

## Document de treball de l'IEB 2010/41

### UNEVEN LANDSCAPES AND THE CITY SIZE DISTRIBUTION

**Sanghoon Lee, Qiang Li**

Cities and Innovation

Document de  
treball de l'IEB

**UNEVEN LANDSCAPES AND  
THE CITY SIZE DISTRIBUTION**

Sanghoon Lee, Qiang Li

The **IEB** research program in **Cities and Innovation** aims at promoting research in the Economics of Cities and Regions. The main objective of this program is to contribute to a better understanding of agglomeration economies and 'knowledge spillovers'. The effects of agglomeration economies and 'knowledge spillovers' on the Location of economic Activities, Innovation, the Labor Market and the Role of Universities in the transfer of Knowledge and Human Capital are particularly relevant to the program. The effects of Public Policy on the Economics of Cities are also considered to be of interest. This program puts special emphasis on applied research and on work that sheds light on policy-design issues. Research that is particularly policy-relevant from a Spanish perspective is given special consideration. Disseminating research findings to a broader audience is also an aim of the program. The program enjoys the support from the **IEB-Foundation**.

The **Barcelona Institute of Economics (IEB)** is a research centre at the University of Barcelona which specializes in the field of applied economics. Through the **IEB-Foundation**, several private institutions (Caixa Catalunya, Abertis, La Caixa, Gas Natural and Applus) support several research programs.

Postal Address:

Institut d'Economia de Barcelona

Facultat d'Economia i Empresa

Universitat de Barcelona

C/ Tinent Coronel Valenzuela, 1-11

(08034) Barcelona, Spain

Tel.: + 34 93 403 46 46

Fax: + 34 93 403 98 32

[ieb@ub.edu](mailto:ieb@ub.edu)

<http://www.ieb.ub.edu>

The IEB working papers represent ongoing research that is circulated to encourage discussion and has not undergone a peer review process. Any opinions expressed here are those of the author(s) and not those of IEB.

**UNEVEN LANDSCAPES AND  
THE CITY SIZE DISTRIBUTION \***

Sanghoon Lee, Qiang Li

**ABSTRACT:** This paper proposes a new explanation for Zipf's law often observed in the top tail of city size distribution. We show that Zipf's law can emerge if city size can be expressed as a product of multiple random factors. Each of the factors need not generate Zipf's law by itself. The key implication is that we cannot reject a model simply because the model does not generate Zipf's law. A single model, typically representing only one factor, may not generate Zipf's law, but if we have many such models together as in reality, Zipf's law may emerge.

JEL Codes: D39, R12

Keywords: City size distribution, Zipf's law, rank-size rule, log-normal distribution.

Sanghoon Lee  
Sauder School of Business  
University of British Columbia  
2053 Main Mall, Vancouver  
BC V6T 1Z2, Canada,  
Email: [sanghoon.lee@sauder.ubc.ca](mailto:sanghoon.lee@sauder.ubc.ca)

Qiang Li  
School of Public Economics and  
Administration  
Shanghai University of Finance and  
Economics  
777 Guoding Road  
Shanghai 200433, China,  
Email: [li.qiang@mail.shufe.edu.cn](mailto:li.qiang@mail.shufe.edu.cn)

---

\* We are very grateful to Tom Holmes, Gilles Duranton, Ralph Winter, Nate Schiavone, Kristian Behrens, and David Cuberes for valuable comments. We also thank seminar participants at University of British Columbia, Yonsei University, the Barcelona Institute of Economics (IEB), Shanghai University of Finance and Economics, and the 2010 Econometric Society World Congress.

# 1 Introduction

City size distribution follows a robust regular pattern. Zipf’s law claims that the population size of a city is inversely proportional to its rank: the second largest city in a country is about a half the size of the largest city, the third largest city is about a third the size of the largest city, and so forth. Zipf’s law has repeatedly been shown to hold in the top tails of city size distribution across different countries and periods (e.g., Rosen and Resnick, 1980; Dobkins and Ioannides, 2001; Ioannides and Overman, 2003; Gabaix and Ioannides, 2004; Soo, 2005).

This paper proposes a new explanation for this empirical regularity, i.e. Zipf’s law in the top tail of city size distribution. Our model assumes that numerous random factors affect city size (e.g., industry composition, road networks, climates, geographic constraints such as mountains and waters, human capital, zoning restriction, etc.), and predicts city size as a product of these factors. We prove that the city size distribution converges to the log-normal distribution by applying the central limit theorem (after a log transformation). The log-normal distribution is, as shown in Eeckhout (2004) and Eeckhout (2009), consistent with Zipf’s law in the top tail. Since modern central limit theorems require only weak conditions, our result applies quite generally; the random factors need not follow any specific distribution, different random factors can come from different distributions, and the factors may be correlated with each other to some degree.

Outside the top tail there is ongoing debate over what distribution best describes the city size distribution. Eeckhout (2004) argues that the city size distribution for all cities follows the log-normal distribution. Rozenfeld, Rybski, Gabaix and Makse (forthcoming) argue that Zipf’s law holds for cities greater over 12,000 population size. Our model generates the log-normal distribution as Eeckhout (2004). However, given the current debate over the city size distribution for all cities, we focus on Zipf’s law in the top tail which has repeatedly been confirmed in the literature.

A number of other explanations have been proposed to explain the empirical city size

distribution. The workhorse in this literature is the dynamic random growth process  $\rightarrow$ (e.g., Simon, 1955; Gabaix, 1999; Eeckhout, 2004; Duranton, 2006, 2007; Rossi-Hansberg and Wright, 2007; Córdoba, 2008); if the growth rate of a city is independent of its size (i.e., Gibrat’s law holds), city size distribution converges to the log-normal distribution or the Zipf distribution with additional conditions.<sup>1</sup> Two static models have been offered as well: Hsu (2009) uses the central place theory and Behrens, Duranton and Robert-Nicoud (2010) use human capital distribution across cities.

Our model is a static version of the random growth models. The random shocks are aggregated in the cross section instead of time. However, being a static model yields unique interpretations and implications. First, the random shocks in our model represent different factors affecting city size, and we show that Zipf’s law may emerge from the interaction of the multiple factors even when each factor does not follow the Zipf distribution. This leads to an important message: one cannot use Zipf’s law to test a model of cities.<sup>2</sup> Classical urban economics models such as Henderson (1974) have sometimes been criticized because they do not generate Zipf’s law (e.g., Krugman, 1996; Gabaix, 1999). However, a typical economic model focuses on one economic force it aims to deliver. A single model alone may not generate Zipf’s law, but when we have many such models together as in reality, Zipf’s

---

<sup>1</sup>Simon (1955) proposes a random growth mechanism to explain the Pareto distribution in city size distribution. Gabaix (1999) shows that Gibrat’s law (i.e., the growth rate of a city being independent of its size) with a lower bound on city size can lead to Zipf’s law (i.e., Pareto distribution with coefficient 1) in the steady state. Eeckhout (2004) argues that the city size distribution for all cities follows the log-normal distribution and provides an economic model that generates it. Duranton (2006) provides an economic foundation to Simon (1955) using the endogenous growth model with product proliferation developed by Romer (1990). Duranton (2007) uses small industry-level shocks to explain fast changes in industry location, slow changes in a city’s relative ranking in population, and very stable city size distribution. Rossi-Hansberg and Wright (2007) propose a dynamic model with endogenous city formation. Their model can generate not only Zipf’s law but also often-observed deviations. Córdoba (2008) translates the Gibrat’s law into more economically meaningful restrictions about preferences, technologies, and the dynamic shocks.

