Document de travail de la série

Etudes et Documents

E 2009.20

Zipf's Law for Cities: On a New Testing Procedure[1]

Sébastien Terra
CERDI-CNRS, Université d'Auvergne
sebastien.terra@u-clermont1.fr

Octobre 2009

[1] The author thanks Grégroire Rota-Graziosi for his valuable comments and suggestions.

Zipf's Law for Cities: On a New Testing Procedure[☆]

Sébastien Terra[a]

[a] *Centre d'Études et de Recherches sur le Développement International (CERDI), 65 boulevard François Mitterrand, 63000 Clermont-Ferrand, France*

Abstract

In this paper, we provide a new framework to assess the validity of Zipf's Law for cities. Zipf's Law states that, within a country, the distribution of city sizes follows a Pareto distribution with a Pareto index equal to 1. We adopt a two-step approach where we formally test if the distribution of city sizes is a Pareto distribution and then we estimate the Pareto index. Through Monte Carlo experiments, we investigate the finite sample performances of this testing procedure and we compare the small-sample properties of a new estimator (the minimum variance unbiased estimator) to those of commonly used estimators. The minimum variance unbiased estimator turns out to be more efficient and unbiased. We use this two-step approach to examine empirically the validity of Zipf's Law on a sample of 115 countries. Zipf's Law is not rejected in most countries (62 out of 115, or 53.9%).

Key words: Zipf's Law, Pareto distribution, Monte Carlo study, Minimum variance unbiased estimator, Developing countries

JEL: C13, C16, R12

---

1. Introduction

The interest for Zipf's Law for cities and the rank-size rule is quite old. Auerbach (1913) and Zipf (1949) were among the first to postulate the existence of an empirical regularity between the population (size) of a city and its rank in the urban hierarchy. Zipf's Law stipulates that, within a country, the distribution of city sizes follows a Pareto distribution with a Pareto index equal to 1. The cross-country investigation of Rosen and Resnick (1980) initiated a series of empirical studies in the 1980s and early 1990s (Alperovich, 1984, 1988; Cameron, 1990; Guérin-Pace, 1995). The seemingly validation of Zipf's Law as an empirical regularity and the inadequacy of existing urban theories in explaining it gave birth to the "mystery of urban hierarchy" (Krugman, 1996). Gabaix (1999a,b) renewed the interest of economists in Zipf's Law in two main areas. The first research area has pursued the exploration of the empirical validity of Zipf's Law (Soo, 2005;

[☆] The author thanks Grégroire Rota-Graziosi for his valuable comments and suggestions.
Email address: sebastien.terra@u-clermont1.fr (Sébastien Terra)

Nitsch, 2005; Soo, 2007; Le Gallo and Chasco, 2008). On the other hand, the second research area has sought to provide theoretical underpinnings to Zipf's Law (Duranton, 2006; Mansury and Gulyas, 2007; Rossi-Hansberg and Wright, 2007; Cordoba, 2008). However, Gan et al. (2006) suggested that Zipf's Law was a mere statistical artifact resulting from a spurious regression. They concluded that Zipf's Law "does not require an economic theory that determines city-size distributions".

Zipf's Law is a useful approach to analyze and describe of the distribution of city sizes. Other measures can be used to describe this distribution: primacy indexes or concentration indexes (Herfindhal or Gini indexes for instance). However, as Soo (2007) suggested, Zipf's Law has been extensively studied for two main reasons. On the one hand, unlike primacy measures, Zipf's Law and the Pareto index provide information on the distribution of the urban system beyond the largest cities. On the other hand, the Pareto index is easy to interpret: it is closely related to the Gini coefficient. High values of the Pareto index imply that city sizes are more uniform whereas low values of the index reveal that the population is more concentrated in a few cities. Moreover, as in Soo (2005) and Ioannides et al. (2008), the Pareto index can be used in investigations of the effects of political variables (Ades and Glaeser, 1995) and geographical features (Brakman et al., 1999; Fujita et al., 1999) on urban structure.

The analysis of the validity of Zipf's Law crucially depends on the estimation of the Pareto index. The choice of an appropriate econometric methodology has been the cornerstone of empirical studies on Zipf's Law. The Pareto index has traditionally been estimated using the OLS estimator. However, using a Monte Carlo study, Gabaix and Ioannides (2004) showed that this estimator was biased in small samples and proposed to use Hill's estimator (Hill, 1975). In an empirical cross-country investigation, Soo (2005) used both methods: his results suggested that Hill's estimator might not be reliable in small samples. Recently, Gabaix and Ibragimov (2007) derived a simple and practical correction to the OLS estimator to minimize its bias in small samples.

In this paper, we review and extend previous works on the choice of an appropriate estimator for the Pareto index. In particular, we develop a new testing procedure to assess the validity of Zipf's Law, based on a two-step approach. In the first step, we use goodness-of-fit tests to determine if the distribution of city sizes within a country follows a Pareto distribution. This first step has seldom been considered in previous studies. In the second step, we estimate the Pareto index and test if it is equal to 1. This second step corresponds to what has been used as a test of the validity of Zipf's Law in previous studies. To estimate the Pareto index, we propose to use a new estimator: the minimum variance unbiased estimator. Through Monte Carlo experiments, we investigate the finite sample performances of the suggested testing procedure and we compare the small-sample properties of this estimator to those of commonly used estimators. The minimum variance unbiased estimator turns out to be more efficient and unbiased. When this estimator is used, the actual size of the testing procedure is close to the nominal size and the test has a high power as soon as the sample size is larger than 50. Consequently, we apply this two-step approach to examine the validity of Zipf's Law on a sample of 115 countries. For each country, we focus on the upper tail of the city size distribution[1]. Zipf's Law is not rejected in most countries (62 out of 115, or 53.9%). Particularly, Zipf's Law is not rejected in many developing countries, especially in Africa and South America. However, Zipf's Law is rejected in most Asian countries and developed countries.

The article is novel for three main reasons. First, contrary to most previous empirical studies, it investigates the validity of Zipf's Law in an integrated framework allowing to test if city sizes follow a Pareto distribution. In most previous studies, Zipf's Law was rejected if the Pareto index was different from 1[2]. However, if the distribution of city sizes is not a Pareto distribution, the meaning of the estimate of the Pareto index is unclear. For instance, in some countries (such as Benin, China or the United States), the Pareto index is not different from 1 but the distribution of city sizes does not follow a Pareto distribution. Using the usual approach, we would not reject Zipf's Law whereas it ought to be. More generally, our results show that the usual approach based only on the value of the Pareto index may be misleading. Indeed, with the usual approach, we would conclude that Zipf's Law is not rejected in 76 countries (66.1%) whereas our testing procedure suggest that Zipf's Law is not rejected in only 62 countries (53.9%).

