Long-run Zipf's and Gibrat's Laws Deviation: Spanish Population Impacts*

Guillermo Peña[§a]

[a]Departamento de Análisis Económico & IEDIS, Universidad de Zaragoza, Spain

gpena@unizar.es

December 8, 2025

Abstract

This paper analyzes the population effects of the deviation from the Zipf and Gibrat laws for all the municipalities of Spain in the last censuses from 1900 to 2020 in NUTS-3 level (provinces). A novel econometric technique of Difference-in-Differences (DiD) with multiple time period treatments is employed. The noncompliance of the Zipf law robustly leads to higher population while the rejection of the Gibrat law leads to lower inhabitants only when climate factors are considered. An example of agglomeration shadows is found at the province of the capital of Spain, Madrid, and its surrounding provinces.

Keywords: Zipf, Gibrat, NUTS-3, shock persistence, difference-in-differences, heterogeneity.

JEL classification: R10, R11, J11, R12.

*This work has received financial support from Fundación Ibercaja/Universidad de Zaragoza (JIUZ2022-CSJ-19), Gobierno de Aragón (S39-23R ADETRE Research Group), and Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación (Grant PID2024-157255NB-I00).

§Corresponding author. Address: Facultad de Economía y Empresa. Gran Vía 2. 50005 Zaragoza, Spain. Tel.: (+34) 976 762 212.

1 Introduction

The analysis of power laws and, specifically, the study of the potential validity of the Zipf's and Gibrat's Laws, holds a long tradition as scientific field inside of the Urban Economics (Clauset, Shalizi, and Newman 2009; Fang, Pang, and Liu 2017; González-Val, Ximénez-de-Embún, and Sanz-Gracia 2024). Zipf proposes a linear stable relationship between the rank and the population of cities, called as Zipf's law. On the other hand, Gibrat postulates a law of proportional growth, or Gibrat's law, where the population growth rate of cities is independent from their initial population. For a extensive review of literature, see Arshad, Hu, and Ashraf (2018) and Nitsch (2005). The present paper follows González-Val, Ximénez-de-Embún, and Sanz-Gracia (2024) study but with the following differences. First, the present paper deals with Spain and its NUTS-3 administration level as geographical reference, the provinces, with few literature on the topic as indicated. Second, in addition to applying the methodology from González-Val, Ximénez-de-Embún, and Sanz-Gracia (2024), this paper goes beyond them and other papers and this is the first time, to the best of our knowledge, that a paper analyzes, by modern advanced Difference-in-Differences techniques for multiple time periods, the impact of the fulfillment of Zipf's and Gibrat's laws on the population size of the provinces. Finally, it is remarkable highlighting the wider sample used. Whereas other authors as the previous ones stop the analysis in 2010, this article goes beyond until data of 2020.

At the sub-national level there is scarce literature. Some examples include studies for the German case (Giesen and Súdekum 2010), an Indian province (Ganapathy and Arumugam 2014; Subbarayan 2009), Colombia (Pérez Valbuena and Meisel Roca 2014), eight Chinese subregions (Ye 2011), 26 provinces from China (Ziqin 2016), analyzed sub-national Chinese administrative areas (Ye 2011), Turkish provinces (Kundak and Dókmeci 2018), Pakistan (Arshad, Hu, and Ashraf 2019) and the United States (González-Val, Ximénez-de-Embún, and Sanz-Gracia 2024). For the Spanish case there is a lack of papers on the sub-national level, an example includes with an analysis with regions (González-Val 2023). This author does not apply Zipf for provinces but urban growth is analyzed.