<sup>2</sup>We thank Gilles Duranton for guiding us to this implication.

law may emerge.

Second, our theory does not require Gibrat’s law. Instead, our theory requires that city size can be expressed as a product of multiple factors. This puts a restriction on the shape of city production function as Gibrat’s law in the random growth models puts a restriction on the dynamic growth pattern across cities. A product, as compared to a sum, implies that there is complementarity among the factors. In order for a city to become large, it has to do well overall. If a city fails in one factor (e.g., climate), it can severely limit city size.

In addition to the complementarity among the factors, our theory requires that city size is determined by numerous small factors in order to apply the central limit theorem. If city size is determined by only a few dominant factors, our theory may not work. We examine this issue by running a simulation. The result suggests that Zipf’s Law can emerge in the top tail even when there is only a small number of factors.

Our model also differs from other urban economics models in that we assume different cities are endowed with exogenously different physical attributes such as rivers and climate. We use them as a subset of the random factors affecting city size. The other models assume that all locations are ex-ante identical and endogenously generate city size distribution.

The approach of using heterogenous natural features to explain the empirical city size distribution was first suggested by Krugman (1996). Our model develops the idea in the following ways. First, we provide an explicit economic model. Second, our model extends the randomness beyond just natural features, to include randomness in other man-made features such as industry composition, road networks; the randomness in these features can happen due to the randomness in policy making, big firms’ location decisions, etc.

Even though the heterogenous features alone can generate the differences in city size, it is unquestionable that the agglomeration economies also play a major role in shaping city size distribution. Our model allows the agglomeration economies in each factor. The agglomeration economies amplify city size differences initiated by the heterogeneous features across cities.

The Zipf coefficient roughly measures the degree of concentration of population among cities (see section 4 for the definition of the Zipf coefficient). The more concentrated the population distribution becomes, the smaller the Zipf coefficient becomes. Since the agglomeration economies amplify city size differences by making big cities even bigger, greater agglomeration economies lead to a lower Zipf coefficient. Dobkins and Ioannides (2001) find that the Zipf coefficient for the U.S. declines from 1.044 in 1900 to .949 in 1990. One interpretation based on our model is that agglomeration economies became more important over this period.

Our model also generates implications that one can test against other types of models. First, against the random growth models, our model predicts that each city has a stable unique equilibrium city size. Thus, in response to a temporary shock, city size tends to go back to the original equilibrium size. On the other hand, the random growth models predict that each shock has permanent effect on city size and thus city size does not have a tendency to go back to the original size. Second, against the models relying only on agglomeration economies assuming ex-ante identical features across cities, our model predicts substantial city size variation even in a period when agglomeration economies play little role. Third, against the models relying only on exogenous heterogeneous features, our model predicts that population distribution will become more concentrated into larger cities in the period when agglomeration economies become more important.

Davis and Weinstein (2002) test these implications. They study the distribution of regional population in Japan from Stone Age to the modern era and obtain the following results. First, after the extensive bombing over Japanese cities during World War II, including two nuclear attacks, most cities returned to their relative position in the distribution of city sizes within about 15 years. This confirms our first implication. Second, throughout history there has always been a great deal of variation in population density across regions, even in the Stone Age when agglomeration economies would not seem to have played an important role. This confirms our second implication. Third, the population distribution became

more concentrated into larger cities in the last century when agglomeration economies became more important as Japan became industrialized and more integrated into the world economy. This confirms our third implication. Based on these findings, Davis and Weinstein (2002) advocate for a hybrid model of agglomeration economies and heterogeneous natural features. Our model is the first one in this direction.

The rest of the paper is structured as follows. Section 2 presents the base model. Section 3 shows that our model generates the log-normal distribution. Section 4 shows that our model is consistent with Zipf’s law in the top tail. Section 5 concludes.

## 2 Base Model

The key idea of this paper is that the city size distribution can generate Zipf’s law in the top tail if city size can be expressed as a product of numerous random factors. This section provides an underlying model to support this idea. The model starts from Roback (1982). The Roback model predicts the wage and rent of a city as a function of its local production amenities and consumption amenities. In order to transform the Roback model into a model of city size distribution, we make two changes. First, we add a housing market and this works as the main congestion force pinning down the population size of a city. Second, we allow local production and consumption amenities to depend on population size to capture the agglomeration economies. As the result, the model predicts city size as an increasing function of production amenities, consumption amenities and land supply.

### 2.1 Description

A continuum of potential city sites are indexed by  $s \in [0, 1]$ . The locations differ *exogenously* in three groups of characteristics: natural consumption amenities  $\mathbf{a} \in \mathbb{R}^J$ , natural production amenities  $\mathbf{o} \in \mathbb{R}^K$ , and physical land supply factors  $\mathbf{l} \in \mathbb{R}^M$ . These natural features capture rivers, mountains, climate, coastal locations, etc. The locations differ *endogenously* in ag-



aggregate consumption amenities  $A \in \mathbb{R}$ , aggregate production amenities  $O \in \mathbb{R}$ , aggregate land supply  $L \in \mathbb{R}$  as well as population size  $N$ , wage  $w$ , and rent  $r$ . These endogenous features encompass the effects of man-made facilities such as restaurants, high ways, zoning restrictions as well as the natural amenities. Note that  $\mathbf{a}, \mathbf{o}, \mathbf{l}$  are vectors but  $A, O, L$  are scalars capturing the aggregate effects. The local consumption amenities  $A$ , production amenities  $O$ , and land supply  $L$  all depend on the natural features  $\mathbf{a}, \mathbf{o}$  and  $\mathbf{l}$  as well as population size  $N$ :

$$A = A(N, \mathbf{a}),$$

$$O = O(N, \mathbf{o}),$$

$$L = L(N, \mathbf{l}).$$

There are two commodities: a composite good and housing. The composite good is freely tradable with zero transportation cost while housing is locally provided. The markets for both goods are perfectly competitive.

$\bar{N}$  workers live in the economy. All workers are homogeneous and freely mobile with zero moving cost. A worker first chooses a city to live in and then chooses her consumption bundle consisting of the composite good  $q$  and housing  $h$ . The utility function  $U(q, h; A)$  is increasing in the consumption amenities  $A$ . Each worker supplies one unit of labor. The decision of a worker can be summarized by the following maximization problem:

$$\max_s V(r_s, w_s; A_s)$$

where

$$V(r_s, w_s; A_s) \equiv \max_{q, h} U(q, h; A_s) \text{ subject to } q + r_s h = w_s.$$

where  $r_s$ ,  $w_s$ , and  $A_s$  are housing rent, wage, and consumption amenities in city  $s$ . We use the composite good as the numeraire.