Then, the estimator that we use to estimate the Pareto index (the minimum variance unbiased estimator) is unbiased and more efficient than traditional estimators especially in small samples. The OLS estimator provides biased results and under-estimated standard errors: it over-rejects Zipf's Law. On the other hand, the method suggested by Gabaix and Ibragimov (2007) suffers from its low precision leading to a lack of power in detecting departures from Zipf's Law. The minimum variance unbiased estimator proposed in this study overcomes these two caveats and provides a more reliable value. Moreover, our approach avoids the pitfalls of the log-log plot which is frequently used to study city size distributions (Eeckhout, 2009; Levy, 2009).

Lastly, we use a sample that is more comprehensive than previous ones. Rosen and Resnick

[1] Eeckhout (2009) recently showed that the Pareto distribution was appealing to study the upper tail of the city size distribution

[2] Eeckhout (2004) and Gan et al. (2006) are notable exceptions; goodness-of-fit tests are used in both studies.

(1980) studied a sample of 44 countries while the sample of Soo (2005) contained 73 countries. Here, the urban systems of 115 countries, covering the period 1970-2009, are analyzed.

This paper is structured as follows. Section 2 describes the methodology and the two-step approach followed to test the validity of Zipf's Law. Section 3 presents the results of the Monte Carlo study investigating the performance of the testing procedure and comparing the small-sample properties of various estimators of the Pareto index. In section 4, we present the data and the main results concerning the empirical validity of Zipf's Law. Finally, section 5 concludes.

2. Empirical methodology

Our empirical methodology is based on a two-step approach. The first step consists in a formal test (a goodness-of-fit test) of the hypothesis that the distribution of city sizes is a Pareto distribution. In the second step, the Pareto index is estimated and a test of its equality to 1 is implemented.

2.1. Goodness-of-fit tests for Pareto distributions

Before describing the goodness-of-fit tests, a reminder of the main properties of the Pareto distribution is in order[3]. The probability density function f, the cumulative distribution function F and the survival function S of a variable X following a Pareto distribution $\mathcal{P}(k, \alpha)$ are:

$$\begin{aligned}
f(x) &= \frac{\alpha k^{\alpha}}{x^{\alpha+1}}, x \geq k > 0 \\
F(x) &= 1 - \left(\frac{x}{k}\right)^{-\alpha}, x \geq k > 0 \\
S(x) &= 1 - F(x) = \left(\frac{x}{k}\right)^{-\alpha}, x \geq k > 0 \qquad (1)
\end{aligned}$$

where k is a scale parameter corresponding to the minimal value of the distribution and $\alpha > 0$ is a shape parameter measuring the thickness of the distribution tail. The parameter α is called the Pareto index (or Pareto exponent or Zipf's coefficient). Zipf's Law holds if $\alpha = 1$.

Goodness-of-fit tests allow to test if the empirical distribution of a variable (here city sizes) follows a known theoretical distribution (here a Pareto distribution). Several tests have been developed: the Kolmogorov-Smirnov test (Kolmogorov, 1933; Smirnov, 1948), the Cramér-von Mises test (Cramér, 1928; von Mises, 1931) and the Anderson-Darling test (Anderson and Darling, 1952, 1954; Stephens, 1974), to name just a few. The null hypothesis of these tests is that the postulated distribution is acceptable whereas the alternative hypothesis is that the data do

[3]Kleiber and Kotz (2003) and Johnson et al. (1994) present additional results and properties for the Pareto distribution.

not follow this distribution. Thus, the general structure of these tests is:

$$H_0 : F_n(x) = F(x;\theta)$$
$$H_1 : F_n(x) \neq F(x;\theta)$$

where F_n is the empirical cumulative distribution function, F is the postulated theoretical cumulative distribution function and θ is a vector of parameters.

To compute these test statistics, a measure of the distance between F_n and F is needed. This distance can be measured by a supremum norm or a quadratic norm. In the case of a supremum norm, the goodness-of-fit test is the Kolmogorov-Smirnov test. The Kolmogorov-Smirnov statistic D is defined by :

$$D = \sup_x |F_n(x) - F(x;\theta)|$$

In the case of a quadratic norm, the goodness-of-fit tests are the Cramér-von Mises and Anderson-Darling tests. The generic statistic of test in the Cramér-von Mises family is given by:

$$Q = n \int_{-\infty}^{\infty} (F_n(x) - F(x;\theta))^2 \, \psi(x) \, dx$$

When $\psi(x) = 1$, Q is the statistic of Cramér-von Mises, denoted by W^2. When $\psi(x) = F(x)\,(1 - F(x))$, Q is the statistic of Anderson and Darling, denoted by A^2. Computational details for these tests can be found in Čížek et al. (2005, chapter 13).

A particular difficulty arises since we want to test the hypothesis that a sample of size n has a given distribution function $F(x;\theta)$ with unknown θ, where $\theta = (k, \alpha))$ in a Pareto distribution. Therefore, we first need to estimate the vector θ, which makes it necessary to adjust the critical values of the tests. Ross (2006, chapter 9) advocates the use of Monte Carlo simulations in this context and suggests the following procedure. First, θ is estimated from the sample and the required test statistic (D, W^2 or A^2) is computed and denoted by d. Then R samples of size n are generated from the distribution $F(x;\hat{\theta})$. For each simulated sample r ($r = 1, \ldots, R$), the parameter vector $\hat{\theta}_r^{SIM}$ is estimated and the test statistic d_r is calculated assuming that the sample is distributed according to $F(x;\hat{\theta}_r^{SIM})$. The p-value p is obtained as the proportion of times that this test statistic is at least as large as d:

$$p = \frac{1}{R} \sum_{r=1}^{R} \mathbb{1}_{[d_r \geq d]}$$

where $\mathbb{1}_A$ is the indicator function that equals 1 if event A occurs.

2.2. Estimators of the Pareto index

Various estimators can be used to estimate the value of the Pareto index. These estimators are based either on a regression estimated by ordinary least squares (OLS), or on the maximum likelihood estimator.

