Received: 11 March 2021 | Revised: 26 May 2021 | Accepted: 3 July 2021

DOI: 10.1111/rsp3.12449

ORIGINAL ARTICLE

The distribution of city sizes in Turkey: A failure of Zipf's law due to concavity

Hasan Engin Duran[1] | **Andrzej Cieślik**[2]

[1]Faculty of Architecture, Department of City and Regional Planning, IYTE Gülbahce Kampüsü, Mimarlık Fakültesi, Şehir ve Bölge Planlama Bölümü, E-Blok, Izmir Institute of Technology, Urla/Izmir, Turkey

[2]Faculty of Economic Sciences, University of Warsaw, Warszawa, Poland

Correspondence
Andrzej Cieślik, University of Warsaw, Faculty of Economic Sciences, Ul. Dluga 44/50, 00-241. Warszawa, Poland.
Email: cieslik@wne.uw.edu.pl

Abstract

The linearity of the distribution of city sizes is often assumed in the existing literature. Although different functional forms were tried, almost all of them impose a certain functional shape. In this study, we investigate the urban hierarchy and Zipf's law using data for 973 Turkish subprovincial cities in 2019. We contribute to the literature in several ways. We force no definite functional form to observe the natural shape and employ nonparametric and quadratic regressions. We incorporate formal procedures of spatial dependence in regression models. We demonstrate that the linear model overestimates the Pareto exponent for small cities and underestimates it for bigger cities. We show that city sizes are unevenly distributed in Turkey. The rank–size rule is not valid in Turkey, either above or below a certain city-size truncation level. Thus, the Pareto exponent estimated from the linear model is not a reliable indicator as quadratic regressions perform much better.

KEYWORDS
concavity, districts (subprovinces) of Turkey, nonparametric regressions, rank–size rule, Zipf's law

JEL CLASSIFICATION
O53; R12

1 | INTRODUCTION

The city size distribution and Zipf's law have long been a subject of interest in theoretical and empirical studies (Eaton & Eckstein, 1997; Gabaix, 1999; Giesen & Suedekum, 2011; Giesen, Zimmermann, & Suedekum, 2010; Ioannides & Overman, 2003; Luckstead & Devadoss, 2014; Moura & Ribeiro, 2006; Rosen & Resnick, 1980; Soo, 2007; Veneri, 2016; Zipf, 1949). Disproportional city sizes might lead to distortions in urban systems, such as segregation, congestion, crime, unemployment, economic efficiency losses, etc. (Arshad, Hu, & Ashraf, 2018; Lu & Wan, 2014). In recent decades, Turkey has experienced massive regional migrations and socioeconomic imbalances (Duran, 2015, 2019a, 2019b). Its population has increased from about 13 million in 1920s to about 83 million in 2019.[1] The rapid population growth has been accompanied by the increased urbanization rate. The share of population living in urban areas was about 25% during 1920s and 92% in 2019.[2] The internal migration from less to more developed cities accelerated urbanization processes after 1980, a year in which Turkey experienced a significant transition towards a liberal and more open economy (Coban, 2013; Deliktas, Onder, & Karadag, 2013; Evcil, Dokmeci, & Kıroglu, 2006; Gedik, 1997, 2003; Kundak & Dokmeci, 2018). During this period, industrialization around Istanbul and Izmir attracted many workers especially from eastern and northern parts of the country. A lack of jobs in less-developed areas contributed to this process. The inefficiency of the agricultural sector, which is the main economic activity in out-migrating regions, was one of the major reasons behind this migration process.

The main goal of this study is to investigate the city size distribution, the validity of Zipf's law, and the rank–size rule using the dataset that covers 973 Turkish subprovinces for the year 2019. It is well known that validity of Zipf's law is strictly related to scaling issue. It is crucial that the scale being too small or too big will affect the results of Zipf's law. Our spatial units, districts (subprovinces), are at the medium scale, not as big as provinces and not as small as neighborhoods.

Zipf's law can be seen as an important initial step in understanding the urban structure and required policies in Turkey. It is an essential analysis in estimating the validity of the rank–size rule as it theoretically predicts the Pareto optimal distribution. The invalidity of linearity and empirical innovations brought by nonparametric analyses are also crucial aspects that are pursued in current empirical analyses. Therefore, in our study, we adopt a nonlinear framework that has not been extensively used in the literature. Although the Pareto optimality of urban systems is usually studied in the log-linear form, we relax this assumption to provide a more general analysis (Arshad, Hu, & Ashraf, 2018; Cordoba, 2008a, 2008b; Duranton, 2007; Gabaix, 1999; Gabaix & Ibragimov, 2007; Lee & Li, 2013; Soo, 2005). By incorporating nonlinearities, the Pareto exponent is allowed to differ across different groups of cities, which enables more relevant and region/city-specific policy implications.

The paper is structured as follows. In section 2, a related literature is surveyed. In section 3, empirical methodologies and dataset properties are explained. In section 4, empirical and econometric results are documented. Then, finally, section 5 summarizes and concludes with directions for further studies and policy recommendations.

2 | LITERATURE REVIEW

The reasons why Zipf's law may or may not hold have been hotly debated in the literature. There are two opposing views. The first view supports the general validity of the law. There are two main hypotheses on the reasons for its validity. The first and the major reason is related to **random growth** of city sizes. The city sizes are argued to grow randomly when shocks to urban populations (e.g., migration, productivity, etc.) occur randomly. Hence, the distribution of city sizes is supposed to be in the Pareto state (Arshad, Hu, & Ashraf, 2018; Cordoba, 2008a, 2008b; Duranton, 2007; Gabaix, 1999; Gibrat, 1931; Lee & Li, 2013; Rossi-Hansberg & Wright, 2007). Similarly, it is stated that Pareto distribution is the result of an evolution that occurs naturally in a steady state, without an underlying theory (Arshad, Hu, & Ashraf, 2018; Batty, 2006; Corominas-Murtra & Solé, 2010; Gan, Li, & Song, 2006; Mansury & Gulyás, 2007; Semboloni, 2001). An alternative but a less popular hypothesis is related to the evolution of **human**

capital that leads to a proportional distribution of city sizes (Arshad, Hu, & Ashraf, 2018; Behrens, Duranton, & Robert-Nicoud, 2014; Christaller, 1933; Hsu, 2012).

With regard to theoretical studies supporting the counter argument, it is often emphasized that city sizes might significantly deviate from the Pareto distribution owing to many reasons. As an explanation of labor migrations towards urban giants, job incentives, higher wages, industrialization, education, and other facilities are often referred to as important drivers (Coban, 2013; Deliktas, Onder, & Karadag, 2013; Evcil, Dokmeci, & Kıroglu, 2006; Gedik, 1997, 2003). The migration from rural to urban areas due to a decline of the agricultural sector as well as economic and natural shocks may also be responsible for deviations from the Pareto distribution (Coban, 2013; Deliktas, Onder, & Karadag, 2013; Evcil, Dokmeci, & Kıroglu, 2006).The empirical studies are also far from reaching a consensus. These studies focus on estimated Pareto exponents that show the degree of closeness of a population distribution to the Pareto optimality. The exponent equal to unity represents the Pareto optimal distribution, whereas deviations from unity in any direction imply disproportionality. The empirical studies that support Zipf's law and validate the rank–size distribution include Gabaix (1999) for US Metropolitan Areas, Giesen and Suedekum (2011) for German cities/urban areas, and Luckstead and Devadoss (2014) for Indian cities. In addition, Nota and Song (2012) analyze the largest cities in the United States in 2000 and find modest deviations of Pareto exponent (0.89) from 1. Similarly, they study also China in the same context for the year 1999 and again find low level of deviations of Pareto exponent (1.09) from 1.