Each city has numerous firms producing the composite good. All firms use the same constant-returns-to-scale technology and thus we can consider one aggregate firm for each

city, which behaves like a perfectly competitive firm. The aggregate firm uses labor  $n$  and buildings which we assume come from the same stock of housing as workers' housing  $h$ . The production function  $F$  is increasing in the production amenities  $O$ . The decision of a firm in city  $s$  can be summarized by the following maximization problem:

$$\max_{n,h} F(n, h; O_s) - w_s n - r_s h$$

where  $n$  and  $h$  are labor and housing input.

All housing is owned by absentee landlords. Instead of explicitly modeling housing developers, we assume for simplicity that housing supply is a function of rent  $r$  and land supply  $L$ :

$$H^S(r; L).$$

## 2.2 Equilibrium

An equilibrium of the model  $\{S, \bar{u}, w_s, r_s, N_s | s \in S\}$  consists of the set of populated sites  $S$ , equilibrium utility level  $\bar{u}$ , and wage  $w_s$ , rent  $r_s$ , population size  $N_s$  for each city  $s \in S$ , satisfying the following five conditions. First, workers get the same utility across all populated locations:

$$(1) \quad V(r_s, w_s; A(N_s; \mathbf{a}_s)) = \bar{u} \text{ for } s \in S$$

where  $\bar{u}$  is the common utility level and  $S$  is the set of locations with positive population size ( $S \equiv \{s | N_s > 0\}$ ).

Second, firms that produce the composite good earn zero profits. Since the firms use constant returns to scale technology, the zero profit condition is equivalent to the unit cost being equal to the unit output price:

$$(2) \quad C(r_s, w_s; O(N_s; \mathbf{o}_s)) = 1 \text{ for } s \in S$$

where  $C$  is the unit cost function.

Third, the housing market in each city clears:

$$(3) \quad H^D(N_s, r_s, w_s; A(N_s; \mathbf{a}_s), O(N_s; \mathbf{o}_s)) = H^S(r_s; L(N_s; \mathbf{l}_s)) \text{ for } s \in S$$

where  $H^D$  is the aggregate housing demand function of workers and firms.

Fourth, economy-wide labor market clears:

$$(4) \quad \int_S N_s ds = \bar{N}.$$

Fifth, unpopulated sites offer lower utility than the common utility level  $\bar{u}$ :

$$(5) \quad V(r_s, w_s; A(0, \mathbf{a}_s)) \leq \bar{u} \text{ for } s \notin S.$$

Equations (1) to (3) determine wage  $w_s$ , housing rent  $r_s$  and population size  $N_s$  for each city  $s \in S$ , when the common utility level  $\bar{u}$  and the set of populated sites  $S$  are given. Equation (4) determines the common utility level  $\bar{u}$  given the set of populated sites  $S$ , since  $N_s$  is a function of  $\bar{u}$ . Equation (5) characterizes the set of populated sites  $S$ .

### 3 Log-Normal Distribution

We obtain the key result by imposing specific functional forms. First, we use the following functional forms for workers' preference, firms' production technology, and housing supply.

$$(6) \quad \begin{array}{ll} \text{Preference:} & U(q, h; A) = A \cdot q^\alpha h^{1-\alpha} \\ \text{Production technology:} & F(n, h; O) = O \cdot n^\beta h^{1-\beta} \\ \text{Housing supply:} & H^s(r; L) = L \cdot r^\gamma \end{array}$$

where  $\alpha, \beta \in (0, 1)$  and  $\gamma > 0$ . These functional forms are quite standard and also have some empirical support. For example, Davis and Ortalo-Magne (forthcoming) show that the income elasticity of housing consumption is 1. The Cobb-Douglas function is arguably the most common production function in economics. The housing supply function can be motivated by the following story. Imagine a city located on a coast so can expand only in

180 degrees. Its land supply  $L$  would be a half as large as that of a city that can expand in 360 degrees.

Second, we assume that consumption amenities  $A_s$ , production amenities  $O_s$  and land supply  $L_s$  can be expressed as the product of multiple underlying factors:

$$(7) \quad A_s = \prod_{j=1}^J A_{j,s}, \quad O_s = \prod_{k=1}^K O_{k,s}, \quad L_s = \prod_{m=1}^M L_{m,s},$$

where

$$(8) \quad \begin{aligned} A_{j,s} &= a_{j,s} N_s^{\lambda_j}, \\ O_{k,s} &= o_{k,s} N_s^{\mu_k}, \\ L_{m,s} &= l_{m,s} N_s^{\nu_m}. \end{aligned}$$

Each factor  $A_{j,s}$ ,  $O_{k,s}$ ,  $L_{m,s}$  consists of exogenous features  $a_{j,s}$ ,  $o_{k,s}$ ,  $l_{m,s}$  and agglomeration economy terms  $N_s^{\lambda_j}$ ,  $N_s^{\mu_k}$ ,  $N_s^{\nu_m}$ . The agglomeration economy parameters  $\lambda_j$ ,  $\mu_k$ ,  $\nu_m$  can differ across different factors.

With these functional forms, we derive our theoretical results. We first show that the equilibrium population size of each city is unique and stable. This result relates to the natural experiment Davis and Weinstein (2002) studied. Suppose that  $S$  and  $\bar{u}$  are fixed. Using equations (2) and (3), we can solve for wage  $w_s$  and rent  $r_s$  as the functions of  $N_s$ . By substituting these into  $w_s$  and  $r_s$  in the indirect utility function  $V$  we can express the indirect utility function in terms of only population size  $N$ . (See Appendix A for details.)

$$(9) \quad \tilde{V}_s(N) = \Phi \left( \frac{1}{N} \right)^{\frac{1-\Omega}{\Phi_A}} \left\{ \prod_{j=1}^J (a_{j,s})^{\Phi_A} \prod_{k=1}^K (o_{k,s})^{\Phi_O} \prod_{m=1}^M l_{m,s} \right\}^{\frac{1}{\Phi_A}}$$

where  $\Phi_A$ ,  $\Phi_O$ , and  $\Phi$  are *positive* constants and  $\Omega \equiv \Phi_A \sum_{j=1}^J \lambda_j + \Phi_O \sum_{k=1}^K \mu_k + \sum_{m=1}^M \nu_m$  is the aggregate agglomeration economy parameter.<sup>3</sup> This indirect utility function  $\tilde{V}_s(N)$  is the utility city  $s$  offers when its population size is  $N$ . Equilibrium population size  $N_s$  is determined by the remaining equation (1):

$$(10) \quad \tilde{V}_s(N_s) = \bar{u}.$$

---

<sup>3</sup>  $\Phi_A = \frac{1+\beta\gamma}{1-\alpha\beta}$ ,  $\Phi_O = \frac{\alpha+\gamma}{1-\alpha\beta}$ ,  $\Phi = (1-\alpha)^{1-\alpha} \alpha^\alpha (1-\beta)^{-\frac{(-1+\beta)(\alpha+\gamma)}{1+\beta\gamma}} \beta (1-\alpha\beta)^{\frac{-1+\alpha\beta}{1+\beta\gamma}}$

Suppose that  $\Omega < 1$ . It is clear from equation (9) that  $\tilde{V}_s(N)$  is continuous and strictly decreasing in  $N$  with  $\lim_{N \rightarrow 0} \tilde{V}_s(N) = \infty$  and  $\lim_{N \rightarrow \infty} \tilde{V}_s(N) = 0$ . Thus, there exists a unique  $N_s$  satisfying equation (10) for any positive  $\bar{u}$  by the intermediate value theorem. In addition, this population size  $N_s$  is stable in Krugman (1991) sense. For example, suppose that a negative temporary shock hits a city  $s$  and its population size decreases below equilibrium city size  $N_s$ . With a smaller population size, the utility the city offers is greater than the common utility level  $\bar{u}$  and thus its population size increases back to equilibrium size  $N_s$ . Note that if  $\Omega > 1$ , equilibrium city size becomes unstable (i.e.,  $\tilde{V}_s(N)$  is increasing in  $N$ ) and this can lead to a black-hole equilibrium where all people go to only one city.