2.2.1. *OLS estimator*

Historically, the estimation by OLS has been the first method used to assess the validity of Zipf's Law. The following regression is estimated by OLS:

$$\ln(R_i) = a - b\ \ln(P_i) + \varepsilon_i\ ,\ i = 1, \ldots, n \tag{2}$$

where P_i is the population of city i and R_i its rank in the decreasing hierarchy of cities. This equation can easily be derived when we assume that the size distribution of cities is a Pareto distribution. As Eeckhout (2004) showed, the rank of a city in the empirical distribution is given by $R = nS(P)$, where n is the number of cities in the sample and $S(P)$ is the probability that a city has a population greater than P (where S is the survival function of a Pareto distribution). Using equation (1), the previous equation can be rewritten as:

$$R = n\left(\frac{P}{k}\right)^{-\alpha},$$

$$\text{or}\ \ \ln R = a - \alpha \ln P$$

where $a = \ln n + \alpha \ln k$ is a constant. The parameter b in equation (2) therefore corresponds to the Pareto index: $b = \alpha$.

2.2.2. *Corrected OLS estimator*

The previous estimator remains a popular approach to estimate the Pareto index. However, as Gabaix and Ioannides (2004) showed, this approach provides biased estimations in small samples. Gabaix and Ibragimov (2007) derived a simple and practical solution to correct this bias. They proposed to use the variable $R_i - \frac{1}{2}$ instead of R_i, and therefore to estimate equation (3) instead of equation (2):

$$\ln(R_i - 1/2) = a - b \ln(P_i) \tag{3}$$

The Pareto index is estimated by b. Gabaix and Ibragimov (2007) showed that the transformation 1/2 is optimal since it minimizes the small-sample bias of the OLS estimator. The valid standard error of the estimated parameter $\hat{b}$ is $\hat{b}\sqrt{\frac{2}{n}}$.

2.2.3. *Hill's estimator*

To correct the bias of the OLS estimator in small samples, Gabaix and Ioannides (2004) proposed to use the estimator derived by Hill (1975). Denoting $P_{(i)}$ the population of the city of rank i, with $P_{(1)} \geq \ldots \geq P_{(n)}$, Hill's estimator is defined as:

$$\hat{\alpha}^{HILL} = \frac{n-1}{\sum_{i=1}^{n-1} \ln P_{(i)} - \ln P_{(n)}} \tag{4}$$

The standard error of the estimated parameter $\hat{\alpha}^{HILL}$ is

$$\sigma\left(\hat{\alpha}^{HILL}\right) = \hat{\alpha}^{HILL}\left(\frac{\sum_{i=1}^{n-1}\left(\tau_i - 1/\hat{\alpha}^{HILL}\right)^2}{n-2}\right)^{\frac{1}{2}}(n-1)^{-\frac{1}{2}} \quad (5)$$

where $\tau_i = i\left(\ln P_{(i)} - \ln P_{(i+1)}\right)$. Gabaix and Ioannides (2004) suggested that the small-sample properties of this estimator may be "very worrisome". The empirical results of Soo (2005) confirmed this analysis.

2.2.4. Maximum likelihood estimator

The maximum likelihood estimator can easily be determined when we assume that the size distribution of cites is a Pareto distribution. The maximum likelihood estimators of α and k are:

$$\hat{\alpha}^{MLE} = n\left[\sum_{i=1}^{n}\ln\left(\frac{x_i}{\hat{k}^{MLE}}\right)\right]^{-1}$$
$$\hat{k}^{MLE} = \min_i x_i$$

Quandt (1966) showed that $\hat{\alpha}^{MLE}$ and $\hat{k}^{MLE}$ are consistent. However, in finite samples, these estimators are not unbiased:

$$E\left(\hat{\alpha}^{MLE}\right) = \frac{n\alpha}{n-2}, n > 2$$
$$E\left(\hat{k}^{MLE}\right) = \frac{n\alpha k}{n\alpha - 1}, n > \frac{1}{\alpha}$$

The statistic $\frac{2N\alpha}{\hat{\alpha}^{MLE}}$ follows a χ^2 distribution with $2(n-1)$ degrees of freedom (Johnson et al., 1994). The 95% confidence interval for α is:

$$\left[\frac{\hat{\alpha}^{MLE}}{2N}\chi^2_{2(n-1),0.025}, \frac{\hat{\alpha}^{MLE}}{2N}\chi^2_{2(n-1),0.975}\right]$$

2.2.5. Minimum variance unbiased estimator

The maximum likelihood estimators of α and k are biased in small samples even if they are asymptotically unbiased. However, Likeš (1969) and Baxter (1980) derived minimum variance unbiased estimators for these parameters:

$$\hat{\alpha}^{MVU} = \left(1 - \frac{2}{n}\right)\hat{\alpha}^{MLE}$$
$$\hat{k}^{MVU} = \left(1 - \frac{1}{(n-1)\hat{\alpha}^{MLE}}\right)\hat{k}^{MLE}$$

The statistic $\frac{2N\alpha}{\hat{\alpha}^{MVU}}$ follows a χ^2 distribution with $2(n-1)$ degrees of freedom (Likeš, 1969). The 95% confidence interval for α is:

$$\left[\frac{\hat{\alpha}^{MVU}}{2N}\chi^2_{2(n-1),0.025}, \frac{\hat{\alpha}^{MVU}}{2N}\chi^2_{2(n-1),0.975}\right]$$

3. A Monte Carlo investigation of the small-sample properties of the testing procedure

We use a Monte Carlo investigation to study the small-sample properties of the testing procedure described in section 2. This investigation allows us to evaluate the actual size (or significance level) of the test (how often does the test reject the null hypothesis H_0 when it is true?) and its power (how often does the test reject H_0 when it is false?). Since a major issue in the literature on Zipf's Law is the choice of an appropriate estimator for the Pareto index, we also investigate the small-sample properties of the estimators described in paragraph 2.2. The asymptotic properties of these estimators are well-known : they are consistent and (at least asymptotically) unbiased (if the model is not misspecified). However, these asymptotic results do not preclude finite sample bias. A Monte Carlo investigation allows us to compare these estimators on the basis of bias and precision.

3.1. The design of the simulation study

We assume that the data generating process is a Pareto distribution $\mathcal{P}(k, \alpha)$. For a sample of size n, we generate n realizations of a random variable following this distribution. Random numbers are generated with the inverse cumulative method (Train, 2003, chapter 9). Random numbers ε following a Pareto distribution are generated as:

$$\varepsilon = \frac{k}{(1-\mu)^{1/\alpha}} \tag{6}$$

where μ is a random draw from a standard Uniform distribution. The sample size varies between experiments with four different sample sizes: 20, 50, 100 and 200. The parameter α also varies between experiments. Three values are used for α : 0.6, 1 and 1.4. For each Monte Carlo experiment, the number of replications is 2,000.