At the same time, there are many other studies that reject the Pareto distribution. Some important examples include Rosen and Resnick (1980), who find that, in 31 (of 44) countries, the Pareto exponent is above 1.00, and for all the countries, it ranged between 0.81 and 1.96; Guerin-Pace (1995), who detects deviations from the Pareto optimality for some of French cities; and Soo (2005), who finds that, in 39 (of 73) countries, the Pareto exponent is above 1.00, whereas in 14 countries, it is below 1.00, and ranging between 0.72 and 1.70 overall. Moreover, Bee, Ricabboni, and Schiavo (2013) and Lalanne (2014) are among the other studies that provide evidence against the applicability of Zipf's law for the United States and Canada, respectively. In a subsequent study, Modica, Reggiani, and Nijkamp (2015) analyze Botswana, Germany, Hungary, and Luxembourg. As a result, in all of these cases, Zipf's law and the Pareto distribution are not supported empirically. In the case of Botswana, the years 2001 and 2011 are analyzed and the Pareto coefficient is found to exceed unity; 1.13 and 1.17. Similarly, in the case of Germany, the years between 1993 and 2007 are analyzed. The Pareto exponent is found to be quite high (1.4) and stable over the years. In the case of Hungary, the parameter is estimated between 1.13 and 1.26 for some years between 1980 and 2011. In the case of Luxembourg, the Pareto exponent is found far lower, ranging between 0.5 and 0.91 for the selected years between 1821 and 2011. Finally, a very recent study by Bajracharya and Sultana (2020) finds that Zipf's law tends to fail in the case of Bangladesh as the related coefficient decreases from 1.015 in 1991 to 0.85 in 2011.

However, the existing studies are criticized for many reasons. First, it is argued that, due to improper **sample selection**, estimation results may be biased. When the entire sample of cities is studied, Zipf's law generally fails (Arshad, Hu, & Ashraf, 2018; Fazio & Modica, 2015). It is often claimed that the upper (higher) tail of an urban hierarchical distribution fits Zipf's law better (Eeckhout, 2004; Fazio & Modica, 2015; Ioannides & Skouras, 2013). Therefore, it is suggested to choose a correct truncation level of a population or to take into account only a certain fraction of cities (Arshad, Hu, & Ashraf, 2018; Li & Sui, 2013; Rosen & Resnick, 1980; Wheaton & Shishido, 1981). Another argument relates to the **improper definition of cities**. When administrative boundaries are used, it is found that Zipf's law fits the data poorly. However, when functional urban areas are employed, the goodness of fit improves greatly (Arshad, Hu, & Ashraf, 2018; Berry & Okulicz-Kozaryn, 2012). As another concern, **small sample sizes** may constitute a problem that creates a bias. In particular, Gabaix and Ioannides (2004) find that the Pareto exponent is biased in small samples. Therefore, robust estimators are proposed to deal with small samples (Arshad, Hu, & Ashraf, 2018; Deliktas, Onder, & Karadag, 2013; Dobkins & Ioannides, 2001; Gabaix & Ibragimov, 2007; Hill, 1975). Finally, also the **assumption of linearity** has been criticized. Soo (2005) adopts a quadratic version of the law and tests the nonlinearity for 73 countries. In his study, convexity is evident for 30 countries, and concavities are found for 20 others. Moreover, alternative functional forms/types are present (Ausloss & Cerquetti, 2016; Chen, 2016).

The number of empirical studies on the city size distribution in Turkey is still very limited. Nevertheless, there are a few notable exceptions. Deliktas, Onder, and Karadag (2013) focus on 81 Turkish provinces during the period 1980–1997 to find that the Pareto exponent ranges between 0.87 and 0.97. However, Duran and Ozkan (2015) find that Zipf's law fails in more recent years. Hence, findings for Turkey are far from clear cut, and there is a need for additional studies. Our study targets several methodological improvements with respect to existing studies. We relax the linearity assumption to allow for potential nonlinearities to avoid the possible estimation bias. Moreover, although different functional forms were proposed in the literature, the majority of them tend to ignore spatial dependence, and almost all impose a certain functional shape. In contrast to previous studies, we incorporate spatial autocorrelation in our estimated models. Furthermore, in our study, we force no definite functional form in order to be able to observe the natural shape and, thus, apply a nonparametric estimation. Having done this, we observe a clear strong concavity and, therefore, apply quadratic regressions. Although linear forms are widely criticized, there exists no study showing explicitly the magnitude of the bias created by an incorrect functional form. We attempt to investigate this issue in our study by calculating the deviations from the optimal Pareto distribution of city sizes and computing the mean absolute error (MAE) and mean absolute percentage error (MAPE) driven by different functional forms (de Myttenaere, Golden, Le Grand, & Rossi, 2016; Wilmott & Matsuura, 2005).

3 | METHODOLOGY AND DATASET

In the literature, Zipf's law is used to evaluate the optimality of city-size distributions through the following equation, also called the rank–size rule (Auerbach, 1913; Gabaix, 1999; Rosen & Resnick, 1980; Singer, 1936; Soo, 2005; Zipf, 1949):

$$R_i = AP_i^{-\beta}, \tag{1}$$

where: R_i represents the i-th city rank, P_i represents its population, and A is the expected population of the most populated city.

The estimated parameter β, known as the Pareto exponent, is claimed to indicate the proportionality of the distribution (Arshad, Hu, & Ashraf, 2018; Deliktas, Onder, & Karadag, 2013; Gabaix & Ioannides, 2004; Nitsch, 2005). When $\beta = 1$,the cities are supposed to follow the Pareto distribution. In such a case, city sizes follow the exact proportionality: the second-biggest city has half of the population of the biggest city, the third city has one-third the population of the biggest city, etc. (Arshad, Hu, & Ashraf, 2018). However, when $\beta < 1$, the distribution of the city sizes is uneven, as they deviate from the Pareto distribution (Soo, 2005; Deliktaş, Onder and Karadag, 2013; Arshad, Hu, & Ashraf, 2018). Similarly, when $\beta > 1$, the population is more evenly distributed than when compared with the Pareto state (Soo, 2005; Deliktas, Onder and Karadag, 2013; Arshad, Hu, & Ashraf, 2018).

Our empirical methodology includes some aspects of the traditional stream in the literature (i.e., the linear form of Zipf's law) but has also some new features (i.e., nonlinear methods). In short, we follow three main steps: (i) a linear test of Zipf's law, (ii) nonlinear modeling, (iii) testing robustness of estimations against spatial dependence, and (iv) calculation of the bias implied by either model.

3.1 | Linear methods

In terms of the research methodology, following many previous studies, we start with the log-linearized form of the rank–size rule (Gabaix & Ioannides, 2004; Nitsch, 2005; Deliktaş, Onder and Karadag, 2013):

$$\ln(R_i) = \alpha - \beta \ln(P_i) + u_i \tag{2}$$

where *i* represents 973 Turkish cities.

However, the ordinary least squares (OLS) estimation of Equation 2 is known to be biased for small samples (Hill, 1975; Gabaix & Ioannides, 2004; Deliktaş, Onder and Karadag, 2013). Thus, the following rank-minus-half form as proposed by Gabaix and Ibragimov (2007) is suggested in the literature (Deliktas, Onder and Karadag, 2013):

$$\ln(R_i - 1/2) = \alpha - \beta \ \ln \ (P_i) + v_i \tag{3}$$

In our study, Equations 2 and 3 are run using simple OLS regressions together with several diagnostic tests.