The intuition behind this stability result is the following. As city size decreases, housing price decreases allowing the city to offer higher utility. However, the downside of losing population is the loss in the agglomeration economies in consumption amenities, production amenities, and land supply. If the agglomeration economy parameter  $\Omega$  is less than 1, the housing effect dominates so the city offers better utility with a smaller population size and this makes equilibrium city size stable. On the other hand, if  $\Omega$  is greater than 1, the agglomeration effect dominates so a city offers better utility with a larger population size. This makes the equilibrium city size unstable and can lead to the black-hole equilibrium.

By solving equation (10), we obtain equilibrium population size  $N_s$ :

$$(11) \quad N_s = \left\{ \left( \frac{\Phi}{\bar{u}} \right)^{\Phi_A} \prod_{j=1}^J (a_{j,s})^{\Phi_A} \prod_{k=1}^K (o_{k,s})^{\Phi_O} \prod_{m=1}^M l_{m,s} \right\}^{\frac{1}{1-\Omega}}.$$

Equilibrium population size  $N_s$  is strictly increasing in production amenities  $a_{j,s}$ , consumption amenities  $o_{k,s}$ , and supply factors  $l_{m,s}$  and strictly decreasing in the common utility level  $\bar{u}$ .

So far we have taken equilibrium utility level  $\bar{u}$  and the set of populated sites  $S$  as given. Equilibrium utility level  $\bar{u}$  is unique given the set of populated sites  $S$ . This follows from equation (4) because the population size of each city is continuous and strictly decreasing in  $\bar{u}$  with  $\lim_{\bar{u} \rightarrow 0} N_s(\bar{u}) = \infty$  and  $\lim_{\bar{u} \rightarrow \infty} N_s(\bar{u}) = 0$  as can be seen in equation (11). The

intermediate value theorem implies that equation (4) is satisfied for only one value of  $\bar{u}$ . The set of populated sites  $S$  is equal to the set of all locations  $[0, 1]$  because the indirect utility  $\lim_{N \rightarrow 0} \tilde{V}_s(N) = \infty$  and zero population size is not stable.<sup>4</sup>

**Proposition 1** *Suppose that  $\Omega < 1$ .*

- (a) *The population size of each city is unique and stable.*
- (b) *Population size  $N_s$  of city  $s$  is increasing in consumption amenity factor  $a_{j,s}$ , production amenity factor  $o_{k,s}$ , and land supply factor  $l_{m,s}$  ( $j \in J, k \in K, m \in M$ ).*

Now we derive our key result that if the exogenous factors  $a_j$ ,  $o_k$ , and  $l_m$  are randomly distributed, population size  $N$  converges in distribution to the log-normal distribution as the number of these factors increases. We interpret  $a_{j,s}, o_{k,s}$ , and  $l_{m,s}$  as the realizations of random variables  $a_j$ ,  $o_k$ , and  $l_m$  so we do not show the city index  $s$  any more. Taking log transformation of equation (11) we obtain

$$(12) \quad \log N = \frac{1}{1 - \Omega} \left\{ \sum_{j=1}^J \Phi_A \log a_j + \sum_{k=1}^K \Phi_O \log o_k + \sum_{m=1}^M \log l_m + \Phi_A \log \left( \frac{\Phi}{\bar{u}} \right) \right\}$$

Mathematically  $\Phi_A \log a_j$ ,  $\Phi_O \log o_k$ , and  $\log l_m$  play the same roles. In order to simplify notations we introduce new symbol  $X_i$  which we use for all the three types. We can reorder the attribute terms

$$(\Phi_A \log a_1, \dots, \Phi_A \log a_J, \Phi_O \log o_1, \dots, \Phi_O \log o_K, \log l_1, \dots, \log l_M)$$

as we like and assign  $X_1, \dots, X_I$  where  $I \equiv J + K + M$ . We rewrite equation (12) using the new notations as

$$(13) \quad \log N^I = \frac{1}{1 - \Omega} \left\{ \sum_{i=1}^I X_i + \Phi_A \log \left( \frac{\Phi}{\bar{u}} \right) \right\}$$

$$(14) \quad = \frac{1}{1 - \Omega} \left\{ \sum_{i=1}^I \hat{X}_i + \sum_{i=1}^I \bar{X}_i + \Phi_A \log \left( \frac{\Phi}{\bar{u}} \right) \right\}$$

---

<sup>4</sup>This is a undesirable feature coming from the functional forms we use. We can fix this by tweaking the agglomeration economies for small cities or by introducing fixed cost to develop a city that has to be paid by the absentee landlords. However, this fix would come at the cost of making the model more complicated and less focused.

where  $\bar{X}_i \equiv E(X_i)$ , and  $\hat{X}_i \equiv X_i - \bar{X}_i$ . Note that  $E(\hat{X}_i) = 0$ . In order to show that city size distribution converges to log-normal distribution, it suffices to show that  $\sum_{i=1}^I \hat{X}_i$  in equation (14) converges to normal distribution as  $I$  increases. We obtain this result by applying the central limit theorem to  $\sum_{i=1}^I \hat{X}_i$ .

The classical central limit theorem states that  $\sum_{i=1}^I \hat{X}_i$  converges in distribution to normal distribution if  $\hat{X}_i$ s are independent and identically distributed. Since this requirement is too restrictive for our purpose, we use a version of modern central limit theorem which relaxes these requirements. This version allows different random variables  $\hat{X}_i$  to come from different distributions and also allows correlation among the variables to some degree. Kourogenis and Pittis (2008) provide an excellent survey of modern central limit theorems. The version we use corresponds to Theorem 4 in Kourogenis and Pittis (2008) which in turn is based on Corollary 1 in Herrndorf (1984). We begin by describing the allowable correlation structure among the random variables using  $\alpha$ -mixing.

