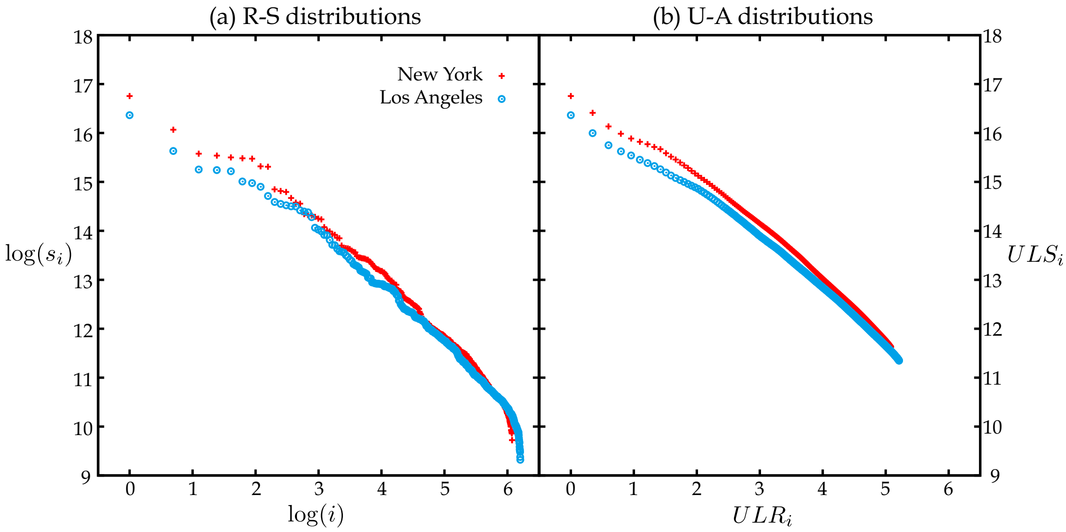

Figure 19: R-S and U-A distributions for $K = 2$

Turning finally to the R-S data, the test results show that (for all λ) economic regions now exhibit a significant CPL property for $K = 2$, but none for larger values of K (so that in terms of Table 4, the first column is again all black dots with blanks everywhere else). The explanation here is due to the even more extreme emphasis of R-S on the lower tails of city-size distributions, which happen to yield a strong CPL property for the $K = 2$ case in Figure 19(a), but which also ignore the upper tail properties that are so crucial for power laws in general.

References

[1] Ahlfedlt, Gabriel, Stephen Redding, Daniel Sturm and Nikolaus Wolf. 2002. "The economics of density: Evidence from the Berlin Wall." Centre for Economic Performance Discussion Paper, 1154, June 2012.

[2] Ait-Sahalia, Yacine and Jefferson Duarte. 2003. "Nonparametric option pricing under shape restrictions", *Journal of Econometrics* 116: 9 – 47.

[3] Akamatsu, Takashi, Tomoya Mori and Yuki Takayama. 2014. "Spatial coordinations among industries and the common power laws for city size distributions." In progress. Institute of Economic Research, Kyoto University.

[4] Allen, Treb and Costas Arkolakis. 2014. "Trade and the topography of the spatial economy." The Quarterly Journal of Economics, forthcoming.

[5] Anas, Alex, Richard Arnott and Kenneth A. Small. 1998. "Urban Spatial Structure." *Journal of Economic Literature* 36(3): 1426-1464.

[6] Anderson, James and Eric van Wincoop. 2004. "Trade costs." *Journal of Economic Literature* 42(3): 691-751.

[7] Behrens, Kristian, Gilles Duranton and Frédéric Robert-Nicoud. 2014. "Productive cities: Sorting, selection, and agglomeration." *Journal of Political Economy*, forthcoming.

[8] Behrens, Kristian, Giordano Mion, Yasusada Murata and Jens Südekum. 2013. "Spatial frictions." Center for Economic Policy Research, No.8572.