3.2 | Nonlinear methods

The nonlinearity in Zipf's law is studied in three possible ways. First, we fit a nonparametric Kernel regression by assuming a normal distribution and polynomial function of order 2 (Cleveland & Devlin, 1988; Fan & Gijbels, 1996; Fan, Heckman, & Wand, 1995; Fan & Marron, 1994; Härdle, 1991; Henderson & Parmeter, 2015; Simonoff, 1996). It estimates the polynomial regression for many small subsamples and combines them afterwards. The advantage of nonparametric regression is that, in contrast to parametric estimations, it does not force any certain functional shape of Zipf's law. The second method is the estimation of models for above and below a threshold level of populations detected by applying Bai and Perron (1998)'s multiple breakpoint test to Equations 2 and 3 (Bai, 1997; Bai & Perron, 1998, 2003a, 2003b).

As a third method of nonlinear modeling, a squared term of an independent variable is added to Equations 2 and 3 to capture the nonlinearity.

$$\ln(R_i \text{or} \, R_i - 1/2) = \alpha + \beta \ \ln \ (P_i) + \gamma [\ln(P_i)]^2 + \varphi_i \tag{4}$$

Although there is no strong theoretical basis for this polynomial equation, it is pursued only to complement/reconfirm the result of nonlinearity/concavity obtained from threshold and nonparametric Kernel regressions. Higher polynomial orders could be tried such as third or fourth, but this would not improve the fit as, in the current form, R^2 is already high, at about 0.96–0.97. So, it is preferred to adopt the second-order polynomial function.

All estimation results obtained from nonlinear regressions are reported in four categories: (i) for total subprovincial populations, (ii) for urban populations of subprovinces, and by using both (iii) rank–size and (iv) rank-minus-half forms.

3.3 | Spatial models

A crucial concern on the technical feasibility is the spatial dependence in regional/city level studies. Ignorance of spatiality may provide misleading or biased results. Therefore, we find it essential. Hence, we estimate spatial models in order to check the robustness of the previous results with respect to this fact.

Hence, we employ Lagrange multiplier lag and error tests to Equations 2, 3, and 4 (Anselin, 1988, 2001; Anselin & Bera, 1998; Anselin, Bera, Florax, & Yoon, 1996; Anselin & Moreno, 2003; Anselin & Rey, 1991; Bera, Doğan, Taşpınar, & Leiluo, 2019; Le Sage, 2008). We use a 973 × 973 raw standardized inverse distance matrix constructed in ARCGIS/ARCMAP program, and tests are performed in R 3.53 by using SP, SPDATA SPDEP packages.[3]

Then, we estimate two types of spatial regressions. First, spatial dependence is assumed in the dependent variable. This model, named spatial autoregressive model (SAR) is expressed as follows (Anselin, 1988, 2001; Anselin &

Bera, 1998; Anselin, Bera, Florax, & Yoon, 1996; Anselin & Moreno, 2003; Anselin & Rey, 1991; Bera, Doğan, Taşpınar, & Leiluo, 2019; Le Sage, 2008):

$$\text{Linear case}: \ \ln(R_i \ or \ R_i - 1/2) = \alpha - \beta \ \ln \ (P_i) + \rho W \ln(R_i) + e_i \tag{5}$$

$$\text{Quadratic case}: \ \ln(R_i \ or \ R_i - 1/2) = \alpha - \beta \ \ln \ (P_i) + \gamma \ln(P_i)^2 + \rho W \ln(R_i) + e_i \tag{6}$$

where ρ captures the spatial spillovers in the dependent variable. W is the spatial weight matrix defined as 973×973 inverse distance matrix.

The second spatial regression is the spatial error model (SEM) that allows spatial interactions between the error terms of neighboring cities (Anselin, 1988, 2001; Anselin & Bera, 1998; Anselin, Bera, Florax, & Yoon, 1996; Anselin & Moreno, 2003; Anselin & Rey, 1991; Bera, Doğan, Taşpınar, & Leiluo, 2019; Le Sage, 2008):

$$\text{Linear case}: \ \ln(R_i \ \text{or} \ R_i - 1/2) = \alpha - \beta \ \ln \ (P_i) + e_i \ z_i = \theta W z_j \tag{7}$$

$$\text{Quadratic case}: \ \ln(R_i \ \text{or} \ R_i - 1/2) = \alpha - \beta \ \ln \ (P_i) + \gamma \ln(P_i)^2 + e_i \ z_i = \theta W z_j \tag{8}$$

where θ captures the spatial spillovers of residuals between neighboring cities.

We estimate Equations 5–8 in both linear and nonlinear settings.

3.4 | Bias calculation driven by linearity

We argue that a wrong functional form may result in serious biases in the estimation of the distribution of city sizes. In particular, it may over- or underestimate the deviations from optimality. Hence, one needs to investigate in detail the extent of the resulting biases. To do so, we calculate the deviations from the Pareto distribution of the city sizes and explore the size of the biases brought by different functional forms. In particular, we first let optimal ranks of provinces be denoted by $r_{pareto,i}$. According to $r_{pareto,i}$, the biggest city was in the first place ($i = 1$) and the others followed in the sequential order as $i = 2, 3, \ldots, n$.

We estimate Equations 5–8 (SAR and SEM models) and obtain the fitted values of the city populations. $r_{linear,i}$ and $r_{quadratic,i}$ represent the fitted-rank values obtained from linear and quadratic models. Then, we calculate the mean absolute errors (MAE) and mean absolute percentage Error (MAPE) in terms of the deviations from the Pareto distribution (de Myttenaere, Golden, Le Grand, & Rossi, 2016; Wilmott & Matsuura, 2005), such that:

$$MAE_{rank\_linear} = \sum\nolimits_{i=1}^{n} | r_{pareto,i} - r_{linear,i} | / n \tag{9}$$

$$MAE_{rank\_quadratic} = \sum\nolimits_{i=1}^{n} | r_{pareto,i} - r_{quadratic,i} | / n$$

$$MAPE_{rank\_linear} = \sum\nolimits_{i=1}^{n} | (r_{pareto,i} - r_{linear,i)/} r_{pareto_i} \ | / n$$

$$MAPE_{rank\_quadratic} = \sum\nolimits_{i=1}^{n} | \left(r_{pareto,i} - r_{quadratic,i} \right) / r_{pareto_i} | / n$$

Hence, the aforementioned MAE and MAPE specifications help demonstrate the extent of the bias driven by the linearity. In addition, the Kolmogorov–Smirnov test is used to examine whether two distributions of populations (or ranks) (i.e., the estimated one and Pareto optimal distribution) have identical or significantly different characteristics (Kolmogorov, 1933; Smirnov, 1948).

3.5 | Dataset

Our dataset covers 973 subprovincial Turkish cities for the most recent 2019 year. In terms of spatial units, we adopt two different types: (i) administrative units (subprovinces) and (ii) urban-only populations of subprovinces as a proxy of functional urban areas. The main reason why we use only post-2013 data is because a new 'Metropolitan Law' was initiated after 2012 that changed borders of many sub-provinces.

It is informative to illustrate the map of Turkish districts (subprovinces), which is presented in Figure 1. The larger districts are observed within metropolitan areas, particularly in Western and Southern urban giants such as Istanbul, Ankara, İzmir, Adana, Bursa, Antalya, etc. In contrast, smaller districts are rather observed in rural areas such as North and Eastern parts.