There is empirical evidence of a constant factor of the inverse proportional relationship between the rank of a city (municipality in Spain) and its population (Auerbach and Ciccone 2023). There is also an enriched formulation Zipf based on the statistical regularity of slope equal to -1 when regressing the logarithm of the rank of the city with respect to the logarithm of its population (Wan et al. 2020; Wang 2021; Wei et al. 2021). The formulation known as Zipf's law (Soo 2005) is "one of the most conspicuous empirical facts in economics" (Ciaschini et al. 2025; Gabaix 1999b). This law is intended to be an effective tool for the analysis among different metropolitan hierarchies (Berezkin et al. 2020; Ciaschini et al. 2025; Nitsch 2005; Peña and Sanz-Gracia 2021; Pérez-Campuzano, Guzmán-Vargas, and Angulo-Brown 2015). Zipf's law mostly holds for cities above a certain population size (Bergs 2018; Gabaix 1999a; Gabaix 1999b). Other laws include rural-urban elements as the Taylor's power law (Leroy R. Taylor 1961), applied to animal and human populations (L. R. Taylor, Woiwod, and Perry 1978).

Eeckhout (2004) proposed that the Pareto exponent in a standard Zipf equation where the Pareto coefficient decreases as more cities are added to the sample. This paper will additionally try to solve this issue for the last available census of Spain, by means of a rolling sample regressions approach. Similar exercise was conducted by (González-Val and Sanz-Gracia 2024) but United States (US; 2000 and 2010 censuses), Spain (2001 and 2011 censuses) and Italy (2001 and 2011 censuses). Nonetheless, the present paper enlarges and updates de sample, comparing for 1900 and 2020. They find that Eeckhout's proposition is only valid once the distribution of data starts to be lognormal; for the Pareto-distributed upper-tail, the estimated exponent does not change with sample size. As Peña and Sanz-Gracia (2021), the former authors estimate by employing rolling sample regressions (Peng 2010). As the present results will show, the upper-tail will show a raise when the sample enlarges. This paper largely focuses on studying the validity of Zipf's and Gibrat's laws considering each one of provinces (NUTS 3 level) from 1900 to 2010, since testing both laws requires long time (Gabaix and Ioannides 2004). Nonetheless, the present paper updates the sample by other authors (González-Val and Sanz-Gracia 2024; González-Val, Ximénez-de-Embún, and Sanz-Gracia 2024) by focusing on Spain and also considering the 2020 year.

Finally, the main novelty of this paper is obtaining the impact of the municipal Zipf's and Gibrat's non-compliance on total provincial population by an advanced and recently proposed Difference-in-Differences econometric technique for multiple time periods developed by Callaway and Sant'Anna (2021). The event is considered when there is a first rejection of these laws. As the result sections will show, it is found an increase/reduction after the (non)compliance of Zipf's/Gibrat's laws. Robustness checks are provided.

The present paper is divided as follows. After this brief introduction, next Section 2 provides the data and variables and the different methodologies applied in this paper, while Section 3-4 provide the results and discusses them. Finally, Section 5 concludes.

2 Data and Methods

2.1 Data

The data employed is found in the Spanish National Statistics (INE from the acronym in Spanish) database. It collects the national census elaborated roughly each ten years from 1900 to 2020, for more than 8000 municipalities (the whole universe of cities at Spain in 2020).

The number of cities (municipalities) for each one of the 50 provinces in 2020 (the date used for homogenizing the sample) totalizes more than 8000 municipalities. Other years have similar number of municipalities, with the exception of some administrative changes along time, preserved in most cases. The similarity of all the samples of municipalities provides comprehensive information about the scarce birth of new cities even focusing in the long-run.

2.2 Methods for Zipf and Gibrat

The Section 3 provides the methodology of the paper. Concretely, for checking the Zipf's law the next Pareto formulation is applied taking into account the correction addressed by Gabaix and Ibragimov (2011):

$$ln(R - 0.5) = \alpha - \beta ln(S), \tag{1}$$

where R is the rankof the city (being 1 assigned to the most populated one, 2 for the second city and so on), S is its population size, α is a parameter constant and β is the so called Pareto coefficient that, by construction, is always positive. When $\beta = 1$, then, we say that the Zipf's law is accepted and the β is called the Zipf exponent. For checking whether the Zipf's law is statistically verified, it is employed the asymptotically correct standard error of the estimated Pareto cefficient ($\hat{\beta}$). So, it is checked whether β is statistically equal to one, by calculating $\hat{\beta}(2/N)^{1/2}$, where N is the total number of cities included in the sample.

