Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection School Of Economics School of Economics

3-2019

Central place theory and the power law for cities

Wen-Tai HSU
Singapore Management University, WENTAIHSU@smu.edu.sg

ZOU XIN
Singapore Management University, xin.zou.2014@phdecons.smu.edu.sg

Follow this and additional works at: https://ink.library.smu.edu.sg/soe_research

Part of the Behavioral Economics Commons, and the Urban Studies and Planning Commons

Citation

HSU, Wen-Tai and ZOU XIN. Central place theory and the power law for cities. (2019). *The mathematics of urban morphology: Modeling and simulation in science, engineering and technology*. 55-75.
Available at: https://ink.library.smu.edu.sg/soe_research/2281

This Book Chapter is brought to you for free and open access by the School of Economics at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Economics by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylds@smu.edu.sg.

Central Place Theory and the Power Law for Cities

Wen-Tai Hsu and Xin Zou

Abstract This chapter provides a review of the link between central place theory and the power laws for cities. A theory of city size distribution is proposed via a central place hierarchy a la Christaller (1933) either as an equilibrium results or an optimal allocation. Under a central place hierarchy, it is shown that a power law for cities emerges if the underlying heterogeneity in economies of scale across good is regularly varying. Furthermore, we show that an optimal allocation of cities conforms with a central place hierarchy if the underlying heterogeneity in economies of scale across good is a power function.

Keywords Central place theory · Zipf's law · City sizes · Dynamic programming · Optimal city hierarchy

1 Introduction

City size distribution is known to be well approximated by a power law with a tail index around 1, i.e., a Zipf's law. To visualize it, we first rank cities by city size (population): #1 is New York, #2 Los Angeles etc., then we plot the city ranks against city sizes on a log-log scale using U.S. 2000 census data for all Metropolitan Statistical Areas (MSAs). The relationship for 362 MSAs, as shown in Fig. 1, is close to a straight line ($R^2 = 0.9857$), and the slope is close to -1 (-0.9491). Even when the smaller cities and towns are consider and the *overall* city size distribution does not follow the power law (Eeckhout 2004), the right tail of the distribution can still be well approximated by a power law. This empirical regularity has been widely

W.-T. Hsu (✉) · X. Zou
School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903, Singapore
e-mail: wentaihsu@smu.edu.sg

X. Zou
e-mail: xin.zou.2014@phdecons.smu.edu.sg

© Springer Nature Switzerland AG 2019

L. D'Acci (ed.), *The Mathematics of Urban Morphology*,
Modeling and Simulation in Science, Engineering and Technology,
https://doi.org/10.1007/978-3-030-12381-9_3

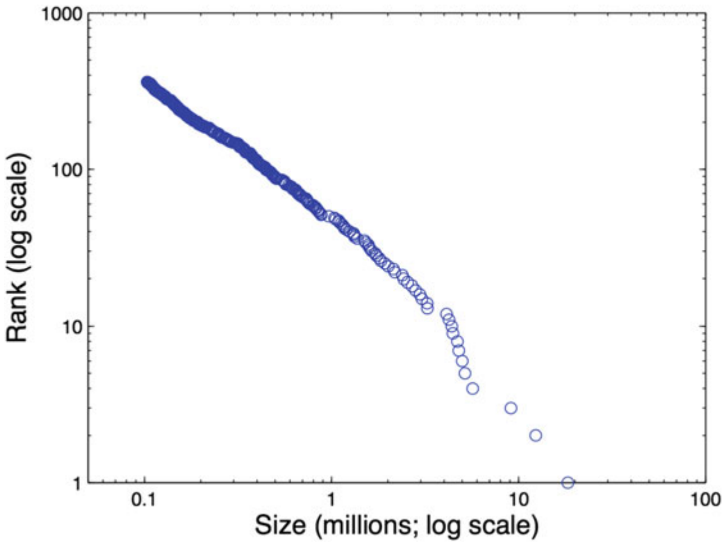

Fig. 1 Zipf plot for all 362 MSAs from 2000 U.S. Census

documented using data spanning across many countries and time periods (see Gabaix and Ioannides (2004) and Dittmar (2011)).

Formally, a power-law can be expressed as the tail probability of a city size distribution following a power function such that

$$P(S > s) = a/s^{\zeta},\ s \geq \underline{s},$$

where a, ζ, and $\underline{s}$ are some positive constants.

One popular explanation for power laws is proportional random growth (see Champernowne (1953), Simon (1955)), based on which (Gabaix 1999) provides an explanation for why the exponent ζ should be close to 1. In this chapter, we introduce a very different theory for the power laws that is proposed by Hsu (2012) and Hsu et al. (2014) and built on the insight of *central place theory*. The key difference from the random growth approach is that central place theory explicitly accounts for the spatial pattern of different sized cities based on geography and the heterogeneity of industries/goods. In such a theory, cities of different sizes play different roles by hosting different industries.

Central place theory was first proposed by Christaller (1933), who described how a city hierarchy emerges from a uniformly populated plane of farmers as consumers. Cities and towns provide a wide variety of goods to these farmers, as these cities and towns have their market areas, they are referred to "central places" of these market areas. The key feature is that goods differ in terms of their degrees of scale economies, and the *hierarchy property* states that if a city provides a good with certain degree of scale economies, it provides all goods with lower scale economies.[1] Put it differently,

[1]This is often called "hierarchy principle" in the literature.

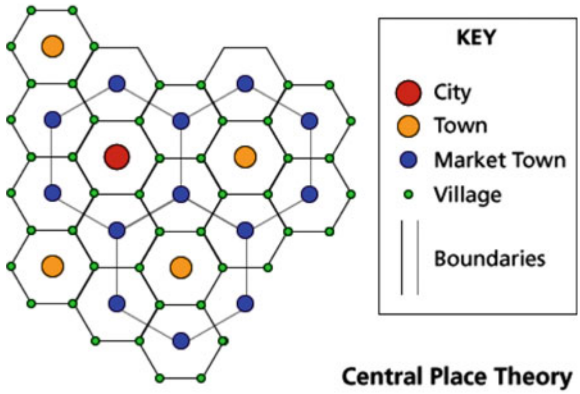

Fig. 2 Central place hierarchy on the plane

larger cities provide all goods that are provided by smaller ones. Some economists started to pay their attention to geography and spatial economics around early 80s. Although central place theory has developed into a main building block in economic geography, its theoretical foundation is weak in economists' point of view. The major problem is that it largely remains a collections of assumptions and statements without rigorous logical deduction, and how the *central place hierarchy* can emerge from an equilibrium with economic agents optimizing given their constraints is unclear. Hence, a microfoundation is needed.

What the picture below shows is a depiction of central place theory in Christaller's form. That is, on a uniformly populated plane of farmers, with uniform soil productivity, market areas of each central place is a hexagon. There are different layers of central places, with central places of the same layer having the same size of market area. A recently proven "honeycomb conjecture" in Hales (2001) provides a ground for why the shape of market areas should be hexagon, as regular hexagons are the most efficient way to partition the 2-D plane with least perimeter used. The hierarchy property is implicitly shown here. When the larger central places provides all goods and services provided by the smaller ones, we also see that the larger central places are selected from a subset of the locations of the smaller central places in a regular way. Of course reality wouldn't conform the central place pattern as in Fig. 2, but it provides a way to comprehend the complex locational patterns of cities and towns on the maps that we see. Put it differently, the seemingly pattern-less, complex map of cities and towns may indeed have some kind of sensible reasoning behind them, only to be obscured by many other factors that distorted the underlying pattern.