[9] Black, Duncan and Vernon Henderson. 2003. "Urban evolution in the USA." *Journal of Economic Geography* 3(4): 343-372.

[10] Brakman, Steven, Harry Garretsen, Charles Van Marrewijk and Marianne van den Berg. 1999. "The return of Zipf: Towards a further understanding of the rank-size distribution." *Journal of Regional Science* 39(1): 183-213.

[11] Christaller, Walter. 1966. *Central Places in Southern Germany*. Translated by Carlisle W. Baskin. Upper Saddle River, N.J.: Prentice Hall. Originally published as *Die zentralen Orte in Süddeutschland* (Gustav Fischer, 1933).

[12] Clauset, Aaron, Cosma Rohilla Shalizi and M.E.J. Newman. 2009. "Power-law distributions in empirical data." *SIAM Review* 51(4): 661-703.

[13] Desmet, Klaus and Jordan Rappaport. 2013. "The settlement of the United States, 1800 to 2000: The long transition towards Gibrat's law." CEPR Discussion Papers 9353.

[14] Desmet, Klaus and Esteban Rossi-Hansberg. 2013. "Urban accounting and welfare." *American Economic Review* 103(6): 2296-2327.

[15] Dobkins, Linda H. and Yannis M. Ioannides. 2001. "Spatial interactions among U.S. cities: 1900–1990." *Regional Science and Urban Economics* 31(6): 701–731.

[16] Donaldson, Dave. 2012. "Railroads of the Raj: Estimating the impact of transportation infrastructure." *American Economic Review*, forthcoming.

[17] Duranton, Gilles. 2006. "Some foundations for Zipf's law: Product proliferation and local spillovers." *Regional Science and Urban Economics* 36: 543-563.

[18] Duranton, Gilles, Peter M. Morrow and Matthew A. Turner. 2013. "Roads and trade: Evidence from the US." Mimeograph. The Wharton School, University of Pennsylvania.

[19] Eeckhout, Jan. 2004. "Gibrat's law for (all) cities." *American Economic Review* 94(5): 1429-1451.

[20] Fajgelbaum, Pablo and Stephen J. Redding. 2014. "External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914," NBER Working Paper, 20217.

[21] Fujita, Masahisa, Paul Krugman and Tomoya Mori. 1999. "On the evolution of hierarchical urban systems." *European Economic Review* 43(2): 209-251.

[22] Gabaix, Xavier. 1999. "Zipf's law for cities: An explanation." *The Quarterly Journal of Economics* 114: 738-767.

[23] Gabaix, Xavier. 2009. "Power laws in economics and finance." *Annual Review of Economics* 1: 255-93.

[24] Gabaix, Xavier and R. Ibragimov. 2011. "Log(Rank-1/2): A simple way to improve the OLS estimation of tail exponents." *Journal of Business & Economic Statistics* 29: 24-39.

[25] Ghosh, Arka P., Diana Hay, Vivek Hirpara, Reza Rastegar, Alexander Roitershtein, Ashley Schulteis, and Jiyeon Suh. 2010. "Random linear recursions with dependent coefficients." *Statistics and Probability Letters* 80: 1597-1605.

[26] Giesen, Kristian and Jens Südekum. 2011. "Zipf's law for cities in the regions and the country." *Journal of Economic Geography* 11(4): 667-686.

[27] Glaeser, Edward L., Giacomo A. M. Ponzetto, and Kristina Tobio. 2011. "Cities, skills, and regional change." Working paper, Harvard University

[28] Goldie, C.M. 1991. "Implicit renewal theory and tails of solutions of random equations." *Annals of Applied Probability* 1: 126-166.

[29] Holmes, Thomas J. and Sanghoon Lee. 2010. "Cities as six-by-six-mile squares: Zipf's law?" NBER Chapters, in: *Agglomeration Economics*: 105-131.

[30] Hering, Laura and Sandra Poncet. 2010. "Market access and individual wages: Evidence from China." *Review of Economics and Statistics* 92(1): 145-159.