3.2. The small-sample properties of the estimators

As the performances of the testing procedure depend on the properties of the estimators used in the second step, we first present the results of the Monte Carlo investigation of the small-sample bias and precision of the estimators of the Pareto index.

Three criteria are used to compare the small-sample properties of the various estimators. We denote by θ the true value of the parameter of interest and $\tilde{\theta}_r^i$ its value estimated using estimator i at replication r $(r = 1, \ldots, R)$ where R is the total number of replications. The first criterion is the bias of the estimator i (B^i) for the parameter θ:

$$B^i = \frac{1}{R}\sum_{r=1}^{R}\tilde{\theta}_r^i - \theta = \overline{\tilde{\theta}^i} - \theta \tag{7}$$

The second criterion is the variance of the estimator i (V^i):

$$V^i = \frac{1}{R-1}\sum_{r=1}^{R}\left(\tilde{\theta}_r^i - \overline{\tilde{\theta}^i}\right)^2 \tag{8}$$

The last criterion is the mean squared error (MSE^i):

$$MSE^i = \frac{1}{R}\sum_{r=1}^{R}\left(\tilde{\theta}_r^i - \theta\right)^2 \tag{9}$$

Table 1 presents detailed numerical values of the bias and precision of the various estimators for $\alpha = 1$ and $k = 10,000$[4]. For a small sample size ($n = 20$), the bias of some estimators is substantial: it approaches 10% for the OLS and maximum likelihood estimators. Consistently with the results of Gabaix and Ioannides (2004), the OLS estimator underestimates the value of the Pareto index. The bias of Hill's estimator and the bias of the corrected OLS estimator of Gabaix and Ibragimov (2007) are more moderate (respectively 6.2% and 4.8%). Finally, the bias of the minimum variance unbiased estimator is very small (-0.6%). As for the mean squared error, the estimator of Gabaix and Ibragimov (2007) has the highest mean squared error due to a relatively poor precision of this estimator. Quite predictably, the minimum variance unbiased estimator has the smallest mean squared error. For larger samples ($n = 200$), the bias of most estimators is negligible: the OLS estimator is the only biased one and underestimates the true value of the Pareto index by about 3.9%. The minimum variance unbiased estimator still has the smallest mean squared error and is virtually unbiased. These Monte Carlo simulations suggest that the minimum variance unbiased estimator has the best small-sample properties, both in terms of bias and mean squared error.

[Table 1 about here.]

3.3. The small-sample properties of the testing procedure

Two criteria are used to investigate the properties of the testing procedure: the size and the power of the test. The size (or level) of the test corresponds to the type I error, that is to say to an incorrect rejection of H_0 when H_0 is true. The power of the test is related to the type II error and measures the probability of rejecting H_0 when it is false. With Monte Carlo studies, the true data generating process is known. Therefore, we can investigate if the testing procedure rejects H_0 when $\alpha = 1$ (to investigate the size of the test) or when $\alpha = 0.6, 1.4$ (to investigate the power of the test). The empirical level of the test is the proportion of rejections of H_0 when it is true while the empirical power is the proportion of rejections of H_0 when it is false.

[4]Complete results are available from the author upon request.

Table 2 presents the empirical size of the testing procedure for different sample sizes. Since this testing procedure is a sequence of two tests with a nominal significance size of 5% for each test, the nominal size of the overall testing procedure is 10% (by Bonferroni inequality). Using the OLS estimator in the second step of the procedure yields a very high type I error: we reject H_0 in at least 75% of cases while it is in fact true. The corrected OLS estimator of Gabaix and Ibragimov (2007) under-rejects H_0 with the Kolmogorov-Smirnov and Cramér-von Mises tests but over-rejects it with the Anderson-Darling test: these results may be due to the lower precision of this estimator. For the other estimators (Hill's estimator, maximum likelihood estimator, and minimum variance unbiased estimator), the empirical size of the test is close to the nominal size for moderately large samples.

[Table 2 about here.]

Table 3 presents the empirical power of the testing procedure for different sample sizes. Since the actual size of the test is larger than the nominal size for the OLS estimator, the evaluation of the power is deeply flawed: the empirical power of the test using this estimator will not be commented. Using the corrected OLS estimator in the second step of the testing procedure yields a lower power than using the Hill's estimator, the maximum likelihood estimator or the minimum variance unbiased estimator. The lower precision of the corrected OLS estimator results in larger confidence intervals leading too often to a non-rejection of H_0 while it is indeed false. When $\alpha = 0.6$, the testing procedure using the three other estimators has a relatively high power, especially for the minimum variance unbiased estimator. When $\alpha = 1.4$, the power of the testing procedure is lower; a reasonable power (80%) is achieved when n is larger than 100.

[Table 3 about here.]

The results of these Monte Carlo simulations suggest that the performance of the testing procedure heavily depends on the choice of the estimator in the second step of the procedure. When the OLS estimator or the corrected OLS estimator of Gabaix and Ibragimov (2007) are used, the probability that a type I or a type II error is committed is very high. The use of these estimators would lead to unreliable results. When the minimum variance unbiased estimator is used, the actual size of the test is very close to the nominal size when n is larger than 50; the procedure also has a high power to detect departures from Zipf's Law when n is greater than 50.

4. Empirical results on the validity of Zipf's Law

This section presents the empirical results on the validity of Zipf's Law on a sample of 115 countries. We follow the two-step approach described in section 2. We first test if the empirical distribution of city sizes follows a Pareto distribution. We then use the approach followed in previous studies and investigate whether the Pareto index is equal to 1. Finally, we bring together these two steps to formally test the validity of Zipf's Law.

4.1. Data

The data on city populations come from Brinkhoff (2008). This source of data from this website has previously been used by Soo (2005) and Ioannides et al. (2008). Soo (2005) discusses the issues of the reliability of the data and the delineation of cities. For some countries, the data was supplemented with information coming from Statistics Finland (2008). In this database, all cities with more than 100,000 inhabitants are included (with a threshold of 90,000 inhabitants for European countries). The population of capitals and main cities of each country is also available in this dataset.

We focus on the upper tail of the city size distribution, with a minimum population threshold of 10,000 inhabitants per city. This population threshold varies between countries. To include a country in the sample, an urban network of at least 20 cities was needed. Therefore, the sample contains data on 115 countries, covering the period 1970-2009. For each country, city populations are often available for several years: the maximal number of available years is 5, for an average number of 2.7. The total number of observations (country $\times$ year) is 305. Our dataset is thus more comprehensive than the sample used by Soo (2005) who studied 73 countries for the period 1972-2001 (yielding 197 observations).