Unfortunately, central place theory in Christaller's form is too difficult the deal with, mainly due to the 2-dimensional geographic space. Economists have made several attempts to try to provide a "formal" central place theory in various ways, and these include Eaton and Lipsey (1982), Quinzii and Thisse (1990), Fujita et al. (1999), Tabuchi and Thisse (2006, 2011), Hsu (2012) and Hsu et al. (2014). The common task of the above-mentioned papers are again the hierarchy property and the locational patern. In all of these "modern" central place theories, the geographic space is either two-location space or a one-dimensional space. What distinguishes

Hsu (2012) and Hsu et al. (2014) from other economic models of central place theory is that these two explain the power laws for cities. However, note that these papers are not the first that link central place hierarchy with the power laws. Indeed, Beckmann (1958) was the first to point out the link between central place theory and the power law. Nevertheless, he provides no economic model for his hierarchical structure, and his hierarchy lacks the dimension of industries, which is crucial in the theory by Hsu (2012) and Hsu et al. (2014).

Similar to Beckmann (1958), there is a large literature in geography that examines the relationships between central place systems and power laws via the angle of fractal analysis. See, for examples, Arlinghaus (1985), Arlinghaus and Arlinghaus (1989), Batty and Longley (1994), Frankhauser (1998), Chen and Zhou (2006), Chen (2011, 2014). This strand of analysis is interesting as it examines different variants of fractals that could link the system of cities with the power laws. Compared with economic theories described above, we can thus see richer morphology of the system of cities without the hands being tied with the need for a microfoundation. In this sense, this strand of literature in geography and the attempts in economics to provide microfoundation for central place theory complement each other.

Before we formally introduce the theory, we first peek at the result—a central place hierarchy either as an equilibrium outcome or as an optimal allocation. Here, we formally define a central place hierarchy as an allocation of production locations such that both the *hierarchy property* and *central place property* hold. A central place hierarchy in a one-dimensional geographic space can be illustrated by Fig. 3. As the vertical axis is the commodity space, the hierarchy property is clearly seen in this figure. The central place property is that any city is located in the middle between two neighboring larger sized cities.

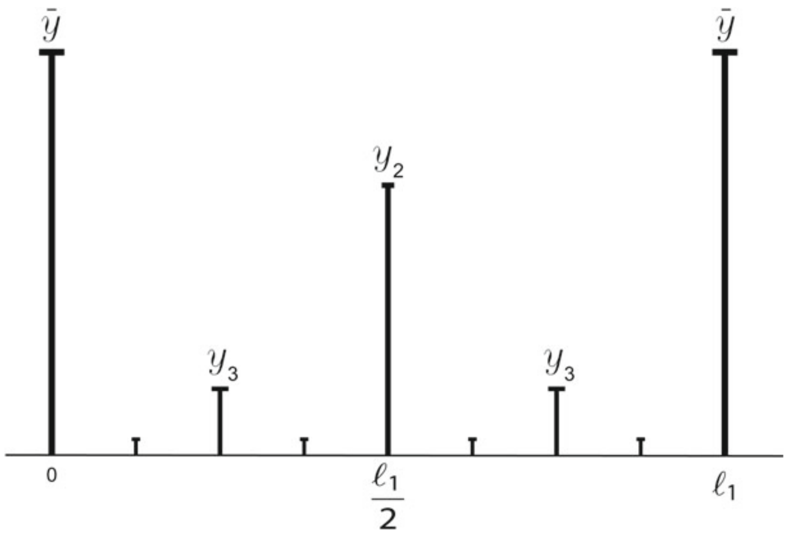

Fig. 3 Central place hierarchy on the line

This chapter is organized as follows: Sects. 2 and 3 explain the city size distribution by formalizing central place theory as an equilibrium outcome. With the hierarchy property imposed, Sect. 4 show that how a central place hierarchy can emerge as an optimal allocation. In particular, we apply dynamic programming to study the social planner's problem. We conclude this chapter in Sect. 5.

2 Central Place Theory

We first lay out a basic model and derive the central place hierarchy as an equilibrium outcome.

2.1 Model and One-Good Equilibrium

The geographic space is the real line, and the location is indexed as $x \in \mathbb{R}$. There is a continuum of consumption goods $z \in [0, z_1]$, where z_1 is exogenously given. There are two types of agents: farmers and firms. The farmers are immobile and are uniformly distributed on the real line with a density of 1. Each farmer demands one unit of each good z in $[0, z_1]$ inelastically.

For any production of good $z \leq z_1$, a fixed cost $y = \phi(z)$ is needed, and the (cumulative) distribution function of y is denoted as $F : [\underline{y}, \bar{y}] \subset \mathbb{R}_+ \to [0, 1]$. Besides the fixed cost, producing one unit of each good requires constant marginal cost c. Moreover, it occurs a transportation cost t per unit per mile traveled. For each good, there is an infinite pool of potential firms. The firms and farmers play the following two-stage game (Lederer and Hurter 1986).

1. Entry and location stage
   The potential firms simultaneously decide whether to enter. Upon entering, each entrant chooses a location and pays the setup cost for the good it produces. Assume the tie-breaking rule: if a potential firm sees a nonnegative opportunity, then it enters.
2. Price competition stage
   The firms deliver goods to the farmers. Given its own and other firms' locations, each firm sets a delivered price schedule over the real line. For each good, each location on the real line is a market in which the firms engage in Bertrand competition. Each farmer decides the specific firm from which to buy each good.

Now let's consider the subgame perfect equilibrium (SPNE) based on the two-stage game setting above. Denote the firm on the left-hand side as A located at $x = 0$ and that on the right-hand side as B located at $x = L$. The marginal costs of delivering the good to a consumer who is x distance from A are thus:

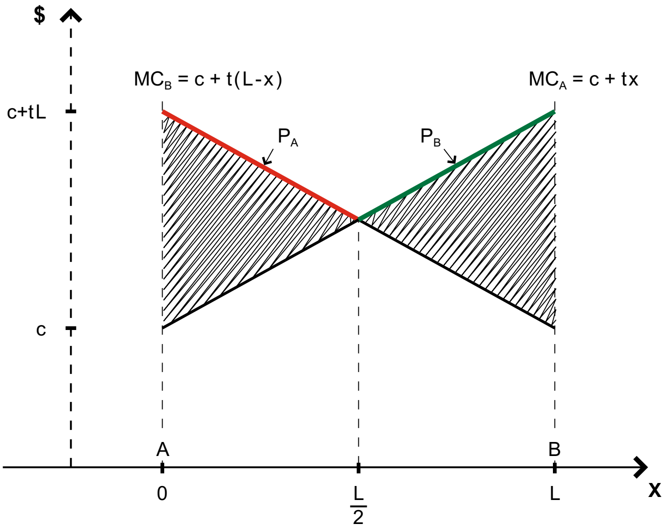

Fig. 4 Second-stage competition: prices and gross profits

$$MC_A = c + tx,$$
$$MC_B = c + t(L - x).$$

Bertrand competition results in firm with the lower marginal cost dominating the market and charging the price of its opponent's marginal cost. Thus, the equilibrium prices at each x on $[0, L]$ can be written as

$$p(x) = \begin{cases} c + t(L - x) & x \in \left[0, \frac{L}{2}\right], \\ c + tx & x \in \left[\frac{L}{2}, L\right]. \end{cases}$$

Figure 4 shows the marginal costs, the equilibrium prices, and the gross profits from the market area between A and B.

For a given good whose fixed cost is y, the firms producing the same given good should be apart equally. Furthermore, since the gross profit of any of these given firms with a market area of L is $\frac{tL^2}{2}$, and firms enter with nonnegative profits, we present the following lemma as the derivation of an SPNE for any arbitrary good.

Lemma 2.1 *Fix the level of fixed cost y and define $\underline{L}(y)$ as the solution to the zero-profit condition $t\left[\underline{L}(y)\right]^2/2 = y$. Thus, $\underline{L}(y) = \sqrt{2y/t}$. There is a continuum of equilibria in which one firm is located at every point in $\{x + nL\}_{n=-\infty}^{\infty}$, where $L \in [\underline{L}(y), 2\underline{L}(y))$ and $x \in [0, \underline{L}(y))$.*

There exists a continuum of equilibria because any L in the interval $\left[\underline{L}(\bar{y}), 2\underline{L}(y)\right)$ is an equilibrium distance; $L \geq \underline{L}(y)$ implies that all firms earn a nonnegative profit (no exit), whereas $L < 2\underline{L}(y)$ implies that any new entrant between any two existing firms must earn a negative profit (no entry).