[31] Hsu, Wen-Tai. 2012. "Central Place Theory and City Size Distribution." *Economic Journal* 122: 903-932.

[32] Ioannides, Yannis and Henry G. Overman. 2003. "Zipf's law for cities: An empirical examination." *Regional Science and Urban Economics* 33(2): 127-137.

[33] Ioannides, Yannis M. and Henry G. Overman. 2004. "Spatial evolution of the US urban system." *Journal of Economic Geography* 4: 131-156.

[34] Ioannides, Yannis M. and Spyros Skouras. 2013. "US city size distribution: Robustly pareto, but only in the tail." *Journal of Urban Economics* 73(1): 18-29.

[35] Kesten, H. 1973. "Random difference equations and renewal theory for products of random matrices." *Acta Mathematica* 131: 207-248.

[36] Kitamura, Y., G. Tripathi and H. Ahn. 2004. "Empirical likelihood-based inference in conditional moment restriction models." *Econometrica* 72, 1667–1714.

[37] Krugman, Paul. 1993. "On the number and location of cities." *European Economic Review* 37(2-3): 293-298.

[38] Lucas, Robert E. Jr. and Esteban Rossi-Hansberg. 2002. "On the internal structure of cities." *Econometrica* 70(4): 1445-1476.

[39] McMillen, D., and C. Redfearn, 2010, "Estimation and hypothesis testing for nonparametric hedonic house price functions." *Journal of Regional Science* 50: 712-733.

[40] Michaels, Guy. 2008. "The effect of trade on the demand for skill - evidence from the interstate highway system." *Review of Economics and Statistics* 90(4): 683-701.

[41] Michaels, Guy, Ferdinand Rauch and Stephen J. Redding. 2012. "Urbanization and structural transformation." *Quarterly Journal of Economics* 127(2): 535-586.

[42] Nishiyama, Yoshihiko, Susumu Osada and Yasuhiro Sato. 2008. "OLS estimation and the *t*-test revisited in rank-size regression." *Journal of Regional Science* 48: 691-716.

[43] Nitsch, Volker. 2005. "Zipf zipped." *Journal of Urban Economics* 57: 86–100.

[44] Redding, Stephen and Daniel Sturm. 2008. "The costs of remoteness: Evidence from German division and reunification." *American Economic Review* 98(5): 1766-1797.

[45] Rossi-Hansberg, Esteban and Mark L. J. Wright. 2007. "Urban structure and growth." *Review of Economic Studies* 74: 597-624.

[46] Rozenfeld, Hernan, Diego Rybski, Xavier Gabaix and Hernan Makse. 2011. "The area and population of cities: New insights from a different perspective on cities." *American Economic Review* 101(5): 2205-2225.

[47] de Saporta, Benoite, 2005. "Tail of the stationary solution of the stochastic equation $Y_{n+1} = a_n X_n + b_n$, with Markovian coefficients." *Stochastic Processes and their Applications* 115: 1954-1978.

[48] Soo, Kwok Tong. 2005. "Zipf's law for cities: A cross country investigation." *Regional Science and Urban Economics* 35(3): 239-263.

[49] Tabuchi, Takatoshi and Jacques-François Thisse. 2011. "A new economic geography model of central places." *Journal of Urban Economics* 69: 240-252.

[50] The US Office of Management and Budget. 2010. "Standards for defining metropolitan and micropolitan statistical areas." *Federal Register* 75, No.123.

[51] The US Bureau of Transport Statistics. 2007. *Commodity Flow Survey*.

[52] The US Census Bureau. 2010. "Census Bureau regions and divisions with state FIPS codes." http://www.census.gov/geo/www/reference.html.

[53] The US Geographical Survey. 2001. *National Land Cover Database*.

[54] Vervaat, Wim. 1979. "On a stochastic difference equation and a representation of non-negative infinitely divisible random variables." *Advances in Applied Probability* 11: 750-783.