**Definition 2** *For a sequence  $\hat{X}_1, \hat{X}_2, \dots$  of random variables, let  $\alpha_i$  be a number such that*

$$|P(A \cap B) - P(A)P(B)| \leq \alpha_i \text{ for } A \in \sigma(\hat{X}_1, \dots, \hat{X}_n) \text{ and } B \in \sigma(\hat{X}_{n+i}, \hat{X}_{n+i+1}, \dots)$$

*where  $\sigma(\mathbf{X})$  is defined as the  $\sigma$ -field generated by  $\mathbf{X}$ . If  $\alpha_i \rightarrow 0$  as  $i \rightarrow \infty$ , the sequence  $\hat{X}_1, \hat{X}_2, \dots$  is said to be  $\alpha$ -mixing.*

If a sequence  $\{\hat{X}_i, i \in \mathbb{N}\}$  is  $\alpha$ -mixing,  $\hat{X}_n$  and  $\hat{X}_{n+i}$  becomes approximately independent as  $i$  increases to infinity. How much they are correlated with each other depends on how fast  $\alpha_i$  converges to 0 as  $i$  increases. Now we state the central limit theorem.

**Theorem 3** *(Herrndorf (1984)) Let  $\{\hat{X}_i, i \in \mathbb{N}\}$  be an  $\alpha$ -mixing sequence of random vari-*

ables satisfying the following conditions.

- 1)  $E(\hat{X}_i) = 0$
- 2)  $\lim_{i \rightarrow \infty} E(\frac{S_i^2}{i}) = \sigma^2, 0 < \sigma^2 < \infty$
- 3)  $\sup_{i \in \mathbf{N}} E|\hat{X}_i|^b < \infty$  for some  $b > 2$
- 4)  $\sum_{i=1}^{\infty} (\alpha_i)^{1-\frac{2}{b}} < \infty$

where  $S_i \equiv \sum_{j=1}^i \hat{X}_j$ . Then  $\frac{1}{\sigma\sqrt{i}}S_i$  converges in distribution to the standard normal distribution  $N(0, 1)$ .

Since we construct  $\hat{X}_i$  so that  $E(\hat{X}_i) = 0$ , the first condition is satisfied. The second condition means that the variances of the partial sum behaves nicely. The third condition requires that the moments of order  $b > 2$  to be uniformly bounded. The fourth condition puts restriction on the  $\alpha$ -mixing rate. The third and the fourth conditions are linked by  $b$ . As  $b$  increases, the third condition becomes harder to satisfy and the fourth condition becomes easier to satisfy. By applying Theorem 3 we obtain our main result.

**Proposition 4** *If the sequence  $\hat{X}_1, \hat{X}_2, \dots$  satisfies the conditions listed in Theorem 3, the city size distribution converges in distribution to log normal distribution. Asymptotically, city size  $N^I$  with  $I$  random factors follows  $\log N \left( \frac{1}{1-\Omega} \left\{ \sum_{i=1}^I \bar{X}_i + \Phi_A \log \left( \frac{\Phi}{u} \right) \right\}, \frac{\sigma^2 I}{(1-\Omega)^2} \right)$ .*

## 4 Zipf's Law

This section shows the relationship between our model and Zipf's law. Zipf's law emerges when city size follows Pareto distribution with the shape parameter 1. Suppose that city size  $N$  follows Pareto distribution with scale parameter  $\tilde{N}$  and shape parameter  $\alpha$ :

$$CDF(N) = 1 - \left( \frac{\tilde{N}}{N} \right)^\alpha.$$



When there are  $M$  cities in total, the rank  $R_s$  of city  $s$  with population size  $N_s$  can be approximated as

$$R_s \approx M (1 - CDF(N_s)) = M \left( \frac{\tilde{N}}{N_s} \right)^\alpha.$$

When  $\alpha = 1$ , we obtain Zipf's law:

$$N_s \approx \frac{1}{R_s} \cdot M \tilde{N}.$$

In other words, the population size of a city is inversely proportional to its rank.

Typically, the parameters are estimated by the Zipf regression:

$$(15) \quad \log R_s = C - \alpha \log N_s$$

where  $C$  is a constant term. Due to data availability most empirical studies use only the largest cities in a country to estimate the Zipf coefficient  $\alpha$  and find that Zipf's law holds well in the top tail (e.g., Rosen and Resnick, 1980; Dobkins and Ioannides, 2001; Soo, 2005). Gabaix and Ioannides (2004) survey the literature and report that most estimates of the Zipf coefficient  $\alpha$  fall into  $[0.8, 1.2]$ .

For the whole distribution including bottom tail there is ongoing debate over what distribution best describes the city size distribution. The issue is that population size for cities in the bottom tail are usually not available. Rozenfeld et al. (forthcoming) use their own algorithm (City Clustering Algorithm) to construct cities from US Census tracts population distribution and find that Zipf's law holds for cities larger than 12,000 inhabitants in the US. Eeckhout (2004) uses US Census places to look at small cities and finds that the size distribution for all cities follows the log-normal distribution. He also argues that the log-normal distribution is hardly distinguishable from Pareto distribution in top tail and thus approximately consistent with Zipf's law in top tail. Ioannides and Skouras (2009) uses data in Eeckhout (2004) and argue that there is a switch from log-normal distribution to Pareto distribution around population size of 100,000.

Our model generates log-normal distribution asymptotically and thus, as in Eeckhout (2004), consistent with Zipf's law in top tail. However, this argument may not work if city

size is determined by a small number of dominant factors or the random factors are too much correlated with each other. We examine this issue as follows. First, we prove analytically that our model can generate the Zipf coefficient equal to 1 by adjusting model parameters, regardless of the number of factors or the degree of correlation. Second, we calculate  $R^2$  in the Zipf regression for simulated samples with different number of factors and different degree of correlation. The  $R^2$  shows how well our model can generate the linear relationship implied by Pareto distribution between log rank and log city size.

Using equations (13) and (15) we can express the Zipf coefficient as

$$\alpha = -\frac{Cov(\log N, \log R)}{Var(\log N)} = (1 - \Omega) \frac{-Cov\left(\sum_{i=1}^I X_i, \log R\right)}{Var\left(\sum_{i=1}^I X_i\right)}.$$

Thus, the Zipf coefficient  $\alpha$  depends on agglomeration economy parameter  $\Omega$ , distribution of random shocks  $\{X_i\}$ , and the number of random shocks  $I$ . Since  $Cov\left(\sum_{i=1}^I X_i, \log R\right)$  is negative, we obtain the following results immediately.

**Proposition 5** 1. *Zipf coefficient  $\alpha$  is proportional to  $1 - \Omega$ , thus decreasing in  $\Omega$ .*

2. *Suppose that we rescale the random shock distribution so that  $X' = \varphi X$  and use  $X'$  to calculate the Zipf coefficient. The Zipf coefficient  $\alpha$  is inversely proportional to  $\varphi$ , thus decreasing in  $\varphi$ .*

Proposition 5.1 links agglomeration economies to the Zipf coefficient. This result is intuitive. A smaller Zipf coefficient means that city size distribution is more uneven; larger agglomeration economies make big cities even bigger creating more uneven city size distribution. Dobkins and Ioannides (2001) show that the Zipf coefficient for the U.S. has declined in the twentieth century. One explanation based on our model is that the agglomeration economies became more important over this period.

Proposition 5.1 suggests that we can always match the Zipf coefficient equal to 1 by adjusting  $\Omega$  or by rescaling the random shock distribution. Note that this result does not depend on the number of factors or the degree of correlation. Gabaix and Ioannides (2004)

show that the OLS estimate of  $\alpha$  is downward biased in a small sample and provide the magnitude of the bias for various sample sizes. This is not a problem for us because we can match the Zipf coefficient with the bias taken into account.