4.2. Step 1: Does the size distribution of cities follow a Pareto distribution?

In this paragraph, we present the empirical results of goodness-of-fit tests: does the empirical distribution of city sizes follow a Pareto distribution? We use the testing procedure outlined in section 2.1, with the Anderson-Darling test[5]. Table 4 summarizes the results for the latest available year[6] classifying for each continent the countries according to the results of goodness-of-fit tests: Pareto distribution (null hypothesis) rejected at the 5% significance level or Pareto distribution not rejected. Globally, in 26 countries out of 115 (22.6%), the null hypothesis is rejected at the 5% significance level: in these countries, the urban structure cannot be represented

[5] The results obtained with the Kolmogorov-Smirnov test or the Cramér-von Mises test are very similar.

[6] Complete results are available from the author upon request.

by a Pareto distribution. On the contrary, in most countries, the hypothesis that city sizes are Pareto distributed is not rejected.

The existence of primate cities in a country may explain why the distribution of city sizes does not follow a Pareto distribution[7]. For instance, in Costa Rica, the urban structure does not follow a Pareto distribution: its urban hierarchy is dominated by the city of San José with 346,800 inhabitants in 2006 while the second largest city (Limón) only has 63,500 inhabitants.

[Table 4 about here.]

4.3. Step 2: Is the Pareto index equal to 1?

This section presents the estimates of the Pareto index for the 115 countries in the sample. In table 5, usual descriptive statistics for the various estimators are presented. The mean value for the Pareto index estimated with the OLS estimator is 1.041. For the corrected OLS estimator, the mean is 1.130. The mean of the estimates with the minimum variance unbiased estimator is 1.039. Globally, theses results are slightly inferior to the values obtained in previous studies. Rosen and Resnick (1980) found that the mean value for the Pareto index (using the OLS estimator) was 1.136. The results obtained by Soo (2005) are somewhat similar: the mean value with the OLS estimator is 1.11, while the mean value with Hill's estimator is 1.17.

[Table 5 about here.]

In table 6, we present the results of the following test: is the Pareto index different from 1 at the 5% significance level? The estimates obtained with the OLS estimator suggest that the Pareto index is significantly superior to 1 in 49 countries (42.6%) and inferior to 1 in 46 (40.0%). On the contrary, the results with the corrected OLS estimator show that the Pareto index is not different from 1 in 98 countries (85.2%). This is a consequence of the low precision of the estimator, as shown in the Monte Carlo simulations described in section 3. For the minimum variance unbiased estimator, the Pareto index is not significantly different from 1 in 76 countries (66.1%), while it is inferior to 1 in 19 countries and higher than 1 in 20 countries.

These results on the closeness of the Pareto index to 1 are quite different from those presented in previous studies (Rosen and Resnick, 1980; Soo, 2005). These two studies showed that the Pareto index was often different from 1. Rosen and Resnick (1980) found that the Pareto index was higher than 1 in 32 countries out of 44 (72.7%). Soo (2005) suggested that the Pareto index was superior to 1 in 39 countries out of 73 (53.4%) and inferior to 1 in 14 countries (19.2%).

[7]Box 7.3 in Brakman et al. (2001) contains an interesting discussion on primate cities and Zipf's Law.

[Table 6 about here.]

Two reasons may explain the differences between our results and the findings of previous studies. First, our sample is more comprehensive than previous ones. Rosen and Resnick (1980) studied a sample of 44 countries while the sample of Soo (2005) contained 73 countries. Here, the urban systems of 115 countries are analyzed. One difference is the coverage of Africa: the sample used by Soo (2005) included 10 African countries while ours contains 29 African countries. Africa has a particular urban pattern characterized by a relatively low mean value of the Pareto index.

Indeed, analyzing the results by continent (tables 7 and 8), we can see that the Pareto index is, on average, very close to 1 in Africa. Furthermore, the Pareto index is not significantly different from 1 in 23 African countries (79.3%). Europe is characterized by a higher value of the Pareto index. This result is consistent with the findings of Soo (2005). South America (with a mean value of 0.934 for the minimum variance unbiased estimator) and Asia-Oceania (0.965) have a Pareto index generally inferior to 1. These results suggest that the urban population in European countries is more evenly distributed than predicted by Zipf's Law. In Asia and South America, the population is more concentrated in some cities.

[Table 7 about here.]

[Table 8 about here.]

To further investigate the impact of the definition of the sample, we can restrict the analysis to a sub-sample corresponding to the countries present in the sample of Soo (2005)[8]. Using the minimum variance estimator, the Pareto index is higher than 1 in 16 countries out of 71 (23.9%) and inferior to 1 in 15 countries (22.4%). The Pareto index is therefore not significantly different from 1 in most countries (53.7%). Using similar countries but a different estimator, the results are different from those of Soo (2005) but are consistent with our results on a larger sample.

The second and principal explanation can be found in the econometric approach. The usual approach (the OLS estimator) provides biased results and under-estimated standard errors, which may explain why the Pareto index is more often different from 1. The method suggested by Gabaix and Ibragimov (2007) seems to be too imprecise (see the Monte Carlo results in section 3), and therefore leading too often to a non-rejection of the difference of the Pareto index from 1. The minimum variance unbiased estimator proposed in this study overcomes the two caveats and provides a more reliable value.

[8]Two countries present in the sample of Soo (2005) are not included in our sample: Mozambique (due to an insufficient number of reliable data for city populations) and Yugoslavia.

4.4. *Summary: Is Zipf's Law rejected?*

This paragraph synthesizes the two steps described in the previous paragraphs. It provides the empirical results of the test of the validity of Zipf's Law. In the previous paragraphs, the significance level of the tests was 5%. Since we are performing multiple significance tests, by Bonferroni inequality, the error rate for the test of the validity of Zipf's Law is 10%.

Table 9 reports the results by continent. The comparison with table 7 shows that the usual approach based only on the value of the Pareto index may be misleading. Indeed, with the usual approach, we would conclude that Zipf's Law is not rejected in 76 countries (66.1%) whereas our testing procedure suggest that Zipf's Law is not rejected in only 62 countries (53.9%). For instance, in some countries (such as Benin, China or the United States), the Pareto index is not different from 1 but the city sizes are not Pareto distributed: hence, Zipf's Law is rejected.