The dataset of this study is downloaded from the internet website of the Turkish Statistical Institute (TURKSTAT). The estimations and tests are performed in Eviews 4, Eviews 10, and R 3.53 program (SP, SPDATA, SPDEP packages).

4 | EMPIRICAL RESULTS

4.1 | Linear baseline results

The linear estimation of Zipf's law is done for Equations 2 and 3. The results are reported in Table 1. The beta coefficient is the main focus of our analysis. In the first two columns results from estimations in which 'rank' variable used as a dependent variable (i.e., Equation 2), and in the remaining columns, estimation results in which 'rank minus half' variable used as a dependent variable (i.e., Equation 3) are shown. The White–Hinkley (Hinkley, 1977; White, 1980) diagonal robust variance–covariance matrix is used.

The Pareto exponent (β) is significantly bigger than 0 at the 1% level in all cases regardless of whether the rank or the rank minus half is used. It is found to be about 0.68 for subprovinces and 0.56 for their urban populations. The Pareto exponent is found to be lower for urban population than total populations, indicating a greater unevenness across urban parts. The estimated exponents are always above 0.00 but also always significantly below 1.00, as

FIGURE 1 Turkish subprovinces (districts); source: https://www.harita.gov.tr/urun/turkiye-mulki-idare-sinirlari/232. The map is produced in ARCGIS, ARCMAP program

TABLE 1 OLS baseline linear regressions, source: Own estimation

Dependent variable:	Ln(*R*)		Ln(*R-1/2*)	
Parameters	Total population	Urban population	Total population	Urban population
α	12.98338***	11.53447***	13.03328***	11.5715***
$-(\beta)$	−0.680534***	−0.560304***	−0.685713***	−0.564385***
γ	–	–	–	–
Wald test (beta = −1) (χ^2 test stat)	343***	869***	296***	771***
White–Hinkley heteroskedasticity Consistent covariance and S.E.	Yes	Yes	Yes	Yes
R^2	0.86	0.83	0.86	0.82
Detected threshold of population (Bai & Perron, 1998)	163,773	150,488	163,773	156,787
N	973	973	973	973

Note:
*Statistical significance at 10%,
**at 5%,
***at 1%, valid also in the remaining parts of the paper. Note: S.E., standard error.

the Wald test statistic (*F*-value) is statistically significant at 1% in all the cases (Gourieroux, Holly, & Monfort, 1982). The detected breaks in Zipf's law are presented in lower rows. These break levels of populations range between 150.488 and 163.773.

We observe that the results do not change much when 'rank minus half' is used instead of 'rank' as the dependent variable. Therefore, from this point onwards, we do not report in the main text the results for which 'rank minus half' is used as the dependent variable (except Figure 2). Instead, we include them only in Appendix 1a–e. They seem mostly consistent with the estimations presented in the paper.

4.2 | Nonlinear model results

We report the results of nonparametric regressions in Figure 2 insipired from Nadaraya (1964) and Watson (1964). The scatterplots in Figure 2 are shown together with the Kernel fit curve. The population (in natural logs) is placed on the *x*-axis, whereas the rank variable (in natural logs) is placed on the *y*-axis. In all cases, regardless of whether the entire subprovince or only the urban population is analyzed, the results consistently display a strong and clear concavity, such that, on the one hand, the negative relationship between the population and the rank of cities is very strong for cities greater than a certain threshold. On the other hand, the relationship is milder when a subsample below a population threshold is considered.

As a second nonlinear modeling, the results of the threshold regressions are summarized in Table 2. The Zipf coefficient is estimated for below and above threshold levels.

For small cities (below threshold level), the estimated Pareto exponent ranges between 0.38 and 0.48, while for large cities, it ranges between 1.84 and 1.95 The results are consistent across urban and total populations and across different dependent variable specifications. The errors are found to follow a normal distribution. So, it appears that there is a remarkable difference in the estimated Pareto exponents between two groups of cities. This result indicates that there is no unique measure of heterogeneity in urban size distribution. Moreover, the upper and lower tails of the city distribution indicate very different outcomes. Both tails are far from the Pareto optimality, but the distribution among small cities is further from the Pareto optimality. In other words, in the lower tail, small cities are more heterogeneously distributed. The linear model seriously overestimates the Pareto exponent for small cities and

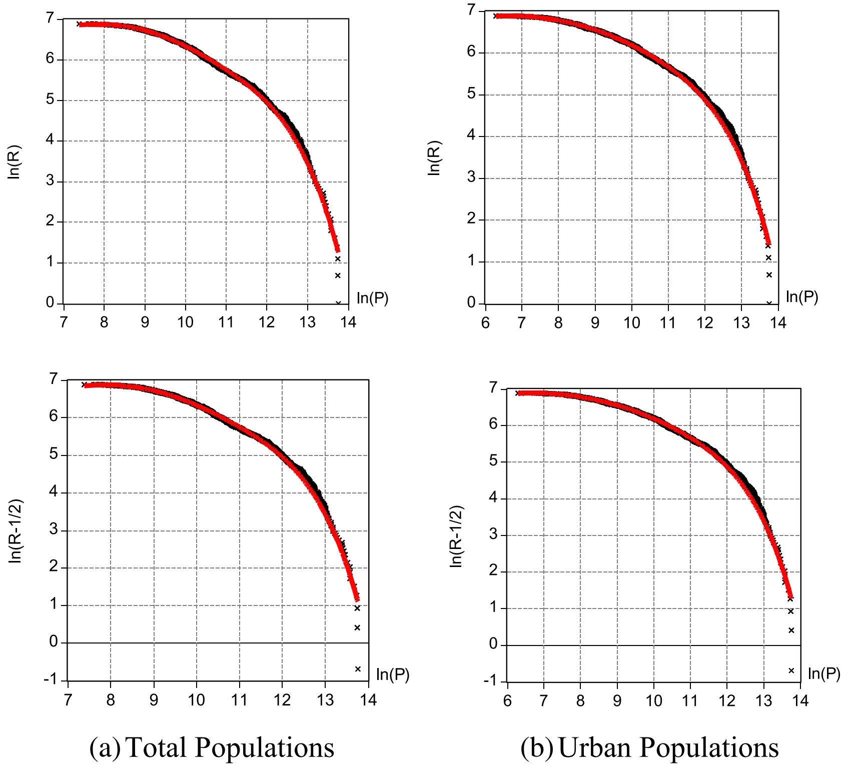

FIGURE 2 Nonparametric estimation of Zipf's law (kernel fit, polynomial order = 2, normal distribution assumed); source: own estimation

TABLE 2 Threshold regressions; source: own estimation

Dependent variable:	Total population		Urban population	
Parameters	Above threshold	Below threshold	Above threshold	Below threshold
c	28.755***	11.02387***	27.35586***	9.879821***
$-(\beta)$	−1.951***	−0.479143***	−1.845344***	−0.380271***
White–Hinkley heteroskedasticity Consistent covariance and S.E.	Yes	Yes	Yes	Yes
R^2	0.91	0.95	0.9	0.94
N	146	827	151	822

Note:

*Statistical significance at 10%,

**at 5%,

***at 1%, valid also in the remaining parts of the paper.

underestimated it for bigger cities. Consequently, Zipf's law does not hold for Turkey either above or below a certain threshold.

Finally, we estimate the quadratic regression model. Table 3 presents the empirical results for Equation 4. In all the estimated regressions, we observe positive and significant β and negative and significant γ. Statistical significance is always observed at the 1% level. Hence, previously observed concave evolution (in Figure 1) is evident since the parameters estimated in quadratic regressions confirm this shape.