Now we examine the goodness of fit  $R^2$  by running a simulation. We proceed as follows. We fix the number of factors and the degree of correlation among the factors, generate 25,000 city sizes using our model, truncate the distribution to include only top 135 cities and run the Zipf regression.<sup>5</sup> We repeat this 2,000 times and report average  $R^2$  and its standard deviation. We vary the number of factors and the degree of correlation and repeat the whole process.

We use the following process to generate random factors  $X_i$ s.

$$\begin{aligned} X_1 &= \varepsilon_1. \\ X_{i+1} &= \rho X_i + (1 - \rho) \varepsilon_{i+1} \end{aligned}$$

where  $\varepsilon_i$  follows iid Uniform distribution  $[0, 1]$ .  $\rho$  captures the degree of correlation: all factors are perfectly correlated if  $\rho = 1$  and are independent if  $\rho = 0$ . We do not report  $\Omega$ ,  $\Phi_A$ ,  $\Phi$  and  $\bar{u}$  because these parameters do not affect  $R^2$  once the random factors  $X_i$ s are determined. In order to see this, we obtain the following equation by inserting  $\log N^I$  in equation (13) into equation (15):

$$\log R^j = C - \alpha \cdot \frac{1}{1 - \Omega} \left\{ \sum_{i=1}^I X_i^j + \Phi_A \log \left( \frac{\Phi}{\bar{u}} \right) \right\}$$

It is clear from this equation that changes in these parameters are fully absorbed by changes in  $\hat{\alpha}$  and  $\hat{C}$  and thus do not affect predicted log rank and thus  $R^2$ . By the same reason rescaling the random shock distribution does not affect  $R^2$  either.

The simulation result in Table 1 shows that our simulated distribution quickly converges close to Pareto distribution in the top tail. When the factors are independent ( $\rho = 0$ ),

---

<sup>5</sup>We use 25,000 clusters because this is close to the numbers of places (25,359) in Eeckhout (2004) and clusters (23,499) in Rozenfeld et al. (forthcoming). We truncate them at the top 135 cities because this is the threshold repeatedly used in the literature.

Table 1: Average  $R^2$ 's from the Zipf regressions on simulated samples

# of Factors	$\rho=0$	$\rho=0.5$	$\rho=0.9$	$\rho=1$
1	0.790 (0.030)	0.791 (0.029)	0.789 (0.030)	0.790 (0.029)
2	0.917 (0.022)	0.917 (0.022)	0.916 (0.022)	0.789 (0.030)
3	0.955 (0.017)	0.955 (0.018)	0.955 (0.018)	0.791 (0.030)
4	0.970 (0.016)	0.970 (0.015)	0.965 (0.016)	0.790 (0.029)
5	0.977 (0.013)	0.976 (0.014)	0.971 (0.015)	0.791 (0.029)
6	0.980 (0.012)	0.980 (0.012)	0.974 (0.014)	0.789 (0.029)
7	0.982 (0.011)	0.982 (0.011)	0.976 (0.014)	0.791 (0.029)
8	0.983 (0.011)	0.982 (0.011)	0.978 (0.013)	0.791 (0.028)
9	0.983 (0.011)	0.983 (0.011)	0.978 (0.013)	0.790 (0.029)
10	0.983 (0.010)	0.983 (0.010)	0.980 (0.013)	0.791 (0.029)

average  $R^2$  becomes greater than .98 around 6 factors. When the factors are correlated (but not perfectly correlated), our simulated distribution still quickly converges to Pareto distribution. For example, when  $\rho = 0.9$ , average  $R^2$  becomes greater than .98 around 10 factors. The .98 is the benchmark  $R^2$  we obtain by running the same simulation but using the Pareto distribution for city size.

## 5 Conclusion

This paper proposes a new explanation for the robust empirical pattern in city size distribution: if city size is determined as a product of numerous factors, this may generate Zipf's law in the top tail. The key implication of our theory is that we cannot reject a model simply because the model does not generate Zipf's law. A typical economic model focuses on one economic force it aims to deliver, abstracting away the other forces existing in reality. Our theory demonstrates that one single model may not generate Zipf's law but Zipf's law may

emerge if we have many such models together as in reality.

## A How to Obtain $\tilde{V}_s(N)$

Using the functional specifications in equations (6) to (8) we can rewrite the equilibrium conditions (1) to (3) as

$$(16) \quad V(r, w; A_s) = A_s \alpha^\alpha (1 - \alpha)^{1-\alpha} w r^{-(1-\alpha)} = \bar{u},$$

$$(17) \quad C(r, w; O_s) = (O_s)^{-1} \beta^{-\beta} (1 - \beta)^{-(1-\beta)} w^\beta r^{1-\beta} = 1,$$

$$(18) \quad H^D(N_s, r, w; A_s, O_s) \equiv N_s \frac{1 - \alpha\beta}{\beta} \frac{w}{r} = Lr^\gamma \equiv H^S(r; L_s).$$

Note that housing is demanded by both workers and firms.

Define  $a_s = \prod_{j=1}^J a_{j,s}$ ,  $o_s = \prod_{k=1}^K o_{k,s}$ ,  $l_s = \prod_{m=1}^M l_{m,s}$ ,  $\lambda = \sum_{j=1}^J \lambda_j$ ,  $\mu = \sum_{k=1}^K \mu_k$ , and  $\nu = \sum_{m=1}^M \nu_m$ . From Equation (17) and Equation (18), we can solve  $r$  and  $w$  in terms of population  $N$  as follows

$$\begin{aligned} r_s &= \left[ (1 - \beta)^{1-\beta} (1 - \alpha\beta)^\beta o_s (l_s)^{-\beta} (N_s)^{\mu + (1-\nu)\beta} \right]^{\frac{1}{1+\beta\gamma}}, \\ w_s &= (1 - \beta)^{\frac{(1+\gamma)(1-\beta)}{1+\beta\gamma}} \beta (1 - \alpha\beta)^{-\frac{1-\beta}{1+\beta\gamma}} (o_s)^{\frac{1+\gamma}{1+\beta\gamma}} (l_s)^{\frac{1-\beta}{1+\beta\gamma}} (N_s)^{\frac{\mu(1+\gamma) - (1-\nu)(1-\beta)}{1+\beta\gamma}}. \end{aligned}$$

Substitute the above into the expression for the indirect utility in Equation (16), we get

$$\tilde{V}_s(N) = \Phi \left( \frac{1}{N} \right)^{\frac{1-\Omega}{\Phi_A}} \left[ (a_s)^{\Phi_A} (o_s)^{\Phi_O} l_s \right]^{\frac{1}{\Phi_A}},$$

where  $\Phi = \alpha^\alpha (1 - \alpha)^{1-\alpha} \beta (1 - \alpha\beta)^{-\frac{1}{\Phi_A}} (1 - \beta)^{(1-\beta)\frac{\Phi_O}{\Phi_A}}$ ,  $\Omega = \nu + \Phi_A \lambda + \Phi_O \mu$ ,  $\Phi_A = \frac{1+\beta\gamma}{1-\alpha\beta}$ , and  $\Phi_O = \frac{\alpha+\gamma}{1-\alpha\beta}$ .