For the check of the rolling sample regression (Peng 2010), cities are incorporated one by one to the sample. It begins with the largest twentieth city, and each time we calculate the estimated Pareto exponent. The first sample size for each country is twenty and it increases one by one until reaching the smallest urban unit. Thus, the largest rank R considered in each regression is equal to the one by one varying sample size N.

The Gibrat Law, on the other hand, holds that the growth of the population is independent from the initial size. Following Ahundjanov and Akhundjanov (2019) expression, which constitutes the third approach proposed by Eeckhout (2004), it has to be checked the significance of the $\hat{\beta}$ (rejection implies deviation from the Gibrat law) or not (non-rejection means verification):

$$ln(S_{i,t}/S_{i,t-1}) = \alpha + \beta ln(S_{i,t-1}) + u_i, \tag{2}$$

where $S_{i,t}$ is the population in city i, for $i = 1, ..., n$, at time t, for $i = 1, ...,$ and u_i is a random error term. There are other ways for checking the Gibrat's law, as by statistical independency tests (Peña, Puente-Ajovín, and Sanz-Gracia 2025) or panel unit root tests (Clark and Stabler 1991; González-Val, Lanaspa, and Sanz-Gracia 2014).

2.3 Methodology for the Causal Inference

The Difference-in-Difference methodology with multiple time treatments developed by Callaway and Sant'Anna (2021) is employed. The dynamic "Average Treatment Treatment" effects (ATTs) with multiple time periods are analyzed. They reflect the impacts on the provincial (NUTS 3 administration level) population of the non-compliance of the Zipf's and Gibrat's laws. The next formula by Doan et al. (2024) is followed in the present paper (Peña 2025):

$$ATT(g,t) = \mathbb{E}\left[\frac{G_g}{\mathbb{E}[G_g]} - \frac{\frac{p_g(X)C}{1-p_g(X)}}{\mathbb{E}\left[\frac{p_g(X)C}{1-p_g(X)}\right]}(Y_t - Y_{g-1})\right]. \tag{3}$$

This method, applied to assess the impacts of the event shock ('treatment', the rejection of Zipf's and Gibrat's laws) on the total of inhabitants, consists of comparing the evolution of the population in the affected ('treated') group with that in a non-affected ('control') one (Sanso-Navarro and Vera-Cabello 2023). In contrast, the not-yet (NY) estimations compare the treated group with the not-yet affected ('not yet treated') group composed by those countries where there is an affection by the event but just considering the years before that event. For the purpose of reaching an improvement in the estimation and inference process, Callaway and Sant'Anna (2021) develop semi-parametric DiD estimators in set-ups with multiple time periods and allow a variation in the timing of treatment, considering $t = 0$ to the period when the event occurs for all the different events and dates.

The method that Callaway and Sant'Anna (2021) proposed, as Doan et al. (2024) affirms, introduces the notion of $ATTs(g,t)$. The different groups are defined by the time treatment initiation. The specific time period is expressed by t. The total of temporal periods is T, denoted as $t = 1, 2, , T$. The Gg denotes a dummy variable, which takes value 1 when a unit is initially treated in period g and zero otherwise. Another dummy variable is C. It takes the value 1 for never treated units and zero otherwise. Each individual is included either within the treatment group (g) or within the never treated group ($C = 1$). A generalized propensity score is defined as $p_g(X) = P(G_g = 1|X, G_g + C = 1)$, which constitutes the probability that a unit receives treatment, given the covariates X and whether it is included within a group (g) of the control group. It can also be considered that $p_g(X)$ is the likelihood of a particular unit with covariates X of receiving treatment. The covariates are X. This score provides the probability of a unit to be treated. The previous expression provides the parameter of interest.