In Africa, Zipf's Law is not rejected in 55.2% of the countries. Thirteen African countries do not verify Zipf's Law. Some, such as Ivory Coast, Tunisia or Benin, have an urban system that cannot be described by a Pareto distribution. Two countries have a Pareto index inferior to 1: Zimbabwe (0.632) and Zambia (0.745). In these two countries, the urban population is more concentrated in few cities than predicted by Zipf's Law. On the contrary, in three countries (South Africa, Ethiopia, Madagascar), the urban population is more evenly distributed. For instance, in 2001, in South Africa, the population of the five main cities are relatively close: Johannesburg has 1,009,000 inhabitants, Soweto 858,600, Kapkaupunki 827,200, Durban 536,600 and Pretoria 525,400.

In Europe and South America, Zipf's Law is also frequently not rejected. On the contrary, in Asia, it is largely rejected.

[Table 9 about here.]

5. Conclusion

In this paper, we proposed a new testing procedure to assess the validity of Zipf's Law. Previous studies (Rosen and Resnick, 1980; Soo, 2005; Nitsch, 2005) suggested that Zipf's Law was rejected in a majority of countries. On the contrary, our results show that Zipf's Law is not rejected in 55% of the countries in the sample. On average, the Pareto index is very close to 1 (1.039).

Two main reasons may explain these differences. On the one hand, our study focused on a more comprehensive sample of 115 developed and developing countries, during the period 1970-2009. Compared with previous studies, developing countries are better represented in our dataset and Zipf's Law is seldom rejected in developing countries (notably in Africa and South

America), which may explain the lower rejection of Zipf's Law. On the other hand, we formally tested the validity of Zipf's Law using a new testing procedure. Our empirical approach rests on an estimator which is virtually unbiased even in small samples, thus circumventing the pitfalls encountered in previous studies.

Moreover, our results suggest that Zipf's Law has a real empirical content. In a recent paper, Gan et al. (2006) suggested that Zipf's Law was a statistical artifact resulting from a spurious regression (regressing the rank of a city, calculated from its population, on its population). However, our results suggest the contrary. Using a rigorous approach, based on the distribution of city sizes (and not on the ranks of cities), we show that Zipf's Law is not rejected in most countries. In other words, the need to provide theoretical foundations for Zipf's Law remains on the research agenda.

References

Ades, A. F., Glaeser, E. L., 1995. Trade and Circuses: Explaining Urban Giants. Quarterly Journal of Economics 110 (1), 195–227.

Alperovich, G., 1984. The size distribution of cities: On the empirical validity of the rank-size rule. Journal of Urban Economics 16 (2), 232–239.

Alperovich, G., 1988. A new testing procedure of the rank size distribution. Journal of Urban Economics 23 (2), 251–259.

Anderson, T. W., Darling, D. A., 1952. Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. The Annals of Mathematical Statistics 23 (2), 193–212.

Anderson, T. W., Darling, D. A., 1954. A Test of Goodness of Fit. Journal of the American Statistical Association 49 (268), 765–769.

Auerbach, F., 1913. Das gesetz der bevolkerungskoncentration. Petermanns Geographische Mitteilungen 59, 74–76.

Baxter, M., 1980. Minimum Variance Unbiased Estimation of the Parameters of the Pareto Distribution. Metrika 27 (1), 133–138.

Brakman, S., Garretsen, H., Van Marrewijk, C., 2001. An Introduction to Geographical Economics: Trade, Location and Growth. Cambridge University Press, New York.

Brakman, S., Garretsen, H., Van Marrewijk, C., van den Berg, M., 1999. The Return of Zipf: Towards a Further Understanding of the Rank-Size Distribution. Journal of Regional Science 39 (1), 183–213.

Brinkhoff, T., 2008. City population. `http://www.citypopulation.de`.

Cameron, T. A., 1990. One-stage structural models to explain city size. Journal of Urban Economics 27 (3), 294–307.

Čížek, P., Härdle, W., Weron, R., 2005. Statistical tools for finance and insurance. Springer, Berlin.

Cordoba, J.-C., 2008. On the Distribution of City Sizes. Journal of Urban Economics 63 (1), 177–197.

Cramér, H., 1928. On the Composition of Elementary Errors: II. Statistical Applications. Skandinavisk Aktuarietidskrift 11, 141–180.

Duranton, G., 2006. Some Foundations for Zipf's Law: Product Proliferation and Local Spillovers. Regional Science and Urban Economics 36 (4), 542–563.

Eeckhout, J., 2004. Gibrat's Law for (All) Cities. American Economic Review 94 (5), 1429–1451.

Eeckhout, J., 2009. Gibrat's law for (All) cities: Reply. American Economic Review 99 (4), 1676–1683.

Fujita, M., Krugman, P., Venables, A. J., 1999. The Spatial Economy: Cities, Regions, and International Trade. MIT Press, Cambridge, MA.

Gabaix, X., 1999a. Zipf's Law and the Growth of Cities. American Economic Review 89 (2), 129–132.

Gabaix, X., 1999b. Zipf's Law for Cities: An Explanation. Quarterly Journal of Economics 114 (3), 739–767.

Gabaix, X., Ibragimov, R., 2007. Rank-1/2: a simple way to improve the OLS estimation of tail exponents. NBER Working Paper.

Gabaix, X., Ioannides, Y. M., 2004. The evolution of city size distributions. In: Henderson, J. V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics. Vol. 4. Elsevier, pp. 2341–2378.

Gan, L., Li, D., Song, S., 2006. Is the Zipf law spurious in explaining city-size distributions? Economics Letters 92 (2), 256–262.

Guérin-Pace, F., 1995. Rank-Size Distribution and the Process of Urban Growth. Urban Studies 32 (3), 551–562.

Hill, B. M., 1975. A Simple General Approach to Inference About the Tail of a Distribution. The Annals of Statistics 3 (5), 1163–1175.

Ioannides, Y. M., Overman, H. G., Rossi-Hansberg, E., Schmidheiny, K., 2008. The effect of information and communication technologies on urban structure. Economic Policy 23 (54), 201–242.

Johnson, N. L., Kotz, S., Balakrishnan, N., 1994. Continuous univariate distributions, 2nd Edition. Vol. 1. John Wiley & Sons, New York.

Kleiber, C., Kotz, S., 2003. Statistical Size Distributions in Economics and Actuarial Sciences. Wiley series in probability and statistics. John Wiley & Sons, Hoboken, NJ.

Kolmogorov, A., 1933. Sulla determinazione empirica di una legge di distribuzione. Giornale dell'Istituto Italiano degli Attuari 4, 83–91.

Krugman, P., 1996. Confronting the Mystery of Urban Hierarchy. Journal of the Japanese and International Economies 10 (4), 399–418.