4.3 | Spatial modeling

The results of spatial autocorrelation tests are reported in Table 4. The results are shown separately for urban and total populations.

In all the cases, strong spatial autocorrelations are observed. Almost all test statistics are positive and statistically significant. Thus, we decide to incorporate such strong spatial correlations by running the SAR and SEM models in both linear and nonlinear contexts.

In the upper panel (Table 5), the linear results are shown for urban and total populations, whereas the quadratic form is estimated in the bottom part. In all the cases, the estimated parameters of Zipf's law are consistent with the previous findings. The estimated parameters have the same signs and significance levels. The concavity is shown to be clearly present. Spatial components lambda or rho are positive and significant in almost all estimated specifications. As a result, it is clearly seen that previously estimated parameters are similar to the parameter estimated by spatial panel models. Therefore, we argue that estimations are robust with respect to presence of spatial interaction components.

4.4 | Calculation of bias and deviations from Pareto optimality

A final analysis is implemented to show the magnitude of the bias created by linear estimations of Zipf's law. The results of MAE and MAPE calculations, referring to Equations (9), are summarized in Table 6.

TABLE 3 Quadratic regressions; source: own estimation

Parameters	Total population	Urban population
α	−4.525348***	−1.30404**
β	2.661379***	2.015199***
γ	−0.156858***	−0.12596***
Wald test (beta = −1) (χ^2 test stat)	–	–
White–Hinkley heteroskedasticity Consistent covariance and S.E.	Yes	Yes
R^2	0.97	0.96
N	973	973

Note:
*Statistical significance at 10%,
**at 5%,
***at 1%, valid also in the remaining parts of the paper.

TABLE 4 Spatial autocorrelation tests; source: own estimation

	Ln(*R*)-linear		Ln(*R*)-quadratic	
Dependent variable	Total population	Urban population	Total population	Urban population
LMerr	491.93***	632.95***	2.8956*	14.979***
Lmlag	97.629***	121.96***	3.2846*	7.1551***
RLMerr	434.36***	554.18***	2.5241	13.598***
RLMlag	40.057***	43.187***	2.9131*	5.7743**

TABLE 5 Spatial autoregressive and spatial error models; source: own estimation

Linear	SAR		SEM	
Parameters	**Total population**	**Urban population**	**Total population**	**Urban population**
α	9.928205***	8.033571***	13.26099***	11.91035***
$-(\beta)$	−0.64253***	−0.52051***	−0.66283***	−0.54255***
γ				
Rho	0.45***	0.53***		
Lambda			0.97***	0.98***
N	973	973	973	973
Quadratic	**SAR**		**SEM**	
Parameters	**Total population**	**Urban population**	**Total population**	**Urban population**
α	−4.69879***	−1.63987***	−4.427***	−1.14235***
$-(\beta)$	2.638482***	1.988121***	2.641794***	1.98301***
γ	−0.1556***	−0.12436***	−0.15588***	−0.12428***
Rho	0.05*	0.075***		
Lambda			0.56	0.81***
N	973	973	973	973

TABLE 6 MAE, MAPE, and Kolmogorov–Smirnov tests: deviations from Pareto optimal distribution in terms of rank, mean absolute error, and Kolmogorov–Smirnov test; source: own estimation

SAR	Linear/ quadratic	Total/urban populations	MAE	MAPE	Kolmogorov–Smirnov (*D*-stat)	*p*-Value
	Linear	Total	0.223473	0.042224	0.14183***	6.32×10^{-9}
	Linear	Urban	0.253316	0.047449	0.148***	1.11×10^{-9}
	Quadratic	Total	0.09302	0.020508	0.10586***	3.68×10^{-5}
	Quadratic	Urban	0.106587	0.022794	0.10791***	2.40×10^{-5}
SEM	Linear	Total	0.219997	0.041558	0.13258***	7.47×10^{-8}
	Linear	Urban	0.248789	0.046557	0.14286***	4.76×10^{-9}
	Quadratic	Total	0.092841	0.020481	0.10586***	3.68×10^{-5}
	Quadratic	Urban	0.105437	0.022622	0.111***	1.24×10^{-5}

According to the linear model, the deviation from optimal ranks is large, such that MAE ranges between 0.22 and 0.25 whereas MAPE ranges between 0.04 and 0.05. This error is far greater than the one implied by the quadratic case in which deviation from optimality ranges between 0.09 and 0.11 (MAE) and 0.02% and 0.025% (MAPE). To compare the distributions, Kolmogorov–Smirnov test is conducted, and results are reported in Table 6 (Kolmogorov, 1933; Smirnov, 1948). The test statistics shows that these two distributions are statistically different from the Pareto ranks as the test (*D*) statistics are always significant at the 1% level. In other words, ranks implied by linear and quadratic functions are statistically different than Pareto ranks. However, differences are much more pronounced for the linear case as test statistics (*D*) are higher. Therefore, we argue that the linear model overestimates deviations from optimality. According to the nonlinear model, cities are actually closer to the Pareto optimality in Turkey. Hence, we can conclude that the linear model provides misleading results.

5 | CONCLUSIONS

Our study provided new findings regarding the urban system in Turkey. In particular, we found the linearity assumption to be too strong. Our estimations indicate that, in Turkey, cities follow a far different path than the Pareto optimal one. For small cities, the Pareto exponent is smaller than 1, whereas for the big cities it is much bigger than 1. On the one hand, this means that sizes of small cities (below a threshold) are very heterogeneously (unevenly) distributed, causing many social and economic problems for developed (unemployment, crime, pollution) or lagging-behind cities (underdevelopment, lack of investments). One the other hand, the sizes of big cities are evenly distributed among themselves, which is also not Pareto optimal. Therefore, the linearity assumption needs to be relaxed in future studies on city-size distributions. One may allow nonlinearity by adding either higher powers to regressions (i.e., a quadratic approach) or employing nonparametric regressions as they do not impose a definite functional form. Moreover, there is a need for deeper analysis of further stages, such as analyses on the determinants and causes of non-Pareto distributions, that were beyond the purpose of the current study.

With regard to urban policies, the uneven distribution of city sizes constitutes a very important economic and social policy problem. Our study shows that there is no unique degree of urban concentration and the Pareto exponent.

On the one hand, it implies that smaller cities in Turkey exhibit a greater deviation from optimality and more uneven distribution compared with bigger cities. Hence, policy recommendations to small and big cities should be separately considered. First, less-developed cities should be economically stimulated. Subsidies and investments should be increased in these places, and productive activities that create jobs should be promoted (i.e., local entrepreneurship, adaptation to the digital economy, and innovation). This is consistent with the new regionalism stream that suggests the determination of place-specific necessities and potentials by local actors (i.e., entrepreneurs, universities, etc.) (Hettne & Söderbaum, 1998). Related projects should be supported by the central government. Export orientation and smart specialization should be also encouraged. In this way, increased investments in less-developed cities could reverse existing migration patterns and mitigate uneven distributions of cities. Moreover, education facilities, transportation, and job incentives should be stimulated in remote places to reverse the existing migration patterns. Well-connected cities, educated workforce, infrastructure, and accessibility can enhance economic efficiency and help less-developed regions to evolve faster (Elburz, Nijkamp, & Pels, 2017; Lucas, 1988). Another relevant policy action could be to support industrialization in out-migrating cities. Since land and labor are less costly in these places, industrial investments can stimulate output and employment. Land reforms should be implemented to increase efficiency of the agricultural sector. Finally, reforms related to restructuring institutions, the promotion of urban services, and the diversity of consumer goods and services sector are also among policies that should be designed for less-developed areas.