Regarding the econometric strategy, given the concerns with the assumption of parallel trends that underlies the used estimation approach, the target is finding a model that minimizes significant, pre-treatment differences in the total provincial number of inhabitants. Each figure of the results shows, for the following empirical estimations, 0 as the treatment period, and -x and x are, respectively, the impacts on the previous (pre-treatment) and posterior (post-treatment) x periods, where x is the maximum available periods for each estimation. The number of periods is homogenized to $x = 10 or 11$ obtaining the results, with the exceptions of the cases when the estimations do not allow that. Each result shows the estimated dynamic ATTs when the treatment group includes those countries that passed from a non-rejection to a rejection of Zipf's or Gibrat's laws in the sample period analyzed, considering that year as $t = 0$, and the control group includes the other countries. As regressor variables, lags (Time) and geographical (Geo) variables have been used. When there are no regressor variables apart from the interest ones, it is called "Unconditional". The rest of models represent conditional estimations, in those cases, according to Callaway and Sant'Anna (2021), the endogenity issues can be solved following the parallel trend hypothesis when the pre-event effects are not significant and statistically significant the post-event impacts. Shaded areas correspond to 95% confidence interval bounds.

3 Zipf and Gibrat Results and Variables for the Inference

Table 1 shows the sample of variables for the 11 censal periods for the 50 provinces (a total of 650 observations per variable). The *Population* and *lnPopulation* variables show the, respectively, censal population per province and census year and its neperian logarithm, used as Time control variable for the conditional Difference-in-Differences (DiD) estimations. With respect to the Geographical control variables, Geo, they appear *nuts1* as the classification of NUTS1 for supra-regions inside of a country according to Eurostat. The *Income_ Quartiles* variable provides the classification of the provinces according to the income in 2020, for distinguishing among subsamples. The *Zipf_ treatment* provides the first year when the Zipf's law is statistically rejected, while the *Zipf_ treatment* provides the same information but for Gibrat. Finally, the *Climate_ code* permits an additional classification of the provinces regarding the climate according to the Köppen climate codes[1].

[Insert Table 1]

The Pareto exponents for each autonomous community per censal year are collected in Table 2, highlighting the paulatine reduction on the Pareto coefficient along the years. It is similar for when the Gibrat law is analyzed for the provinces, with the p-value collected in Table 3. It is observed a smoothy reduction on the p-value, leading to higher rejections of the Gibrat law in the last periods.

[Insert Tables 2-3]

Figures 1-2 show graphically (Figure 1) and with maps (Figure 2) the reduction on the verification of both laws along time. Figure 1 above on the left reflects the initial higher verification of the Gibrat law for the first censal year of the sample, 1910, while it is non-rejected after that. As explained above with Pareto, there is no fulfillment of the Zipf law for the full sample, for entire Spain, where the graph above on the right is representative of that, every year farther from the Zipf law. With the rolling sample regression, it can be seen for 1900 and 2020 that the Eeckhout law sustaining the reduction of the Pareto exponent as well as the sample widens is neither verified, passing from acceptance above the 1 to rejection below it passing by a shape of "bell".

[Insert Figure 1]

Figure 2 also reflects the farther deviation from the Zipf law every year along the time span of the sample. Only a few provinces, mostly in the Center-East of the peninsula, reject it in 1910, until 2020 when almost half of Spanish provinces rejects the law. The verification of the Zipf law was in a wide majority of provinces in 1900 and in 2020 only a few provinces of the Northweast and South of Spain fulfill it. The figure also shows the ES4 region (Center of Spain without Madrid) with this region being the closest one to

[1] https://climateknowledgeportal.worldbank.org/country/spain

Madrid that has experienced an agglomeration shadow, leading to less population. This fact of potential correlation between rejection/compliance of Zipf/Gibrat laws and higher population growth will be confirmed in the empirical inferencial exercise.

[Insert Figure 2]

The previous found data obtained as results, which are useful by themselves, constitute the basis for constructing the target treatment variables from Table 1 and will be also useful for estimating their impacts on population in the next Section 4.