Le Gallo, J., Chasco, C., 2008. Spatial Analysis of Urban Growth in Spain, 1900-2001. Empirical Economics 34 (1), 59–80.

Levy, M., 2009. Gibrat's law for (All) cities: Comment. American Economic Review 99 (4), 1672–1675.

Likeš, J., 1969. Minimum Variance Unbiassed Estimates of the Parameters of Power-Function and Pareto's Distribution. Statistical Papers 10 (2), 104–110.

Mansury, Y., Gulyas, L., 2007. The Emergence of Zipf's Law in a System of Cities: An Agent-Based Simulation Approach. Journal of Economic Dynamics and Control 31 (7), 2438–2460.

Nitsch, V., 2005. Zipf Zipped. Journal of Urban Economics 57 (1), 86–100.

Quandt, R., 1966. Old and New Methods of Estimation and the Pareto Distribution. Metrika 10 (1), 55–82.

Rosen, K. T., Resnick, M., 1980. The size distribution of cities: An examination of the Pareto law and primacy. Journal of Urban Economics 8 (2), 165–186.

Ross, S. M., 2006. Simulation, 4th Edition. Elsevier Academic Press, Amsterdam ; Boston.

Rossi-Hansberg, E., Wright, M. L. J., 2007. Urban Structure and Growth. Review of Economic Studies 74 (2), 597–624.

Smirnov, N., 1948. Table for Estimating the Goodness of Fit of Empirical Distributions. The Annals of Mathematical Statistics 19 (2), 279–281.

Soo, K. T., 2005. Zipf's Law for cities: a cross-country investigation. Regional Science and Urban Economics 35 (3), 239–263.

Soo, K. T., 2007. Zipf's Law and Urban Growth in Malaysia. Urban Studies 44 (1), 1–14.

Statistics Finland, 2008. City and region database. `http://tilastokeskus.fi/tup/maanum/taulukot_en.html`.

Stephens, M. A., 1974. EDF Statistics for Goodness of Fit and Some Comparisons. Journal of the American Statistical Association 69 (347), 730–737.

Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, New York.

von Mises, R., 1931. Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und Theoretischen Physik. Deuticke.

Zipf, G., 1949. Human Behavior and the Principle of Least Effort. Addison-Wesley Press, Cambridge, MA.

Table 1: Small-sample properties of estimators - Monte Carlo simulations for $\alpha = 1$ and $k = 10,000$ (2,000 replications)

Estimator[a]	Mean	Absolute Bias	Relative Bias (%)	Variance	MSE
			n=20		
OLS	0.900	-0.100	-9.998	0.084	0.094
OLS2	1.048	0.048	4.819	0.110	0.112
HILL	1.062	0.062	6.200	0.075	0.079
MLE	1.118	0.118	11.790	0.083	0.097
MVU	1.006	0.006	0.611	0.067	0.067
			n=50		
OLS	0.920	-0.080	-8.041	0.034	0.040
OLS2	1.007	0.007	0.689	0.039	0.039
HILL	1.020	0.020	1.969	0.023	0.023
MLE	1.041	0.040	4.050	0.024	0.026
MVU	0.999	-0.001	-0.112	0.022	0.022
			n=100		
OLS	0.940	-0.060	-6.039	0.018	0.022
OLS2	0.997	-0.003	-0.278	0.020	0.020
HILL	1.005	0.005	0.516	0.011	0.011
MLE	1.015	0.015	1.532	0.011	0.011
MVU	0.995	-0.005	-0.499	0.010	0.010
			n=200		
OLS	0.961	-0.039	-3.881	0.009	0.011
OLS2	0.998	-0.002	-0.152	0.009	0.009
HILL	1.006	0.006	0.561	0.005	0.005
MLE	1.011	0.011	1.066	0.005	0.005
MVU	1.001	0.001	0.055	0.005	0.005

[a] OLS = OLS estimator
OLS2 = corrected OLS estimator
HILL = Hill's Estimator
MLE = maximum likelihood estimator
MVU = minimum variance unbiased estimator

Table 2: Empirical size (%) of the test[a] - Monte Carlo simulations for $\alpha = 1$ and $k = 10,000$ (2,000 replications)

Estimator[b]	$n = 20$	$n = 50$	$n = 100$	$n = 200$
Kolmogorov-Smirnov test				
OLS	75.80	82.30	87.45	88.60
OLS2	7.30	7.65	7.95	7.60
HILL	12.75	11.10	10.80	9.65
MLE	10.35	10.50	10.30	9.00
MVU	13.35	10.95	10.60	9.20
Cramér-von Mises test				
OLS	75.75	82.35	87.45	88.60
OLS2	7.35	7.65	7.75	7.95
HILL	12.55	11.40	11.00	9.40
MLE	10.55	10.10	10.50	8.55
MVU	12.70	11.20	10.70	8.95
Anderson-Darling test				
OLS	80.65	84.75	89.25	89.75
OLS2	19.25	20.55	21.10	18.30
HILL	17.10	13.10	11.95	9.65
MLE	15.65	11.30	11.40	9.40
MVU	13.25	11.30	10.65	8.60

[a] The nominal size of the test is 10%.

[b] OLS = OLS estimator
OLS2 = corrected OLS estimator
HILL = Hill's Estimator
MLE = maximum likelihood estimator
MVU = minimum variance unbiased estimator

Table 3: Empirical power (%) of the test - Monte Carlo simulations for $\alpha = 0.6, 1.4$ and $k = 10,000$

Estimator[a]	$n = 20$	$n = 50$	$n = 100$	$n = 200$
		$\alpha = 0.6$		
Kolmogorov-Smirnov test				
OLS	96.90	99.80	100.00	100.00
OLS2	50.00	81.90	97.90	99.95
HILL	72.65	96.00	99.75	100.00
MLE	65.75	95.00	99.85	100.00
MVU	80.40	97.75	99.85	100.00
Cramér-von Mises test				
OLS	96.80	99.80	100.00	100.00
OLS2	50.45	82.15	97.85	99.95
HILL	72.95	95.90	99.75	100.00
MLE	65.90	95.00	99.80	100.00
MVU	80.25	97.70	99.85	100.00
Anderson-Darling test				
OLS	98.35	99.90	100.00	100.00
OLS2	62.70	88.90	99.00	100.00
HILL	74.35	96.10	99.75	100.00
MLE	67.35	95.15	99.80	100.00
MVU	80.35	97.70	99.85	100.00
		$\alpha = 1.4$		
Kolmogorov-Smirnov test				
OLS	77.15	90.60	96.90	99.75
OLS2	5.50	22.40	55.70	90.55
HILL	27.25	59.25	88.85	99.60
MLE	30.20	65.80	91.45	99.80
MVU	19.25	55.25	87.10	99.60
Cramér-von Mises test				
OLS	77.25	90.85	96.85	99.80
OLS2	5.75	22.50	56.00	90.60
HILL	26.40	59.30	88.85	99.60
MLE	30.15	65.60	91.40	99.80
MVU	18.80	54.90	87.00	99.60
Anderson-Darling test				
OLS	77.20	89.65	96.15	99.75
OLS2	17.10	26.35	54.05	88.60
HILL	29.95	60.25	89.20	99.65
MLE	33.75	66.30	91.55	99.80
MVU	18.65	55.10	87.00	99.60