## References

BEHRENS, K., G. DURANTON AND F. ROBERT-NICOUD, “Productivity Cities: Sorting, Selection and Agglomeration,” Working paper, Department of Economics, University du Quebec a Montreal, 2010.

- CÓRDOBA, J.-C., “On the Distribution of City Sizes,” *Journal of Urban Economics* 63 (January 2008), 177–197.
- DAVIS, D. R. AND D. E. WEINSTEIN, “Bones, Bombs, and Break Points: The Geography of Economic Activity,” *American Economic Review* 92 (December 2002), 1269–1289.
- DAVIS, M. A. AND F. ORTALO-MAGNE, “Household Expenditures, Wages, Rents,” *Review of Economic Dynamics* (forthcoming).
- DOBKINS, L. H. AND Y. M. IOANNIDES, “Spatial Interactions among U.S. Cities: 1900–1990,” *Regional Science and Urban Economics* 31 (November 2001), 701 – 731.
- DURANTON, G., “Some Foundations for Zipf’s Law: Product Proliferation and Local Spillovers,” *Regional Science and Urban Economics* 36 (July 2006), 542–563.
- , “Urban Evolutions: The Fast, the Slow, and the Still,” *American Economic Review* 97 (March 2007), 197 – 221.
- ECKHOUT, J., “Gibrat’s Law for (All) Cities,” *American Economic Review* 94 (December 2004), 1429–1451.
- , “Gibrat’s Law for (All) Cities: Reply,” *American Economic Review* 99 (September 2009), 1676–83.
- GABAIX, X., “Zipf’s Law for Cities: An Explanation,” *Quarterly Journal of Economics* 114 (August 1999), 739–767.
- GABAIX, X. AND Y. M. IOANNIDES, “The Evolution of City Size Distributions,” in J. V. Henderson and J. F. Thisse, eds., *Handbook of Regional and Urban Economics* volume 4 of *Handbook of Regional and Urban Economics*, chapter 53 (Elsevier, 2004), 2341–2378.
- HENDERSON, J. V., “The Sizes and Types of Cities,” *American Economic Review* 64 (September 1974), 640–56.

- HERRNDORF, N., “A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables,” *Annals of Probability* 12 (February 1984), 141–153.
- HSU, W., “Central Place Theory and City Size Distribution,” Working paper, Department of Economics, Chinese University of Hong Kong, August 2009.
- IOANNIDES, Y. M. AND H. G. OVERMAN, “Zipf’s Law for Cities: An Empirical Examination,” *Regional Science and Urban Economics* 33 (March 2003), 127–137.
- IOANNIDES, Y. M. AND S. SKOURAS, “Gibrat’s Law for (All) Cities: A Rejoinder,” Discussion Papers Series 0740, Department of Economics, Tufts University, 2009.
- KOUROGENIS, N. AND N. PITTIS, “Mixing Conditions, Central Limit Theorems and Invariance Principles: A Survey of the Literature with Some New Results on Heteroscedastic Sequences,” *SSRN eLibrary* (2008).
- KRUGMAN, P., “Increasing Returns and Economic Geography,” *Journal of Political Economy* 99 (June 1991), 483–99.
- , “Confronting the Mystery of Urban Hierarchy,” *Journal of the Japanese and International Economies* 10 (December 1996), 399 – 418.
- ROBACK, J., “Wages, Rents, and the Quality of Life,” *Journal of Political Economy* 90 (December 1982), 1257–1278.
- ROMER, P. M., “Endogenous Technological Change,” *Journal of Political Economy* 98 (October 1990), S71–102.
- ROSEN, K. T. AND M. RESNICK, “The Size Distribution of Cities: An Examination of the Pareto Law and Primacy,” *Journal of Urban Economics* 8 (September 1980), 165 – 186.
- ROSSI-HANSBERG, E. AND M. L. J. WRIGHT, “Urban Structure and Growth,” *Review of Economic Studies* 74 (April 2007), 597 – 624.

- ROZENFELD, H. D., D. RYBSKI, X. GABAIX AND H. A. MAKSE, “The Area and Population of Cities: New Insights from a Different Perspective on Cities,” *American Economic Review* (forthcoming).
- SIMON, H. A., “On a Class of Skew Distribution Functions,” *Biometrika* 42 (December 1955), 425–440.
- SOO, K. T., “Zipf’s Law for Cities: A Cross-country Investigation,” *Regional Science and Urban Economics* 35 (May 2005), 239–263.



2009

- 2009/1. Rork, J.C.; Wagner, G.A.: "Reciprocity and competition: is there a connection?"
- 2009/2. Mork, E.; Sjögren, A.; Svaleryd, H.: "Cheaper child care, more children"
- 2009/3. Rodden, J.: "Federalism and inter-regional redistribution"
- 2009/4. Ruggeri, G.C.: "Regional fiscal flows: measurement tools"
- 2009/5. Wrede, M.: "Agglomeration, tax competition, and fiscal equalization"
- 2009/6. Jametti, M.; von Ungern-Sternberg, T.: "Risk selection in natural disaster insurance"
- 2009/7. Solé-Ollé, A.; Sorribas-Navarro, P.: "The dynamic adjustment of local government budgets: does Spain behave differently?"
- 2009/8. Sanromá, E.; Ramos, R.; Simón, H.: "Immigration wages in the Spanish Labour Market: Does the origin of human capital matter?"
- 2009/9. Mohnen, P.; Lokshin, B.: "What does it take for and R&D incentive policy to be effective?"
- 2009/10. Solé-Ollé, A.; Salinas, P.: "Evaluating the effects of decentralization on educational outcomes in Spain?"
- 2009/11. Libman, A.; Feld, L.P.: "Strategic Tax Collection and Fiscal Decentralization: The case of Russia"
- 2009/12. Falck, O.; Fritsch, M.; Heblich, S.: "Bohemians, human capital, and regional economic growth"
- 2009/13. Barrio-Castro, T.; García-Quevedo, J.: "The determinants of university patenting: do incentives matter?"
- 2009/14. Schmidheiny, K.; Brülhart, M.: "On the equivalence of location choice models: conditional logit, nested logit and poisson"
- 2009/15. Itaya, J.; Okamura, M.; Yamaguchi, C.: "Partial tax coordination in a repeated game setting"
- 2009/16. Ens, P.: "Tax competition and equalization: the impact of voluntary cooperation on the efficiency goal"
- 2009/17. Geys, B.; Revelli, F.: "Decentralization, competition and the local tax mix: evidence from Flanders"
- 2009/18. Konrad, K.; Kovenock, D.: "Competition for fdi with vintage investment and agglomeration advantages"
- 2009/19. Loretz, S.; Moorey, P.: "Corporate tax competition between firms"
- 2009/20. Akai, N.; Sato, M.: "Soft budgets and local borrowing regulation in a dynamic decentralized leadership model with saving and free mobility"
- 2009/21. Buzzacchi, L.; Turati, G.: "Collective risks in local administrations: can a private insurer be better than a public mutual fund?"
- 2009/22. Jarkko, H.: "Voluntary pension savings: the effects of the finnish tax reform on savers' behaviour"
- 2009/23. Fehr, H.; Kindermann, F.: "Pension funding and individual accounts in economies with life-cyclers and myopes"
- 2009/24. Esteller-Moré, A.; Rizzo, L.: "(Uncontrolled) Aggregate shocks or vertical tax interdependence? Evidence from gasoline and cigarettes"
- 2009/25. Goodspeed, T.; Haughwout, A.: "On the optimal design of disaster insurance in a federation"
- 2009/26. Porto, E.; Revelli, F.: "Central command, local hazard and the race to the top"
- 2009/27. Piolatto, A.: "Plurality versus proportional electoral rule: study of voters' representativeness"
- 2009/28. Roeder, K.: "Optimal taxes and pensions in a society with myopic agents"
- 2009/29. Porcelli, F.: "Effects of fiscal decentralisation and electoral accountability on government efficiency evidence from the Italian health care sector"
- 2009/30. Troumpounis, O.: "Suggesting an alternative electoral proportional system. Blank votes count"
- 2009/31. Mejer, M.; Pottelsberghe de la Potterie, B.: "Economic incongruities in the European patent system"
- 2009/32. Solé-Ollé, A.: "Inter-regional redistribution through infrastructure investment: tactical or programmatic?"
- 2009/33. Joanis, M.: "Sharing the blame? Local electoral accountability and centralized school finance in California"
- 2009/34. Parcero, O.J.: "Optimal country's policy towards multinationals when local regions can choose between firm-specific and non-firm-specific policies"
- 2009/35. Cordero, J.M.; Pedraja, F.; Salinas, J.: "Efficiency measurement in the Spanish cadastral units through DEA"
- 2009/36. Fiva, J.; Natvik, G.J.: "Do re-election probabilities influence public investment?"
- 2009/37. Haupt, A.; Krieger, T.: "The role of mobility in tax and subsidy competition"
- 2009/38. Viladecans-Marsal, E.; Arauzo-Carod, J.M.: "Can a knowledge-based cluster be created? The case of the Barcelona 22@district"