4 Inference Results

This section reports the principal results following the methodology previously collected in Section 2. Table 4 shows the ATTs for the Zipf treatment variable, so the event is the first non-fulfillment of the Zipf law, with the *Population* variable as the dependent one. The first model is unconditional, i.e., without any control variable, and conditional to temmporal lags (Time), geographical factors (Geo) and both. The results show that the "parallel trend hypothesis" that ensures the avoidance of endogeneity for conditional models as Callaway and Sant'Anna (2021) state is verified since before the event the average ATTs are not statistically significant but statistically significant after the event. Concretely, the impact is positive and significant, indicating a raise on around 450,000 to 850,000 more inhabitants after the event in the long-run. Figure 3 shows similar results than Table 6 by estimating with the same variables but for dealing with Not-Yet (NY) treated provinces as control group, i.e., instead NY and never treated countries as control group, only the first ones are included. The pre-event effects appear in blue, before the initial date $T = 0$, whilst the post-treatment impacts in red, after the initial censal year when the event occurs. The bounds appear indicating the region of the 95% confidence interval.

[Insert Table 4 & Figure 3]

Table 5 and Figure 4 show the only analyzed cases where the Gibrat treatment fulfills the parallel trend hypothesis: when climate is involved. Table 5 shows the estimations for population as dependent variable with controls of Time and Geo for three models with

Zipf law deviation as treatment (the three first ones) and another (the last one) with the Gibrat law deviation used as treatment. The difference with other models is the division by subsamples regarding the climate conditions of the provinces. The results confirm the raise on the population after rejecting Zipf law, in these cases around one million of new inhabitants in warm regions (*bsk*, *csb* , *csa*), but the previous hypothesis is also fulfilled for the deviation from the Gibrat law when warm-climate NUTS3 regions are analyzed. Nonetheless, as expected in the previous Section 3, the effect is the opposite than with the Zipf law: there is a diminution on the population of around 360,000 inhabitants. Figure 4 focuses on the non-compliance of the Gibrat law, first (above) for unconditional models, with never treated (left) and only NY provinces (right) as control groups, showing a reduction on the population. Nonetheless, unconditional models cannot ensure the avoidance on endogeneity issues according the parallel trend hypothesis, and using the climate-code as control variables leads to statistically non-significant effects before the rejection of the Gibrat law and statistically relevant impacts after the event, confirming the hypothesis. The results clearly collect a decrease on population in the long-run after the deviation.

[Insert Table 5 & Figure 4]

Table 6 divides the sample into NUTS1 regions, intead of by climatic regions as performed in the previous Table 5 and Figure 4. The results are focused, from left to right, on the Center of Spain (ES4), ES4 with NY, ES4 plus Madrid uniprovincial autonomus community (ES3) and ES3 and ES4 with NY. These results show, for geographical subsamples, the robustness of the increase on population after the deviation from the Zipf law. Table 6 and Figure 5 also confirm the previous results but for each one of the quartiles of GDP per capita in which the provinces have been divided for, respectively, never treated and NY treated samples as control groups. It is remarkable the higher amount of average ATTs, i.e., on the raise of inhabitants, in the upper-middle and higher income subsamples, passing from around 200,000-210,000 more inhabitants to roughly 1,100,000-1,200,000 inhabitants. Theorefore, in this last case, Madrid would have again a huge weight as seen in Table 5 and Figure 5 and in the previous Section 3.

[Insert Tables 6-7 & Figure 5]

The previous results show the trade off in the fulfillment of Zipf and Gibrat laws at the same time: verifying Zipf law would lead to lower population than rejecting it, while compliance with Gibrat law leads to the opposite effect. It is remarkable how regions as the Center of Spain without Madrid province non-reject/reject the Zipf/Gibrat laws, leading to lower population, in contrast to the closer Madrid, in the center of the country, was the opposite case, with a significant and robust increase of the population after non-complying the Zipf law and verifying the Gibrat's one.