On the other hand, several policy actions can be formulated for big cities. Our study implies that their size is more evenly distributed compared with small cities. Uniformity in large cities brings various socioeconomic and environmental problems. One of the most important problems is the high unemployment driven by massive net in migration to these cities and a rapid increase in the labor supply. Due to this problem, a high number of skilled workers become inutile. To mitigate this, job-creating industries should be promoted rather than sectors that create jobless growth. Technical skills of workers should be promoted via formal and informal education programs to reduce the frictional unemployment and the mismatch between employees and employers. Another important problem is negative environmental externalities induced by the agglomeration of economic activities in/around these large cities. This problem is particularly pronounced in highly industrialized cities. Uncontrolled industrialization is seen as quite dangerous as it triggers pollution in air, soil, and water. Hence, it becomes a serious threat to sustainability. So, in these cities, institutions, monitoring, and regulating mechanisms should function well to reduce such environmental problems including pollution, congestion, etc. By following the most effective of these policies, city sizes might more closely approach the Pareto optimum.

ORCID

Hasan Engin Duran https://orcid.org/0000-0002-0743-9943
Andrzej Cieślik https://orcid.org/0000-0002-7834-7384

ENDNOTES

[1] Data Source: Turkish Statistical Institute (www.tuik.gov.tr)

[2] Data Source: Turkish Statistical Institute (www.tuik.gov.tr)

[3] Pebesma et al. (2021) for SP package, Bivand, Nowosad, Lovalace, Monmonier, and Snow, 2020 for spData package, and Bivand et al., 2021 for SPDEP package. The related data and shape file were obtained from: https://www.harita.gov.tr/urun/turkiye-mulki-idare-sinirlari/232

REFERENCES

Anselin, L. (1988). *Spatial econometrics: Methods and models*. New York: Springer.

Anselin, L. (2001). Rao's score test in spatial econometrics. *Journal of Statistical Planning and Inference*, *97*(1), 113–139. https://doi.org/10.1016/S0378-3758(00)00349-9

Anselin, L., & Bera, A. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah & D. Giles (Eds.), *Handbook of applied economic statistics*. New York: Marcel Dekker.

Anselin, L., Bera, A. K., Florax, R., & Yoon, J. M. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, *26*(1), 77–104. https://doi.org/10.1016/0166-0462(95)02111-6

Anselin, L., & Moreno, R. (2003). Properties of tests for spatial error components. *Regional Science and Urban Economics*, *33*(5), 595–618. https://doi.org/10.1016/S0166-0462(03)00008-5

Anselin, L., & Rey, S. (1991). Properties of tests for spatial dependence in linear regression models. *Geographical Analysis*, *23*(2), 112–131.

Arshad, S., Hu, S., & Ashraf, B. N. (2018). Zipf's law and city size distribution: A survey of the literature and future research agenda. *Physica a: Statistical Mechanics and its Applications*, *492*(C), 75–92. https://doi.org/10.1016/j.physa.2017.10.005

Auerbach, F. (1913). Das Gesetz der Bevölkerungskonzentration. *Petermanns Geographische Mitteilungen*, *59*, 74–76.

Ausloss, M., & Cerquetti, R. (2016). A universal rank-size law. *PLoS ONE*, *11*(11), e0166011. https://doi.org/10.1371/journal.pone.0166011

Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory*, *13*(3), 315–352. https://doi.org/10.1017/S0266466600005831

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*(1), 47–78. https://doi.org/10.2307/2998540

Bai, J., & Perron, P. (2003a). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22. https://doi.org/10.1002/jae.659

Bai, J., & Perron, P. (2003b). Critical values for multiple structural change tests. *The Econometrics Journal*, *6*(1), 72–78. https://doi.org/10.1111/1368-423X.00102

Bajracharya, P., & Sultana, S. (2020). Rank-size distribution of cities and municipalities in Bangladesh. *Sustainability*, *12*(4643), 1–26. https://doi.org/10.3390/su12114643

Batty, M. (2006). Rank clocks. *Nature*, *444*, 592–596. https://doi.org/10.1038/nature05302

Bee, M., Ricabboni, M., & Schiavo, S. (2013). The size distribution of US cities: Not Pareto, even in the tail. *Economic Letters*, *120*(2), 232–237. https://doi.org/10.1016/j.econlet.2013.04.035

Behrens, K., Duranton, G., & Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, *122*(3), 507–553. https://doi.org/10.1086/675534

Bera, A. K., Doğan, O., Taşpınar, S., & Leiluo, Y. (2019). Robust LM tests for spatial dynamic panel data models. *Regional Science and Urban Economics*, *76*, 47–66. https://doi.org/10.1016/j.regsciurbeco.2018.08.001

Berry, B. J. L., & Okulicz-Kozaryn, A. (2012). The city size distribution debate: Resolution for US urban regions and megalopolitan areas. *Cities*, *29*(1), 17–23.

Bivand, R., Altman, M., Anselin, L., Assunção, R., Berke, O., Bernat, A., Blanchet, G., Blankmeyer, E., Carvalho, M., Christensen, B., Chun, Y., Dormann, C., Dray, S., Gómez-Rubio, V., Gubri, M., Halbersma, R., Krainski, E., Legendre, P., Lewin-Koh, N., Li, A., Li, H., Ma, J., Mallik, A., Millo, G., Mueller, W., Ono, H., Peres-Neto, P., Piras, G., Reder, M., Sauer, J., Tiefelsdorf, M., Westerholt, R., Wolf, L., & Yu, D. (2021). SPDEP package R. https://cran.r-project.org/web/packages/spdep/spdep.pdf

Bivand, R., Nowosad, J., Lovalace, R., Monmonier, M., & Snow, G. (2020). spData Package R. https://cran.r-project.org/web/packages/spData/spData.pdf.

Chen, Y. (2016). The evolution of Zipf's law indicative of city development. *Physica a: Statistical Mechanics and its Applications*, *443*, 555–567. https://doi.org/10.1016/j.physa.2015.09.083

Christaller, W. (1933). *Die zentralen Orte in Süddeutschland*. Jena: Gustav Fischer.

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, *83*(403), 596–610. https://doi.org/10.1080/01621459.1988.10478639

Coban, C. (2013). Different periods of internal migration in Turkey from the perspective of development. *American International Journal of Contemporary Research*, *3*(10), 58–65.

Cordoba, J. C. (2008a). On the distribution of city sizes. *Journal of Urban Economics*, *63*(1), 177–197. https://doi.org/10.1016/j.jue.2007.01.005

Cordoba, J. C. (2008b). A generalized Gibrat's law. *International Economic Review*, *49*(4), 1463–1468. https://doi.org/10.1111/j.1468-2354.2008.00518.x

Corominas-Murtra, B., & Solé, R. V. (2010). Universality of Zipf's law. *Physcial Review E*, *82*(1), 011102. https://doi.org/10.1103/PhysRevE.82.011102

de Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean absolute percentage error for regression models. *Neurocomputing*, *192*(5), 38–48. https://doi.org/10.1016/j.neucom.2015.12.114

Deliktas, E., Onder, A. O., & Karadag, M. (2013). The size distribution of cities and determinants of City growth in Turkey. *European Planning Studies*, *21*(2), 251–263. https://doi.org/10.1080/09654313.2012.722922

Dobkins, L. H., & Ioannides, Y. M. (2001). Dynamic evolution of the size distribution of US cities. In J.-M. Huriot & J.-F. Thisse (Eds.), *Economics of cities: Theoretical perspectives*. Cambridge: Cambridge Univ Press.