2010

- 2010/1. De Borger, B.; Pauwels, W.: "A Nash bargaining solution to models of tax and investment competition: tolls and investment in serial transport corridors"
- 2010/2. Chirinko, R.; Wilson, D.: "Can Lower Tax Rates Be Bought? Business Rent-Seeking And Tax Competition Among U.S. States"
- 2010/3. Esteller-Moré, A.; Rizzo, L.: "Politics or mobility? Evidence from us excise taxation"
- 2010/4. Roehrs, S.; Stadelmann, D.: "Mobility and local income redistribution"

- 2010/5, **Fernández Llera, R.; García Valiñas, M.A.:** "Efficiency and elusion: both sides of public enterprises in Spain"
- 2010/6, **González Alegre, J.:** "Fiscal decentralization and intergovernmental grants: the European regional policy and Spanish autonomous regions"
- 2010/7, **Jametti, M.; Joanis, M.:** "Determinants of fiscal decentralization: political economy aspects"
- 2010/8, **Esteller-Moré, A.; Galmarini, U.; Rizzo, L.:** "Should tax bases overlap in a federation with lobbying?"
- 2010/9, **Cubel, M.:** "Fiscal equalization and political conflict"
- 2010/10, **Di Paolo, A.; Raymond, J.L.; Calero, J.:** "Exploring educational mobility in Europe"
- 2010/11, **Aidt, T.S.; Dutta, J.:** "Fiscal federalism and electoral accountability"
- 2010/12, **Arqué Castells, P.:** "Venture capital and innovation at the firm level"
- 2010/13, **García-Quevedo, J.; Mas-Verdú, F.; Polo-Otero, J.:** "Which firms want PhDs? The effect of the university-industry relationship on the PhD labour market"
- 2010/14, **Calabrese, S.; Eppe, D.:** "On the political economy of tax limits"
- 2010/15, **Jofre-Monseny, J.:** "Is agglomeration taxable?"
- 2010/16, **Dragu, T.; Rodden, J.:** "Representation and regional redistribution in federations"
- 2010/17, **Borck, R.; Wimbersky, M.:** "Political economics of higher education finance"
- 2010/18, **Dohse, D.; Walter, S.G.:** "The role of entrepreneurship education and regional context in forming entrepreneurial intentions"
- 2010/19, **Åslund, O.; Edin, P-A.; Fredriksson, P.; Grönqvist, H.:** "Peers, neighborhoods and immigrant student achievement - Evidence from a placement policy"
- 2010/20, **Pelegrín, A.; Bolance, C.:** "International industry migration and firm characteristics: some evidence from the analysis of firm data"
- 2010/21, **Koh, H.; Riedel, N.:** "Do governments tax agglomeration rents?"
- 2010/22, **Curto-Grau, M.; Herranz-Loncán, A.; Solé-Ollé, A.:** "The political economy of infrastructure construction: The Spanish "Parliamentary Roads" (1880-1914)"
- 2010/23, **Bosch, N.; Espasa, M.; Mora, T.:** "Citizens' control and the efficiency of local public services"
- 2010/24, **Ahamdanech-Zarco, I.; García-Pérez, C.; Simón, H.:** "Wage inequality in Spain: A regional perspective"
- 2010/25, **Folke, O.:** "Shades of brown and green: Party effects in proportional election systems"
- 2010/26, **Falck, O.; Heblich, H.; Lameli, A.; Südekum, J.:** "Dialects, cultural identity and economic exchange"
- 2010/27, **Baum-Snow, N.; Pavan, R.:** "Understanding the city size wage gap"
- 2010/28, **Molloy, R.; Shan, H.:** "The effect of gasoline prices on household location"
- 2010/29, **Koethenbueger, M.:** "How do local governments decide on public policy in fiscal federalism? Tax vs. expenditure optimization"
- 2010/30, **Abel, J.; Dey, I.; Gabe, T.:** "Productivity and the density of human capital"
- 2010/31, **Gerritse, M.:** "Policy competition and agglomeration: a local government view"
- 2010/32, **Hilber, C.; Lyytikäinen, T.; Vermeulen, W.:** "Capitalization of central government grants into local house prices: panel data evidence from England"
- 2010/33, **Hilber, C.; Robert-Nicoud, F.:** "On the origins of land use regulations: theory and evidence from us metro areas"
- 2010/34, **Picard, P.; Tabuchi, T.:** "City with forward and backward linkages"
- 2010/35, **Bodenhorn, H.; Cuberes, D.:** "Financial development and city growth: evidence from Northeastern American cities, 1790-1870"
- 2010/36, **Vulovic, V.:** "The effect of sub-national borrowing control on fiscal sustainability: how to regulate?"
- 2010/37, **Flamand, S.:** "Interregional transfers, group loyalty and the decentralization of redistribution"
- 2010/38, **Ahlfeldt, G.; Feddersen, A.:** "From periphery to core: economic adjustments to high speed rail"
- 2010/39, **González-Val, R.; Pueyo, F.:** "First nature vs. second nature causes: industry location and growth in the presence of an open-access renewable resource"
- 2010/40, **Billings, S.; Johnson, E.:** "A nonparametric test for industrial specialization"