2.2 *Hierarchy Equilibrium*

An equilibrium is a collection of locations of firms, delivered price schedules, and farmers' consumption choices such that the allocation for each good y is an SPNE. In this section, we consider a *hierarchy equilibrium* in which the *hierarchy property* holds.

Definition 2.1 A *hierarchy equilibrium* is an equilibrium in which, at any production location, the set of goods produced must take the form $[\underline{y}, y]$ for some level of fixed cost y.

In a hierarchy equilibrium, there exists a decreasing sequence $\bar{y} = y_1 > y_2 > \cdots > y_I \geq \underline{y}$, for some $I \in \mathbb{N} \cup \{\infty\}$, denoting the cutoffs. A hierarchy equilibrium is said to satisfy the *central place property* if the market area of the firms producing $(y_{i+1}, y_i]$ is half that of the firms producing $(y_i, y_{i-1}]$.

Definition 2.2 A hierarchy equilibrium that satisfies the central place property is called a *central place hierarchy*.

In fact, any hierarchy equilibrium is a central place hierarchy. Let $L_1 = m\underline{L}[\bar{y}],\ m \in [1, 2)$, and $L_i = L_1/2^{i-1}$. Due to the hierarchy property, any production location produces goods in the range of $[\underline{y}, y_i]$ for some y_i so that it is called a layer-i city, and the cutoff y_i is given by the zero-profit condition:

$$y_i = \frac{tL_i^2}{2} \quad \forall\ 1 \leq i \leq I, \tag{2.1}$$

where the number of layers is:

$$I = \llcorner \frac{2\ln(m) + \ln(\bar{y}/\underline{y})}{2\ln(2)} + 1 \lrcorner, \tag{2.2}$$

Figure 5 depicts four layers of such location configuration in which we can see that both the hierarchy and central place property are satisfied and the foregoing construction is an equilibrium.

Proposition 2.2 (Central place hierarchy) *For each $L_1 = m\underline{L}(\bar{y}),\ m \in [1, 2)$, let $L_i = L_1/2^{i-1}$, $i \in [1, I]$, y_i be given by the zero-profit condition (2.1), and the number of layers I be given by (2.2). Fix an $x \in \mathbb{R}$, and set the grid for $(y_{i+1}, y_i]$ as $\{x + nL_i\}_{n=-\infty}^{\infty}$. Then, for each $m \in [1, 2)$, the location configuration so constructed is the unique hierarchy equilibrium and satisfies the central place property.*

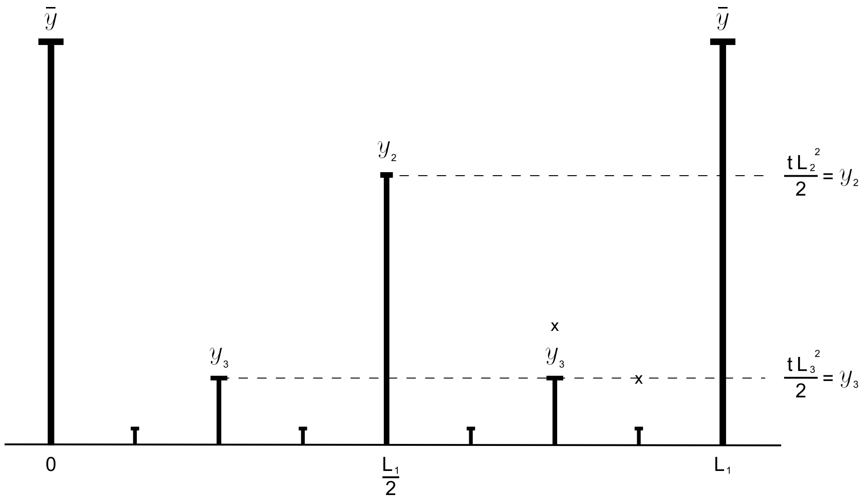

Fig. 5 A central place hierarchy. *Notes* The layer-i cities produce goods in $[\underline{y}, y_i]$. The cutoffs y_i are determined by the zero-profit conditions. The market areas for goods $(y_{i+1}, y_i]$ are half of that for $(y_i, y_{i-1}]$

3 Power Law for Cities

We now explain how a power law for cities emerges. In a central place hierarchy, the output of the firms in range $(y_{k+1}, y_k]$ is $L_k[F(y_k) - F(y_{k+1})]$. Define the size of a layer-i city by the total units produced in that city (as a measure of the level of economic activity):

$$Y_i = \sum_{k=i}^{I} L_k[F(y_k) - F(y_{k+1})]$$

Figure 6 illustrates the definitions of Y_i. The green (shaded with lines) and red (shaded with dots) areas represent the total quantity produced in a layer-1 and layer-2 city, respectively.

For every layer-1 city, there is one layer-2 city and 2^{i-2} layer-i cities. Thus, the total number of cities up to layer-i is

$$R_i = 1 + 1 + \sum_{k=3}^{i} 2^{k-2} = 2^{i-1}$$

Note that R_i represents the rank, by the rank-size rule, since the rank doubles from layer-i to the next layer-$i+1$, Zipf's law can be approximated if city size shrinks by around half from layer-i to the next layer. Similarly, if city size shrinks by an approximately constant fraction from any layer-i to the next, then the power law is

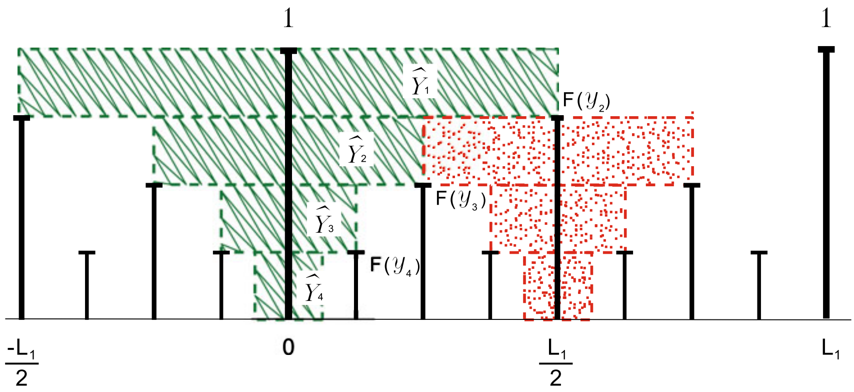

Fig. 6 City size. *Notes* The green (shaded with lines) and red (shaded with dots) areas denote the size of a layer-1 and layer-2 city, respectively. Both shaded areas are composed of rectangles, each of which represents the total production of the respective range of goods

approximated. It means that power laws can be generated by the fractal structures.[2] Therefore, a natural question arises: under what conditions can we guarantee such fractal structures?

There is, indeed, a simple but powerful condition that directly links central place hierarchies and the power law, regardless of the underlying economics behind that hierarchy. Given a central place hierarchy, the location patterns of cities of different layers are fixed, and different underlying economics matter only in relation to how the fractions of goods ($z_i = F(y_i)$) in the different layers are determined. The following proposition specifies the condition for the fractions of goods that renders the central place hierarchy a fractal structure.