Duran, H. E. (2015). Non-linear regional income divergence and policies: Turkey case. *Regional Science Inquiry*, *7*(2), 107–114.

Duran, H. E. (2019a). Asymmetries in regional development: Does TFP or capital accumulation matter for spatial inequalities? *Journal of Economic Asymmetries*, *20*, e00119. https://doi.org/10.1016/j.jeca.2019.e00119

Duran, H. E. (2019b). Regional unemployment dynamics in Turkey. *Regional Science Inquiry*, *XI*(1), 9–23.

Duran, H. E., & Ozkan, S. P. (2015). Trade openness, urban concentration and city-size growth in Turkey. *Regional Science Inquiry*, *7*(1), 35–46.

Duranton, G. (2007). Urban evolutions: The fast, the slow, and the still. *American Economic Review*, *97*(1), 197–221. https://doi.org/10.1257/aer.97.1.197

Eaton, J., & Eckstein, Z. (1997). Cities and growth: Theory and evidence from France and Japan. *Regional Science and Urban Economics*, *27*(4–5), 443–474. https://doi.org/10.1016/S0166-0462(97)80005-1

Eeckhout, J. (2004). Gibrat's law for (all) cities. *American Economic Review*, *94*(5), 1429–1451. https://doi.org/10.1257/0002828043052303

Elburz, Z., Nijkamp, P., & Pels, E. (2017). Public infrastructure and regional growth: Lessons from meta-analysis. *Journal of Transport Geography*, *58*, 1–8. https://doi.org/10.1016/j.jtrangeo.2016.10.013

Evcil, A. N., Dokmeci, V., & Kıroglu, G. B. (2006) Regional migration in Turkey: Its directions and determinants, 46th European congress of the regional science association, august 30–September 3, Volos, Greece.

Fan, J., & Gijbels, I. (1996). *Local polynomial modelling and its applications*. London: Chapman & Hall.

Fan, J., Heckman, N. E., & Wand, M. P. (1995). Local polynomial regression for generalized linear models and quasi likelihood functions. *Journal of the American Statistical Association*, *90*(429), 141–150. https://doi.org/10.1080/01621459.1995.10476496

Fan, J., & Marron, J. S. (1994). Fast implementations of nonparametric curve estimators. *Journal of Computational and Graphical Statistics*, *3*(1), 35–56.

Fazio, G., & Modica, M. (2015). Pareto or log-normal? Best fit and truncation in the distribution of all cities. *Journal of Regional Science*, *55*(5), 736–756.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics*, *114*(3), 739–767. https://doi.org/10.1162/003355399556133

Gabaix, X., & Ibragimov, R. (2007). Rank-1/2: A simple way to improve the OLS estimation of tail exponents, NBER technical working paper no. 342, National Bureau of economic analysis, New York.

Gabaix, X., & Ioannides, Y. M. (2004). The evolution of city size distributions in press. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of regional and urban economics 4* (pp. 2341–2378). Netherlands: Elsevier.

Gan, L., Li, D., & Song, S. (2006). Is the Zipf law spurious in explaining city-size distributions? *Economic Letters*, *92*(2), 256–262. https://doi.org/10.1016/j.econlet.2006.03.004

Gedik, A. (1997). Internal migration in Turkey, 1965-1985: Test of conflicting findings in the literature. *Review of Urban and Regional Development Studies*, *9*(2), 170–179. https://doi.org/10.1111/j.1467-940X.1997.tb00104.x

Gedik, A. (2003). Differential urbanisation in Turkey, 1955–97. *Tijdschrift voor Economische en Sociale Geografie*, *94*(1), 100–111. https://doi.org/10.1111/1467-9663.00240

Gibrat, R. (1931). *Les inégalités économiques*. Paris: Libraire du Recueil Siray.

Giesen, K., & Suedekum, J. (2011). Zipf's law for cities in the regions and the country. *Journal of Economic Geography*, *11*(4), 667–686.

Giesen, K., Zimmermann, A., & Suedekum, J. (2010). The size distribution across all cities double Pareto lognormal strikes. *Journal of Urban Economics*, *68*(2), 129–137. https://doi.org/10.1016/j.jue.2010.03.007

Gourieroux, C., Holly, A., & Monfort, A. (1982). Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters. *Econometrica*, *50*(1), 63–80. https://doi.org/10.2307/1912529

Guerin-Pace, F. (1995). Rank-size distribution and the process of urban growth. *Urban Studies*, *32*(3), 551–562. https://doi.org/10.1080/00420989550012960

Härdle, W. (1991). *Smoothing techniques with implementation in S*. New York: Springer Verlag.

Henderson, D. J., & Parmeter, C. F. (2015). *Applied nonparametric econometrics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511845765

Hettne, B., & Söderbaum, F. (1998). The new regionalism approach. *Politeia*, *17*(3), 6–21.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, *3*(5), 1163–1174.

Hinkley, D. V. (1977). Jackknifing in unbalanced situations. *Technometrics*, *19*(3), 285–292. https://doi.org/10.1080/00401706.1977.10489550

Hsu, W. T. (2012). Central place theory and city size distribution. *Economic Journal*, *122*(563), 903–932. https://doi.org/10.1111/j.1468-0297.2012.02518.x

Ioannides, Y., & Skouras, S. (2013). US city size distribution: Robustly Pareto, but only in the tail. *Journal of Urban Economics*, *73*(1), 18–29. https://doi.org/10.1016/j.jue.2012.06.005

Ioannides, Y. M., & Overman, H. G. (2003). Zipf's law for cities: An empirical examination. *Regional Science and Urban Economics*, *33*(2), 127–137. https://doi.org/10.1016/S0166-0462(02)00006-6

Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano Degli Attuari*, *4*(1), 83–91.

Kundak, S., & Dokmeci, V. (2018). A rank-size rule analysis of the city system at the country and province level in Turkey. *International Journal of Architecture and Planning*, *6*(1), 77–98.

Lalanne, A. (2014). Zipf's law and Canadian urban growth. *Urban Studies*, *51*(8), 1725–1740. https://doi.org/10.1177/0042098013498623

Le Sage, P. J. (2008). An introduction to spatial econometrics. *Revue d'Économie Industrielle*, *123*(3), 19–44.

Lee, S., & Li, Q. (2013). Uneven landscapes and city size distributions. *Journal of Urban Economics*, *78*, 19–29. https://doi.org/10.1016/j.jue.2013.05.001

Li, S., & Sui, D. (2013). Pareto's law and sample size: A case study of China's urban system, 1984–2008. *GeoJournal*, *78*, 615–626. https://doi.org/10.1007/s10708-012-9455-9

Lu, M., & Wan, G. (2014). Urbanization and urban systems in the People's Republic of China: Research findings and policy recommendations. *Journal of Economic Surveys*, *28*(4), 671–685. https://doi.org/10.1111/joes.12078

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*, 3–42. https://doi.org/10.1016/0304-3932(88)90168-7

Luckstead, J. S., & Devadoss, S. (2014). A comparison of city size distributions for China and India from 1950 to 2010. *Economic Letters*, *124*(2), 290–295. https://doi.org/10.1016/j.econlet.2014.06.002

Mansury, Y., & Gulyás, 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*, 2438–2460. https://doi.org/10.1016/j.jedc.2006.08.002

Modica, M., Reggiani, A., & Nijkamp, P. (2015). A Comparative Analysis of Gibrat's and Zipf's Law on Urban Population, Quaderni - Working Paper DSE N° 1008. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2611960