[a] OLS = OLS estimator
OLS2 = corrected OLS estimator
HILL = Hill's Estimator
MLE = maximum likelihood estimator
MVU = minimum variance unbiased estimator

Table 4: Results of the Anderson-Darling goodness-of-fit test[a]

Region	H_0 not rejected	H_0 rejected	Total
Africa	21	8	29
America, North	1	1	2
America, South	15	4	19
Asia and Oceania	16	7	23
Europe	31	4	35
Middle East	5	2	7
World	89	26	115

[a] H_0 : the empirical distribution is a Pareto distribution.
H_1 : the empirical distribution is not a Pareto distribution.

Table 5: Estimates of the Pareto index (α)

Estimator[a]	Obs	Mean	Standard deviation	Minimum	Maximum
OLS	115	1.041	0.226	0.669	1.695
OLS2	115	1.130	0.236	0.728	1.887
HILL	115	1.059	0.252	0.523	1.766
MLE	115	1.080	0.255	0.530	1.778
MVU	115	1.039	0.250	0.517	1.754

[a] OLS = OLS estimator
OLS2 = corrected OLS estimator
HILL = Hill's Estimator
MLE = maximum likelihood estimator
MVU = minimum variance unbiased estimator

Table 6: Is the Pareto index (α) equal to 1 at the 5% confidence level?

Estimator[a]	$\alpha < 1$	$\alpha = 1$	$\alpha > 1$	Total
OLS	46	20	49	115
OLS2	1	98	16	115
HILL	16	78	21	115
MLE	16	76	23	115
MVU	19	76	20	115

[a] OLS = OLS estimator
OLS2 = corrected OLS estimator
HILL = Hill's Estimator
MLE = maximum likelihood estimator
MVU = minimum variance unbiased estimator

Table 7: Is the Pareto index (α) equal to 1 at the 5% confidence level? - Analysis by continent

Estimator[a]/ Region	$\alpha < 1$	$\alpha = 1$	$\alpha > 1$	Total
OLS				
Africa	11	8	10	29
America, North	1	0	1	2
America, South	12	3	4	19
Asia and Oceania	7	5	11	23
Europe	10	4	21	35
Middle East	5	0	2	7
OLS2				
Africa	0	27	2	29
America, North	0	1	1	2
America, South	0	17	2	19
Asia and Oceania	1	20	2	23
Europe	0	26	9	35
Middle East	0	7	0	7
HILL				
Africa	1	25	3	29
America, North	0	2	0	2
America, South	4	14	1	19
Asia and Oceania	7	12	4	23
Europe	3	20	12	35
Middle East	1	5	1	7
MLE				
Africa	1	23	5	29
America, North	0	2	0	2
America, South	4	13	2	19
Asia and Oceania	7	13	3	23
Europe	2	21	12	35
Middle East	2	4	1	7
MVU				
Africa	2	23	4	29
America, North	0	2	0	2
America, South	5	13	1	19
Asia and Oceania	7	13	3	23
Europe	3	21	11	35
Middle East	2	4	1	7

[a] OLS = OLS estimator
OLS2 = corrected OLS estimator
HILL = Hill's Estimator
MLE = maximum likelihood estimator
MVU = minimum variance unbiased estimator

Table 8: Estimates of the Pareto index (α) - Analysis by continent

Estimator[a]/ Region	Obs	Mean	Standard deviation	Minimum	Maximum
OLS					
Africa	29	0.997	0.225	0.700	1.693
America, North	2	1.098	0.223	0.940	1.256
America, South	19	0.966	0.217	0.669	1.695
Asia and Oceania	23	1.037	0.187	0.713	1.458
Europe	35	1.124	0.228	0.756	1.653
Middle East	7	1.010	0.307	0.767	1.610
OLS2					
Africa	29	1.097	0.235	0.811	1.825
America, North	2	1.112	0.214	0.960	1.263
America, South	19	1.045	0.240	0.728	1.887
Asia and Oceania	23	1.119	0.208	0.782	1.621
Europe	35	1.213	0.218	0.871	1.733
Middle East	7	1.118	0.345	0.814	1.779
HILL					
Africa	29	1.063	0.242	0.653	1.578
America, North	2	0.989	0.053	0.952	1.027
America, South	19	0.953	0.232	0.633	1.744
Asia and Oceania	23	0.984	0.209	0.523	1.333
Europe	35	1.189	0.254	0.769	1.766
Middle East	7	0.954	0.271	0.737	1.537
MLE					
Africa	29	1.088	0.248	0.675	1.604
America, North	2	0.991	0.051	0.954	1.027
America, South	19	0.969	0.235	0.654	1.778
Asia and Oceania	23	1.003	0.217	0.530	1.372
Europe	35	1.209	0.252	0.790	1.778
Middle East	7	0.980	0.277	0.754	1.578
MVU					
Africa	29	1.038	0.237	0.632	1.553
America, North	2	0.988	0.054	0.949	1.026
America, South	19	0.936	0.230	0.612	1.710
Asia and Oceania	23	0.965	0.203	0.517	1.323
Europe	35	1.168	0.257	0.748	1.754
Middle East	7	0.929	0.265	0.719	1.497

[a] OLS = OLS estimator
OLS2 = corrected OLS estimator
HILL = Hill's Estimator
MLE = maximum likelihood estimator
MVU = minimum variance unbiased estimator

Table 9: Is Zipf's Law rejected?[a]

Region	Zipf's Law not rejected	Zipf's Law rejected	Total
Africa	16	13	29
America, North	1	1	2
America, South	12	7	19
Asia and Oceania	10	13	23
Europe	20	15	35
Middle East	3	4	7
World	62	53	115

[a] In this table, results obtained with the minimum variance unbiased in the second step are presented.