Proposition 3.1 (Bounds on fraction ratios) *Suppose that there are infinitely many layers in a central place hierarchy. Let z_i denote the fraction of goods produced in a layer-i city, and let $\Delta_k = z_k - z_{k+1}$. Suppose that there is a $\delta > 0$ and a $\rho > 1$, such that for all $i \in \mathbb{N}$,*

$$\frac{\delta}{\rho} \leq \frac{\Delta_{i+1}}{\Delta_i} \leq \rho\delta.$$

Then,

$$\frac{1}{2}(\rho^{-1} - 1)\delta \leq \frac{Y_{i+1}}{Y_i} - \frac{\delta}{2} \leq \frac{1}{2}(\rho - 1)\delta. \tag{3.1}$$

Observe that δ is approximately the ratio of the increments (Δ_k) when ρ is approaching 1, hence, the slope is

[2] A fractal structure is a structure in which smaller parts of it resemble the entire structure.

$$\frac{\ln(R_{i+1}/R_i)}{\ln(Y_{i+1}/Y_i)} \approx \frac{\ln(2)}{\ln(\delta/2)} = -\frac{\ln(2)}{\ln(2) - \ln(\delta)}.$$

The Zipf's law requires only that the increments of the fraction of goods of two adjacent layers do not vary too much ($\delta \approx 1$), whereas the power law relaxes the ratio of increments between two layers from 1.

Besides the constrains on $z_i = F(y_i)$, the further question is how the behavior of $F(.)$ translates into a power law for cities. In order to answer this formally, we first need to introduce a few basic concepts of regular variation.

Definition 3.1 A measurable, positive function g is said to be regularly varying at zero (at infinity) if, for any $u > 0$, and for some $\alpha \in \mathbb{R}$,

$$\lim_{y \downarrow 0\ (\to \infty)} \frac{g(uy)}{g(y)} = u^{\alpha}.$$

If $\alpha = 0$, then g is said to be slowly varying. A function g is regularly varying with index α if and only if there exists a slowly varying function $\ell(y)$ such that

$$g(y) = y^{\alpha}\ell(y).$$

In what follows, $g \in RV_\alpha$ denotes that g is regularly varying at zero with index α. Suppose that $\underline{y} = 0$, and hence there are infinitely many layers. Recall from Proposition 2.2 that $y_{k+1} = y_k/4$ for all $k \geq 2$. Observe that the ratio between the increments between two layers can be written as

$$\delta_k \equiv \frac{\Delta_{k+1}}{\Delta_k} = \frac{F(y_{k+1}) - F(y_{k+2})}{F(y_k) - F(y_{k+1})} = \frac{1 - \frac{F(y_{k+1}/4)}{F(y_{k+1})}}{\frac{F(4y_{k+1})}{F(y_{k+1})} - 1}.$$

According to Definition 3.1, if $F \in RV_\alpha$, then in a small enough neighbourhood of 0, there are infinitely many k's such that

$$\delta_k \approx \frac{1 - \left(\frac{1}{4}\right)^{\alpha}}{4^{\alpha} - 1} = \left(\frac{1}{4}\right)^{\alpha}. \tag{3.2}$$

By Proposition 3.1, the power law is approximated with a tail index close to $1/(1 + 2\alpha)$.

A distribution function F on $(0, \bar{y}]$ can be regularly varying only with a nonnegative index α because an $F \in RV_\alpha$ on $(0, \bar{y}]$ with $\alpha < 0$ must be decreasing in a small neighbourhood of 0, which violates the requirement of a distribution function. However, a distribution function can be defined via a transformation of a non-increasing function $G \in RV_\alpha$ with $\alpha < 0$:

$$F(y) \equiv \frac{G(\underline{y}) - G(y)}{G(\underline{y}) - G(\bar{y})}, \tag{3.3}$$

where the domain of F is $[\underline{y}, \bar{y}]$ for some $\underline{y} > 0$. For a $\underline{y}$ that is close enough to zero, such an F behaves like a regularly varying function with a negative index α. This is because, for a $\underline{y}$ close enough to 0, there exists a sufficiently small neighbourhood of $\underline{y}$ such that, for all y_{k+1} in that neighbourhood:

$$\delta_k = \frac{F(y_{k+1}) - F(y_{k+2})}{F(y_k) - F(y_{k+1})} = \frac{G(y_{k+2}) - G(y_{k+1})}{G(y_{k+1}) - G(y_k)} = \frac{1 - \frac{G(y_{k+1}/4)}{G(y_{k+1})}}{\frac{G(4y_{k+1})}{G(y_{k+1})} - 1} \approx \frac{1 - \left(\frac{1}{4}\right)^{\alpha}}{4^{\alpha} - 1} = \left(\frac{1}{4}\right)^{\alpha}.$$

In any case, when the index α associated with the distribution function is positive (negative), then the slope of the Zipf plot is smaller (greater) than 1. The following proposition summarises the foregoing discussion and provides statements based on the density functions.

Proposition 3.2 (Regularly varying distributions) *Let $\delta = (1/4)^{\alpha}$ and fix any $\rho > 1$. Then, for a sufficiently small $\underline{y} \geq 0$, there exists an integer $K > 0$ such that condition (3.1) holds for all layers $I \geq i \geq K$ (with the possibility that $I = \infty$), if one of the following conditions is met:*

(a) the distribution function of fixed cost $F \in RV_{\alpha}$ with $\alpha \in [0, \infty)$;
(b) $G \in RV_{\alpha}$ with $\alpha \in (-1/2, 0)$ such that F is defined by (3.3);
(c) the density function of fixed cost $f \in RV_{\alpha-1}$, for $\alpha \in (-1/2, \infty)$.

In all cases, the approximate slope of the Zipf plot, i.e., the plot of log of rank on log of size, is $-1/(1 + 2\alpha)$.

Therefore, the power law arises when $F(.)$ has a regularly varying right tail, which is rather general, as it includes several well-known, commonly used distributions, such as the Pareto, Weibull, F, Beta (which subsumes the uniform), and Gamma, which subsumes the Chi-square, exponential, and Erlang.

Note that the sizes of cities of the same layer are equal and such city size distribution is understood as an approximation to the power laws. To fully conform with the relatively smooth city size distribution, some random factors to create differences among cities of the same layer is needed. Indeed, all central-place models for the city size distribution would need this random disturbance. See, for example, Beckmann (1958, p. 245).

4 A Dynamic Programming Approach to Central Place Theory

In this section, taking hierarchy property as given, we consider the social planner's problem about how to construct the optimal city hierarchy via a dynamic programming formulation (Hsu et al. 2014). There must be one and only one immediate smaller city between two neighboring larger-sized cities in any optimal solution. Furthermore, if the fixed cost of setting up a city is a power function, then the immediate smaller city will lie in the middle, which confirms the locational pattern suggested by Christaller, and further provides a rationale for central place theory of city hierarchy.

4.1 The Sequence Problem

The basic environment is similar to Sect. 2.1, we further assume the *hierarchy property*: at any location x, if a good $z \le z_1$ is produced, then all $\tilde{z} \in [0, z]$ are also produced. We label a location that produces all goods up to z as a z-city. Denote the cost of setting up a z-city as $\Phi(z) \equiv \int_0^z \phi(u)\, du$. According to the hierarchy property, z also refers to a city's size. Assume that two *neighboring* z_1-cities are located at 0 and ℓ_1, respectively. Thus, the social planner needs to determine ℓ_1 first,[3] then taken ℓ_1 as given, to find the optimal *city hierarchy*, i.e., the locations and sizes of cities on the interval $(0, \ell_1)$.

Given ℓ_1, let the discrete set of cities on $(0, \ell_1)$ be denoted as

$$W \equiv \left\{ \begin{array}{c} \left(z_i, L_{z_i}, I\right) | z_i \in (0, z_1],\ i = 1, 2, \ldots, I,\ I \in \mathbb{N} \cup \{\infty\},\ z_i > z_{i+1}, \\ L_{z_i} \text{ is the set of locations of } z_i\text{-cities} \end{array} \right\}.$$

That is, z_i is the i-th largest among all cities on $(0, \ell_1)$. For now, there may be multiple z_i cities, and L_{z_i} and $|L_{z_i}|$ denote the set of locations and the number of z_i -cities on $(0, \ell_1)$, respectively. The number I is the number of *layers* of cities, and I can be (countably) infinite. The optimization problem, given ℓ_1, is to search for a city hierarchy W that minimizes the per capita cost of production to serve every consumer a unit of each good in $[0, z_1]$:

$$C^*(\ell_1, z_1) \equiv \inf_W \frac{1}{\ell_1} \left[\sum_{z_i} |L_{z_i}| \Phi(z_i) + \text{total transport cost} \right], \tag{4.1}$$

4.2 The Dynamic Programming Problem

The following two lemmas provide key characteristics of an optimal hierarchy that enables us to set up the planner's problem as a dynamic programming problem.