Moura, N. J., & Ribeiro, M. B. (2006). Zipf law for Brazilian cities. *Physica a: Statistical Mechanics and its Applications*, *367*, 441–448. https://doi.org/10.1016/j.physa.2005.11.038

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and its Applications*, *9*(1), 141–142. https://doi.org/10.1137/1109020

Nitsch, V. (2005). Zipf zipped. *Journal of Urban Economics*, *57*(1), 86–100. https://doi.org/10.1016/j.jue.2004.09.002

Nota, F., & Song, S. (2012). Further analysis of the Zipf's law: Does the rank-size rule really exist? *Journal of Urban Management*, *1*(2), 19–31. https://doi.org/10.1016/S2226-5856(18)30058-X

Pebesma, E., Bivand, R., Rowlingson, B., Gomez-Rubio, V., Hijmans, R., Sumner, M., MacQueen, D., Lemon, L., Lindgren, F., O'Brien, J., & O'Rourke, J. (2021). SP Package R. https://cran.r-project.org/web/packages/sp/sp.pdf

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. https://doi.org/10.1016/0094-1190(80)90043-1

Rossi-Hansberg, E., & Wright, M. L. J. (2007). Urban structure and growth. *The Review of Economic Studies*, *74*(2), 597–624. https://doi.org/10.1111/j.1467-937X.2007.00432.x

Semboloni, F. (2001). Agents with dycotomic goals which generate a rank-size distribution, CASA working paper no. 33, Centre for Advanced Spatial Analysis, London
Simonoff, J. S. (1996). *Smoothing methods in statistics*. New York: Springer-Verlag.
Singer, H. W. (1936). The "Courbe des populations". A parallel to Pareto's law. *The Economic Journal*, *46*(182), 254–263. https://doi.org/10.2307/2225228
Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, *19*(2), 279–281. https://doi.org/10.1214/aoms/1177730256
Soo, K. T. (2005). Zipf's law for cities: A cross-country investigation. *Regional Science and Urban Economics*, *35*(3), 239–263. https://doi.org/10.1016/j.regsciurbeco.2004.04.004
Soo, K. T. (2007). Zipf's law and urban growth in Malaysia. *Urban Studies*, *44*(1), 1–14. https://doi.org/10.1080/00420980601023869
Veneri, P. (2016). City size distribution across the OECD: Does the definition of cities matter? *Computers, Environment and Urban Systems*, *59*, 86–94. https://doi.org/10.1016/j.compenvurbsys.2016.05.007
Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series a*, *26*(4), 359–372.
Wheaton, W. C., & Shishido, H. (1981). Urban concentration, agglomeration economies, and the level of economic development. *Economic Development and Cultural Change*, *30*(1), 17–30. https://doi.org/10.1086/452537
White, H. (1980). A Heteroskedasticity-consistent covariance matrix estimator and a direct test for Heteroskedasticity. *Econometrica*, *48*(4), 817–838. https://doi.org/10.2307/1912934
Wilmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error MAE over the root mean square error RMSE in assessing average model performance. *Climate Research*, *30*(1), 79–82. https://doi.org/10.3354/cr030079
Zipf, G. K. (1949). *Human behaviour and the principle of least effort: An introduction to human ecology*. Reading, MA: Adison-Wesley Press.

How to cite this article: Duran, H. E., & Cieślik, A. (2021). The distribution of city sizes in Turkey: A failure of Zipf's law due to concavity. *Regional Science Policy & Practice*, *13*(5), 1702–1719. https://doi.org/10.1111/rsp3.12449

ELECTRONIC SOURCES

www.tuik.gov.tr (Turkish Statistical Institute)

https://www.harita.gov.tr/urun/turkiye-mulki-idare-sinirlari/232

APPENDIX A.

Zipf law modeling by using rank minus half as the dependent variable: ln($R - 1/2$)

TABLE 1A Threshold regressions; source: own estimation

Dependent variable:	**Total population**		**Urban population**	
Parameters	Above threshold	Below threshold	Above threshold	Below threshold
c	29.59276***	11.02864***	28.58516***	9.909069***
$-(\beta)$	−2.01925***	−0.47973***	−1.94328***	−0.38364***
White–Hinkley heteroskedasticity Consistent covariance and S.E.	Yes	Yes	Yes	Yes
R^2	0.89	0.95	0.88	0.94
N	146	827	146	827

Note:
*Statistical significance at 10%,
**at 5%,
***at 1%, valid also in the remaining parts of the paper.

TABLE 1b Quadratic regressions; source: own estimation

Parameters	Total population	Urban population
α	−4.92437***	−1.55274**
β	2.741884***	2.068436***
γ	−0.16088***	−0.12876***
Wald test (beta = −1) (χ^2 test stat)	–	–
White-Hinkley heteroskedasticity Consistent covariance and S.E.	Yes	Yes
R^2	0.97	0.96
N	973	973

Note:
*Statistical significance at 10%,
**at 5%,
***at 1%, valid also in the remaining parts of the paper.

TABLE 1c Spatial autocorrelation tests, source: Own estimation

	Linear		Quadratic	
Dependent variable	Total population	Total population	Total population	Urban population
LMerr	458.86***	458.86***	1.5484	10.729***
Lmlag	95.615***	95.615***	2.8945*	6.4013**
RLMerr	402.08***	402.08***	1.27	9.53***
RLMlag	38.838***	38.838***	2.6161	5.2024**

TABLE 1d Spatial autoregressive and spatial error models; source: own estimation

Linear Parameters	Total population-SAR	Urban population-SAR	Total population-SEM	Urban population-SEM
α	9.919447***	8,014942***	13.297***	11.93605***
$-(\beta)$	−0.64662***	−0.5236***	−0.66753***	−0.54623***
Gamma				
Rho	0.46***	0.54***		
Lambda			0.97***	0.98***
N	973	973	973	973
Quadratic Parameters	**Total population-SAR**	**Urban population-SAR**	**Total population-SEM**	**Urban population-SEM**
α	−5.10195***	−1.89609***	−4.84322***	−1.39536***
$-(\beta)$	2.717823***	2.04015***	2.725677***	2.036609***
Gamma	−0.15956***	−0.12709***	−0.16008***	−0.1271***
Rho	0.05*	0.08**		
Lambda			0.44	0.77***
N	973	973	973	973

TABLE 1e MAE, MAPE, and Kolmogorov–Smirnov tests: deviations from Pareto optimal distribution in terms of rank, mean absolute error and Kolmogorov–Smirnov test; source: own estimation

SAR	Linear/ quadratic	Dependent variable	Total/urban populations	MAE	MAPE	Kolmogorov–Smirnov (*D*-stat)	*p*-Value
	Linear	Ln($R - 1/2$)	Total	0.227983	0.043286	0.14286***	4.76×10^{-9}
	Linear	Ln($R - 1/2$)	Urban	0.257929	0.048509	0.14902***	8.25×10^{-10}
	Quadratic	Ln($R - 1/2$)	Total	0.098976	0.022058	0.10894***	1.93×10^{-5}
	Quadratic	Ln($R - 1/2$)	Urban	0.112493	0.0243	0.10997***	1.55×10^{-5}
SEM	Linear	Ln($R - 1/2$)	Total	0.224618	0.04263	0.13361***	5.73×10^{-8}
	Linear	Ln($R - 1/2$)	Urban	0.25331	0.04758	0.14286***	4.76×10^{-9}
	Quadratic	Ln($R - 1/2$)	Total	0.098969	0.022059	0.111***	1.24×10^{-5}
	Quadratic	Ln($R - 1/2$)	Urban	0.111361	0.024124	0.11305***	7.95×10^{-6}