Lemma 4.1 *It is never optimal to have an interval without any city in it.*

[3] We do not discuss how to decide the optimal distance ℓ_1 here, since our focus is the city hierarchy between any two largest cities. Nevertheless, this is very crucial question for presenting a complete and meaningful model, please see Sect. 3.5 in Hsu et al. (2014) for the solution.

Proof Consider adding a z'-city in the middle in between with $z' \le z$. Then, the savings in transport cost per good is

$$2\int_0^{\ell/2} txdx - 4\int_0^{\ell/4} txdx = t\ell^2/8$$

so the he net saving from having a z'-city is given by

$$S\left(z';\ell\right) \equiv \int_0^{z'} \left[\frac{t\ell^2}{8} - \phi(y)\right] dy$$

Because ϕ is continuous and strictly increasing, and $\phi(0) = 0$, $S\left(z';\ell\right) > 0$ for sufficiently small $z' > 0$, given ℓ. The result follows from the fact that there always exists sufficiently small z' such that adding a z'-city improves the allocation. □

Lemma 4.2 *It is never optimal to have two cities of the same size $z' < z_1$ without a larger city in between.*

The intuition behind this lemma is that, without a larger city in between, the two neighboring cities cannot be the same size in the optimal solution. We can produce another good more with a infinitesimally higher set-up cost at one of the two cities in order to save the transport cost, in which case there always exists a better allocation whose net cost is less. A full proof can be found in Sect. 2.3 in Hsu et al. (2014).

Lemmas 4.1 and 4.2 indicate that in between two z_1-cities it is optimal to place one and only one immediate sub-city, which is denoted as a z_2-city. Notice that z_2-city is not necessarily located in the middle.[4] Let $\ell_{2,1}$ and $\ell_{2,2}$ represent the distances from the z_2-city to the z_1-city on the left and right side, respectively. When the values of z_2, $\ell_{2,1}$ and $\ell_{2,2}$ are chosen, the recursive nature of the problem becomes apparent: given z_2, $\ell_{2,1}$ and $\ell_{2,2}$, we search for the optimal solutions for endless iterated bifurcations as the one given z_1 and ℓ_1. Figure 7 illustrates the city building process for the first three rounds of bifurcations. In general, the i-th round of bifurcation involves setting up cities of sizes $z_{i+1,1}, z_{i+1,2}, \ldots, z_{i+1,K_i}$, where $K_i = 2^{i-1}$, which divides intervals of length $\ell_{i,1}, \ell_{i,2}, \ldots, \ell_{i,K_i}$, respectively. Formally, let $\boldsymbol{z}_1 \equiv \{z_1\}$, and for all $i \in \mathbb{N}$, let $\boldsymbol{\ell}_i \equiv \left\{\ell_{i,k}\right\}_{k=1}^{K_i}$ and $\boldsymbol{z}_{i+1} \equiv \left\{z_{i+1,k}\right\}_{k=1}^{K_i}$, where $\ell_{1,1} \equiv \ell_1$ and $z_{2,1} \equiv z_2$. We define

$$\Gamma_1\left(\ell_1, z_1\right) \equiv \Gamma\left(\ell_1, z_1\right) \equiv \left\{\left(\boldsymbol{\ell}_2, \boldsymbol{z}_2\right) | z_2 \in [0, z_1], \ell_{2,1}, \ell_{2,2} \in (0, \ell_1) \text{ and } \ell_{2,1} + \ell_{2,2} = \ell_1\right\}. \tag{4.2}$$

and for $i \ge 2$,

$$\Gamma_i\left(\boldsymbol{\ell}_i, \boldsymbol{z}_i\right) \equiv \left\{ \begin{array}{c} \left(\boldsymbol{\ell}_{i+1}, \boldsymbol{z}_{i+1}\right) | z_{i+1,2k-1}, z_{i+1,2k} \in \left[0, z_{i,k}\right] \text{ for all } k = 1, 2, \ldots, K_{i-1}, \\ \ell_{i+1,2k-1}, \ell_{i+1,2k} \in \left(0, \ell_{i,k}\right) \text{ and } \ell_{i+1,2k-1} + \ell_{i+1,2k} = \ell_{i,k}, \text{ for all } k = 1, 2, \ldots, K_i. \end{array} \right\}$$

[4] For example, let $t = z_1 = \ell_1 = 1$. Consider a discontinuous setup cost requirement function: for an arbitrarily small $e \in (0, 1)$, $\phi(y) = 1/13$ for $y \in [0, e]$ and $\phi(y) = 1$ for $y \in (e, 1]$. It is readily verified that, in between two z_1-cities, the per capita cost is minimized by evenly placing two immediate sub-cities with $z' = e$.

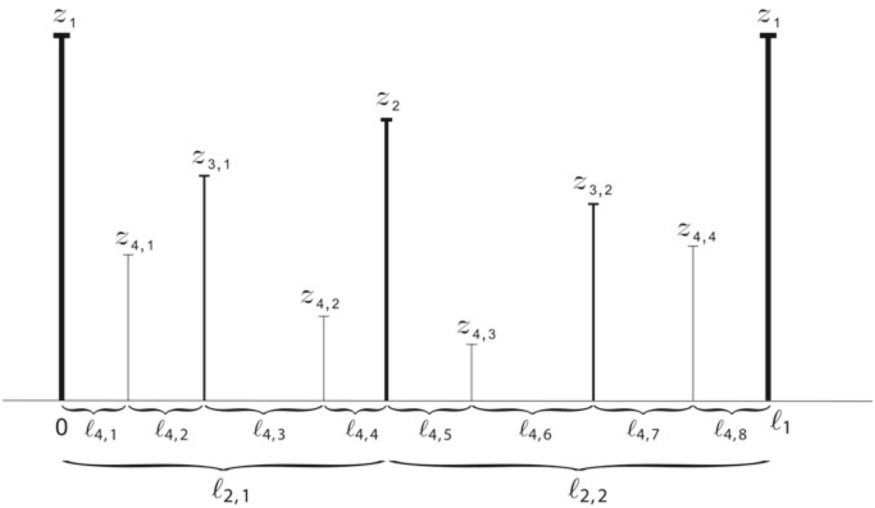

Fig. 7 Illustration of sequence problem

Then, define

$$\Pi(\ell_1, z_1) \equiv \left\{ (\boldsymbol{\ell}_i, \boldsymbol{z}_i)_{i=1}^{\infty} \mid (\boldsymbol{\ell}_{i+1}, \boldsymbol{z}_{i+1}) \in \Gamma_i (\boldsymbol{\ell}_i, \boldsymbol{z}_i), \text{ for all } i = 1, 2, \ldots \right\}.$$

Any $(\boldsymbol{\ell}, \boldsymbol{z}) \equiv (\boldsymbol{\ell}_i, \boldsymbol{z}_i)_{i=1}^{\infty} \in \Pi(\ell_1, z_1)$ is called a *feasible* sequence, given (ℓ_1, z_1).

Let ℓ be the distance between the two neighboring larger-sized z-cities, hence the distances of a immediate sub-city z' to the two cities are $\alpha\ell$ and $(1-\alpha)\ell$, where $\alpha \in (0, 1)$. The savings in transport costs for each good in $[0, z']$ is

$$s^1(\ell, \alpha) \equiv 2\int_0^{\frac{\ell}{2}} txdx - \left(2\int_0^{\frac{\alpha\ell}{2}} txdx + 2\int_0^{\frac{(1-\alpha)\ell}{2}} txdx\right) = \frac{t\ell^2}{2}\alpha(1-\alpha), \tag{4.3}$$

Then, the optimal magnitude of z' is determined by

$$s^1(\ell, \alpha) = \frac{t\ell^2}{2}\alpha(1-\alpha) = \phi(z'). \tag{4.4}$$

The left-hand side of (4.4) is the savings in transport costs when increasing z' marginally, whereas the right-hand side is the corresponding setup cost. If z' is low such that $\phi(z') < \frac{t\ell^2}{2}\alpha(1-\alpha)$, it incurs positive net savings (savings in transport costs net of setup costs) by increasing z'. Similarly, when $\phi(z') > \frac{t\ell^2}{2}\alpha(1-\alpha)$, one can improve the allocation by decreasing z'. In sum, Lemmas 4.1 and 4.2 and (4.4) imply that in any optimal city hierarchy the following constraint holds:

$$\left\{\begin{array}{l} z_{i+1,2k-1}, z_{i+1,2k} \in \left(0, z_{i,k}\right) \\ \ell_{i+1,2k-1}, \ell_{i+1,2k} \in \left(0, \ell_{i,k}\right), \ell_{i+1,2k-1} + \ell_{i+1,2k} = \ell_{i,k} \\ z_{i+1,k} = \phi^{-1}\left(\frac{t}{2}\ell_{i+1,2k-1}\ell_{i+1,2k}\right) \end{array}\right\}. \tag{4.5}$$

Equivalently, any optimal city hierarchy is associated with a sequence $\boldsymbol{\alpha} = \left\{\alpha_{i,k}\right\}$ such that $\ell_{i+1,2k-1} = \alpha_{i,k}\ell_{i,k}$ (hence $\ell_{i+1,2k} = \left(1 - \alpha_{i,k}\right)\ell_{i,k}$) and (4.5) holds.

Note that in defining the choice set of $(\boldsymbol{\ell}, \boldsymbol{z}) \equiv (\boldsymbol{\ell}_i, z_i)_{i=1}^{\infty}$ by Γ_i and Π above, we leave (4.4) implicit and take the closure of $\left(0, z_{i,k}\right)$. According to Lemmas 4.1 and 4.2, we know that situations in which $z_{i+1,2k-1}$ or $z_{i+1,2k}$ equals 0 or $z_{i,k}$ are never optimal (except possibly for $i = 1$), but we do not lose any generality by including this possibility. When the choice of $z_{i+1,2k-1}$, according to (4.4) and given $\ell_{i,2k-1}$, is such that $z_{i+1,2k-1} > z_{i,k}$, one can always relabel i, k to ensure that the constraint $z_{i+1,2k-1}, z_{i+1,2k} \in \left[0, z_{i,k}\right]$ is obeyed. Thus, the choice set defined by Π encompasses all possible candidates for an optimal city hierarchy. In other words, any sequence $(\boldsymbol{\ell}, \boldsymbol{z})$ that satisfies all constraints in (4.5) is included in $\Pi\left(\ell_1, z_1\right)$. If one would like to make the constraint (4.4) explicit, one could redefine Γ_i by replacing $z_{i+1,2k-1}, z_{i+1,2k} \in \left[0, z_{i,k}\right]$ with

$$z_{i+1,2k-1} = \min\left\{\phi^{-1}\left(\frac{t}{2}\ell_{i+1,4k-1}\ell_{i+1,4k-2}\right), z_{i,k}\right\}$$

and

$$z_{i+1,2k} = \min\left\{\phi^{-1}\left(\frac{t}{2}\ell_{i+1,4k-3}\ell_{i+1,4k-4}\right), z_{i,k}\right\}.$$

Suppose the social planner has two z-cities with distance ℓ and nothing in between them, then the total cost in this interval of ℓ is

$$A\left(\ell, z\right) \equiv \Phi\left(z\right) + \frac{zt\ell^2}{4}.$$

Note that only one setup cost of a z-city is counted in this definition. When a z'-city divides an interval of ℓ bounded by two cities producing at least up to z, the total cost for the range of goods $(z', z]$ is given by $A\left(\ell, z\right) - A\left(\ell, z'\right) = \Phi\left(z\right) - \Phi\left(z'\right) + \left(z - z'\right)t\ell^2/4$. We can view the per capita cost for the goods $[0, z_1]$ on ℓ_1 as the sum of the per capita cost of different ranges of goods on different market areas within ℓ_1. Namely, the sequence problem is

$$C^*(\ell_1, z_1) \equiv \inf_{\substack{(\boldsymbol{\ell}, \boldsymbol{z}) \in \Pi(\ell_1, z_1), \\ z_1 > 0 \text{ given.}}} \frac{1}{\ell_1}\left[\begin{array}{l} A\left(\ell_1, z_1\right) - A\left(\ell_1, z_2\right) \\ + \sum_{i=2}^{\infty}\sum_{k=1}^{K_{i-1}}\left[\begin{array}{c} A\left(\ell_{i,2k-1}, z_{i,k}\right) - A\left(\ell_{i,2k-1}, z_{i+1,2k-1}\right) \\ + A\left(\ell_{i,2k}, z_{i,k}\right) - A\left(\ell_{i,2k}, z_{i+1,2k}\right) \end{array}\right] \end{array}\right]. \tag{SP}$$

Besides sequence problem, it can also be represented by a dynamic programming problem. Given state variables ℓ and z, the social planner needs to decide the size and location of the immediate sub-city, z'-city. Denote the length of the intervals to the left/right of z'-city as ℓ_l/ℓ_r. Then, $\ell_l + \ell_r = \ell$. Alternatively, let $\ell_l = \alpha\ell$ and $\ell_r = (1-\alpha)\,\ell$ for $\alpha \in (0,1)$. We present the following dynamic programming problem.

$$\begin{aligned} C(\ell,z) &= \inf_{\ell_l,\ell_r\in(0,\ell),\ell_l+\ell_r=\ell,z'\in[0,z]} \frac{1}{\ell}\left[A(\ell,z) - A\left(\ell,z'\right) + \ell_l C(\ell_l,z') + \ell_r C(\ell_r,z')\right] \\ &= \inf_{\alpha\in(0,1),z'\in[0,z]} \frac{1}{\ell}[A(\ell,z) - A\left(\ell,z'\right)] + \alpha C(\alpha\ell,z') + (1-\alpha)C((1-\alpha)\ell,z'). \end{aligned} \tag{DP}$$

Equivalently, we can consider infimum of the total cost of all the goods $[0, z_1]$ on the interval of length ℓ_1. By defining $D^*(\ell_1, z_1) = \ell_1 C^*(\ell_1, z_1)$, and $D(\ell, z) = \ell C(\ell, z)$, it transforms (SP) and (DP) to (SP^D) and (DP^D) as following:

$$D^*(\ell_1,z_1) = \inf_{\substack{(\boldsymbol{\ell},z)\in\Pi(\ell_1,z_1),\\ z_1>0\text{ given.}}} \begin{array}{c} A(\ell_1,z_1) - A(\ell_1,z_2) \\ +\sum_{i=2}^{\infty}\sum_{k=1}^{K_{i-1}} \left[\begin{array}{c} A\left(\ell_{i,2k-1},z_{i,k}\right) - A\left(\ell_{i,2k-1},z_{i+1,2k-1}\right) \\ +A\left(\ell_{i,2k},z_{i,k}\right) - A\left(\ell_{i,2k},z_{i+1,2k}\right) \end{array}\right], \end{array} \tag{SP^D}$$

and

$$D(\ell,z) = \inf_{\alpha\in(0,1),z'\in[0,z]} A(\ell,z) - A\left(\ell,z'\right) + D\left(\alpha\ell,z'\right) + D\left((1-\alpha)\,\ell,z'\right). \tag{DP^D}$$

For showing the equivalence between the sequence problem(SP^D) and the respective dynamic programming problem (DP^D), i.e., the *principle of optimality*, we present the following corollary.

Corollary 1 *For any two positive real numbers ℓ_1 and z_1, let $X = [0, \ell_1] \times [0, z_1]$, and let $\mathcal{D}(X)$ denote the set of all real-valued continuous functions $d : X \to \mathbb{R}_+$ such that*

$$0 \le d(\ell,z) \le A(\ell,z). \tag{4.6}$$

Then a feasible sequence $\left(\boldsymbol{\ell}^*, \boldsymbol{z}^*\right) \in \Pi(\ell_1, z_1)$ attains the infimum in (SP^D) if and only if it satisfies (DP^D) recursively, i.e.,

$$D^*(\ell^*_{i,k},z^*_{i,k}) = A\left(\ell^*_{i,k},z^*_{i,k}\right) - A\left(\ell^*_{i,k},z^*_{i+1,k}\right) + D^*(\ell^*_{i+1,2k-1},z^*_{i+1,k}) + D^*(\ell^*_{i+1,2k},z^*_{i+1,k}).$$

Now, the question left is: under what conditions can we be sure that there is a unique solution to (SP^D) as well as (DP^D)?

Corollary 2 *Let $\mathcal{D}(X)$ be given by Corollary 1. Let $T : \mathcal{D}(X) \to \mathcal{D}(X)$ be given by, for each $d \in \mathcal{D}(X)$,*

$$Td(\ell,z) \equiv \inf_{\alpha\in(0,1),z'\in[0,z]} A(\ell,z) - A\left(\ell,z'\right) + d(\alpha\ell,z') + d((1-\alpha)\ell,z'). \tag{4.7}$$

Then,

(i) Td *is continuous. Hence,* T *is a self-mapping on* $\mathcal{D}(X)$.
(ii) The minimum is attained; so inf *in the definition of* T *in (4.7) can be replaced with* min. *Moreover, the set of minimizers is an upper hemi-continuous correspondence on* X.
(iii) D^* *is the unique solution to* (DP^D) *in* $\mathcal{D}(X)$ *and hence the unique fixed point of the mapping* T *on* $\mathcal{D}(X)$.
(iv) For any $d \in \mathcal{D}(X)$, *the sequence* $\{T^n d\}$ *converges to* D^*.

Corollary 2 allows the numerical solution for any arbitrary ϕ/or any initial guess. We have implemented this iterative method via Matlab in which the approximation works well. The proofs of corollary 1 and 2 can be seen from Sect. 3.2 and 3.3 in Hsu et al. (2014), which starts from the routine pursue of principle of optimality as in almost all the dynamic programming problems. For a very intriguing and thorough introduction of such recursive problems in economics, please refer to Ljungqvist and Sargent (2012).

4.3 The Central Place Property

It turns out that the central place property holds when the setup cost is a power function: $\phi(z) = az^b$, for $a > 0$ and $b > 0$. Under this functional form, the total setup cost for a z-city is $\Phi(z) = \frac{a}{b+1} z^{b+1}$. The power function assumption of ϕ, in fact, means that the distribution of setup costs across goods is also a power function. Let Y denote the random variable of setup cost for a good. Then, for $y \in [0, \phi(z_1)]$,

$$\Pr[Y \leq y] = \frac{\phi^{-1}(y)}{z_1} = \frac{1}{z_1}\left(\frac{y}{a}\right)^{\frac{1}{b}}.$$

As shown in the hierarchy equilibrium, Sect. 2.2, this distribution of setup cost is a prototype of a class of distributions that leads to a power law distribution of city size.

Recall that it is possible that the optimal $z_2 = z_1$ if ℓ_1 is too large. Note from (4.4) that savings $s^1(\ell_1, \alpha)$ is bounded by $s^1(\ell_1, 1/2) = t\ell_1^2/8$. Define $\bar{\ell}(z)$ by the solution of ℓ in the following equation.

$$\frac{t\ell^2}{8} = \phi(z). \tag{4.8}$$

Then, for any $\ell_1 < \bar{\ell}(z_1)$, optimal $z_2 < z_1$, and thus the two z_1-cities with distance ℓ_1 are neighboring. For the rest of the analyses in this paper, we impose the condition that $\ell_1 < \bar{\ell}(z_1)$.

Proposition 4.3 *Suppose that* $\ell_1 < \bar{\ell}(z_1)$, *where* $\bar{\ell}(z)$ *is defined as the solution to (4.8). Suppose that the setup cost function* $\phi(y) = az^b$, *for positive constants* a *and* b. *Then, the central place property holds.*

Proof For ease of presentation, let $a = 1$. A general $a > 0$ does not change the result. From (4.4),

$$z' = \left(\frac{t\ell^2}{2}\alpha(1-\alpha)\right)^{\frac{1}{b}} = \left(\frac{t}{2}\alpha(1-\alpha)\right)^{\frac{1}{b}} \ell^{\frac{2}{b}} \equiv \kappa(\alpha)\,\ell^{\frac{2}{b}}. \tag{4.9}$$

Equation (4.9) implicitly assumes that $z' < z$. Recall that Lemma 4.2 rules out $z_{i+1,2k-1} = z_{i,k}$ or $z_{i+1,2k-1} = z_{i,k}$ as an optimal solution, and hence (4.9) is necessarily true for all optimal choices of $z_{i,k}$, except possibly for $i = 2$. However, the constraint $\ell_1 < \bar{\ell}(z_1)$ ensures that optimal $z_2 < z_1$.

Recall from (4.5) that there is a sequence $\boldsymbol{\alpha} = \{\alpha_{i,k}\}$ associated with any sequence $(\boldsymbol{\ell}, \boldsymbol{z}) \in \Pi(\ell_1, z_1)$. The fact that the optimal solution of z' is separable in ℓ and α implies that, except for z_1, we can write $z_{i,k} = \ell_1^{2/b} h_{i,k}(\boldsymbol{\alpha})$ and $\ell_{i,k} = \ell_1 g_{i,k}(\boldsymbol{\alpha})$, for some functions $h_{i,k}$ and $g_{i,k}$. Thus, both $A(\ell_{i,2k-1}, z_{i,k}) - A(\ell_{i,2k-1}, z_{i+1,2k-1})$ and $A(\ell_{i,2k}, z_{i,k}) - A(\ell_{i,2k}, z_{i+1,2k})$ are multiplicatively separable in $\ell_1^{2(b+1)/b}$ and some functions of $\boldsymbol{\alpha}$. Thus, for some function H, (SP) can be rewritten as

$$C^*(\ell_1, z_1) = \frac{z_1^{b+1}}{(b+1)\,\ell_1} + \frac{z_1 t \ell_1}{4} + \ell_1^{\frac{b+2}{b}} H(\boldsymbol{\alpha}^*).$$

By Corollarys 1 and 2, an optimal $\boldsymbol{\alpha}^*$ exists, and as such, $H(\boldsymbol{\alpha}^*)$ is well defined. Note that $H(\boldsymbol{\alpha}^*) < 0$, and $\ell_1^{(b+2)/b}|H(\boldsymbol{\alpha}^*)|$ is the per capita savings from building the optimal city hierarchy. Given the equivalence between (SP) and (DP), observe that the negative of per capita savings from having an optimal city hierarchy in an interval of ℓ is given by

$$\tilde{S}(\ell, z) \equiv C(\ell, z) - \frac{A(\ell, z)}{\ell} = C(\ell, z) - \frac{z^{b+1}}{(b+1)\,\ell} - \frac{zt\ell}{4} = \ell^{\frac{b+2}{b}} H(\boldsymbol{\alpha}^*). \tag{4.10}$$

This says that the $\tilde{S}$ function is homogenous of degree $(b+2)/b$ in ℓ and independent of z. With a little abuse of notation, we write $\tilde{S}(\ell) = \tilde{S}(\ell, z)$. Given (4.9) and (4.10), (DP) can be rewritten as

$$\tilde{S}(\ell) = \min_{z' \in (0,z), \alpha \in (0,1)} \frac{A(\alpha\ell, z') + A((1-\alpha)\ell, z') - A(\ell, z')}{\ell} + \left[\alpha^{\frac{2(b+1)}{b}} + (1-\alpha)^{\frac{2(b+1)}{b}}\right]\tilde{S}(\ell) \tag{4.11}$$

Thus,

$$\tilde{S}(\ell) = \frac{-b}{b+1}\left(\frac{t}{2}\right)^{\frac{b+1}{b}} \ell^{\frac{b+2}{b}} \max_{\alpha \in (0,1)} \frac{[\alpha(1-\alpha)]^{\frac{b+1}{b}}}{1 - \alpha^{\frac{2(b+1)}{b}} - (1-\alpha)^{\frac{2(b+1)}{b}}}. \tag{4.12}$$

We show in the separate appendix in Hsu et al. (2014) that the unique solution to the maximization problem in (4.12) is $\alpha = 1/2$. □

Observe that the optimal sequence $\boldsymbol{\alpha}^*$ does not depend on ℓ_1. The recursive nature implies that for all i, k, the optimal sequence in the interval of $\ell_{i,k}$, i.e., $\left\{\alpha_{i',k}\right\}_{i'\geq i}$, does not depend on the magnitude of $\ell_{i,k}$. Thus, under this power function distribution of setup costs, the optimal city hierarchy in any interval of $\ell_{i,k}$ resembles that of the entire one in ℓ_1. As Sect. 3 shows that this *scale-free* property gives the city hierarchy a *fractal structure*; specifically, the structure of the smaller part of the hierarchy resembles that of the larger.

5 Concluding Remarks

This chapter presents a parsimonious model in which central place hierarchies arise from both equilibria and optimal allocations. In the equilibrium model, we show that hierarchy property can arise as an equilibrium outcome. Even though there are possible deviations from the hierarchy property that could still constitute an equilibrium, Hsu (2012) shows that such deviations can be ruled out by adding in the home market effect of the central places. In the social planner's problem, the problem of location choice is complex, but we show that if the distribution of the fixed cost follows a power function, the central place property holds. As such distribution is also regularly varying, a power law also emerges from an optimal allocation.

One potential criticism is that the primary industry (agriculture, fishing, forestry, mining, etc.) and its employment have become a small fraction of the economy in developed countries, and thus central place theory is not quite relevant for these countries. Such a criticism is, however, not well grounded because the reason why a central place hierarchy emerges has nothing to do with the number/fraction of immobile consumers in the economy. It is the fact that they are immobile and dispersed over the entire geographic space that prompts the central places that serve them. A smaller relative size of the primary industry may make the cities and towns more sparse in spacing and smaller in their sizes, but it does not qualitatively alter the fractal structure of the central place hierarchy. Hence, it does not affect the power law for cities either.

Recall that the hierarchy property posits that larger cities are more diverse not only because they have more industries than smaller cities but also because they specialise in industries with more scale economies. This is arguably a reasonable view, especially when we look at industries in broader classifications. Mori et al. (2008) and Landman et al. (2011) find that the hierarchy property holds well for 3-digit JSIC but dissipates for 4-digit JSIC (the finest Japanese industrial classification). These findings hint that, at the finest levels, specialisation matters to the extent that heavy industries can be located in small towns due to numerous factors outside central place theory. The development of a comprehensive theory that produces more realistic patterns of diversity and specialisation is a desirable direction for future research.

References

Arlinghaus, S. L. (1985). Fractals take a central place. *Geografiska Annaler: Series B, Human Geography*, 67(2):83–88.

Arlinghaus, S. L. and Arlinghaus, W. C. (1989). The fractal theory of central place geometry: a diophantine analysis of fractal generators for arbitrary löschian numbers. *Geographical Analysis*, 21(2):103–121.

Batty, M. and Longley, P. A. (1994). *Fractal cities: a geometry of form and function*. Academic press.

Beckmann, M. J. (1958). City hierarchies and the distribution of city size. *Economic Development and Cultural Change*, 6(3):243–248.

Champernowne, D. G. (1953). A model of income distribution. *The Economic Journal*, 63(250):318–351.

Chen, Y. (2011). Fractal systems of central places based on intermittency of space-filling. *Chaos, solitons & fractals*, 44(8):619–632.

Chen, Y. (2014). Multifractals of central place systems: Models, dimension spectrums, and empirical analysis. *Physica A: Statistical Mechanics and its Applications*, 402:266–282.

Chen, Y. and Zhou, Y. (2006). Reinterpreting central place networks using ideas from fractals and self-organized criticality. *Environment and Planning B: Planning and Design*, 33(3):345–364.

Christaller, W. (1933). *Central Places in Southern Germany*, translated by C. W. Baskin (1966), Englewood Cliffs, NJ: Prentice-Hall.

Dittmar, J. (2011). Cities, markets, and growth: the emergence of Zipf's law. Working paper.

Eaton, B. C. and Lipsey, R. G. (1982). An economic theory of central places. *The Economic Journal*, 92(365):56–72.

Eeckhout, J. (2004). Gibrat's law for (all) cities. *American Economic Review*, 94(5):1429–1451.

Frankhauser, P. (1998). The fractal approach. A new tool for the spatial analysis of urban agglomerations. *Population: An English Selection*, pages 205–240.

Fujita, M., Krugman, P. R., Venables, A. J. and Fujita, M. (1999). *The Spatial Economy: Cities, Regions and International Trade*, volume 213. Wiley Online Library.

Gabaix, X. (1999). Zipf's law for cities: an explanation. *The Quarterly Journal of Economics*, 114(3):739–767.

Gabaix, X. and Ioannides, Y. M. (2004). The evolution of city size distributions. In *Handbook of Regional and Urban Economics*, volume 4, pages 2341–2378. Elsevier.

Hales, T. C. (2001). The honeycomb conjecture. *Discrete & Computational Geometry*, 25(1):1–22.

Hsu, W.-T. (2012). Central place theory and city size distribution. *The Economic Journal*, 122(563):903–932.

Hsu, W.-T., Holmes, T. J. and Morgan, F. (2014). Optimal city hierarchy: A dynamic programming approach to central place theory. *Journal of Economic Theory*, 154:245–273.

Landman, B. A., Huang, A. J., Gifford, A., Vikram, D. S., Lim, I. A. L., Farrell, J. A., Bogovic, J. A., Hua, J., Chen, M., Jarso, S., et al. (2011). Multi-parametric neuroimaging reproducibility: a 3-t resource study. *Neuroimage*, 54(4):2854–2866.

Lederer, P. J. and Hurter, A. P. Jr. (1986). Competition of firms: Discriminatory pricing and location?, *Econometrica*, 54(3):623–640

Ljungqvist, L. and Sargent, T. J. (2012). *Recursive Macroeconomic Theory*. The MIT Press.

Mori, T., Nishikimi, K. and Smith, T. E. (2008). The number-average size rule: a new empirical relationship between industrial location and city size. *Journal of Regional Science*, 48(1):165–211.

Quinzii, M. and Thisse, J.-F. (1990). On the optimality of central places. *Econometrica: Journal of the Econometric Society*, pages 1101–1119.

Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika*, 42(3/4):425–440.

Tabuchi, T. and Thisse, J.-F. (2006). Regional specialization, urban hierarchy, and commuting costs. *International Economic Review*, 47(4):1295–1317.

Tabuchi, T. and Thisse, J.-F. (2011). A new economic geography model of central places. *Journal of Urban Economics*, 69(2):240–252.