5 Concluding remarks

Testing the validity of Zipf's and Gibrat's laws requires long time intervals (Gabaix and Ioannides 2004; González-Val, Ximénez-de-Embún, and Sanz-Gracia 2024).The exploratory results show a reduction on the compliance of both laws. The present paper perfoms this testing for all the municipalities of Spain in the last censuses from 1900 to 2020, for the entire sample and in a provincial level for the 50 NUTS-3 Spanish regions (provinces). After that, these estimated finding data are employed as data-sample for the analysis of the causality of the deviation from those laws on the total provincial population. For analyzing this, the novel econometric technique of Difference-in-Differences (DiD) with multiple time period treatments developed by Callaway and Sant'Anna (2021) is employed.

Both the exploratory and inference results are conclusive: the un-fulfillment of the Zipf law leads to higher population in the long run while the rejection of the Gibrat law is accompanied by a reduction on the number of inhabitants. So, the verification of both laws leads to opposite results. Examples of these cases are the capital of Spain, Madrid, allocated in the Center of Spain, and its circunvalant provinces: while the province of the capital rejects the Zipf law and non-rejects the Gibrat law and its population rises due to this, on the other hand, its surrounding provinces comply more with the Zipf law and deviate from the Gibrat law, with this fact leading this NUTS-1 region (ES4) of the Center of Spain to lower population.

As the exploratory results show, there is a potential correlation between income, population and warm climate data, so the previous results on population could be extended to income and climate and similar results would be obtained due to the agglomeration economies. Nonetheless, the scarce economic and climate data availability for this long period of time (1900-2020) constitutes a limitation of the paper, but at the same time an opportunity for further research in the future, when longer-time economic and climatic data will be available.

References

Ahundjanov, Behzod B. and Sherzod B. Akhundjanov (2019). "Gibrat's law for CO2 emissions". *Physica A: Statistical Mechanics and its Applications* 526 (2019), p. 120944. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2019.04.180. URL: https://www.sciencedirect.com/science/article/pii/S0378437119305904.

Arshad, Sidra, Shougeng Hu, and Badar Nadeem Ashraf (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 (2018), pp. 75–92. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2017.10.005. URL: https://www.sciencedirect.com/science/article/pii/S0378437117310130.

— (2019). "Zipf's law, the coherence of the urban system and city size distribution: Evidence from Pakistan". *Physica A: Statistical Mechanics and its Applications* 513 (2019), pp. 87–103. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2018.08.065. URL: https://www.sciencedirect.com/science/article/pii/S0378437118310069.

Auerbach, Felix and Antonio Ciccone (2023). "The Law of Population Concentration". *Environment and Planning B: Urban Analytics and City Science* 50.2 (2023), pp. 290–298. DOI: 10.1177/23998083221147139. eprint: https://doi.org/10.1177/23998083221147139. URL: https://doi.org/10.1177/23998083221147139.

Berezkin, D., A. Goryachev, D. Gritsenko, A. Lychkina, and A. Zolotarev (2020). "Zipf's Law for Russian Cities: Analysis of New Indicators". *Economy of Regions* 16.3 (2020), pp. 935–947. DOI: 10.17059/ekon.reg.2020-3-20.

Bergs, Rolf (2018). "Segmentation of Urban and Non-Urban Space in the Netherlands – Testing Zipf's Law with Nocturnal Satellite Imagery". *Review of Regional Research* 38.2 (2018), pp. 111–140.

Callaway, Brantly and Pedro H.C. Sant'Anna (2021). "Difference-in-differences with multiple time periods". *Journal of Econometrics* 225.2 (2021), pp. 200–230. DOI: 10.1016/j.jeconom.2020.12.001.

Ciaschini, Clio, Rosanna Salvia, Margherita Carlucci, and Luca Salvati (2025). "Testing the Zipf's law under heterogeneous urban-rural hierarchies with spatial quantile regressions". *Applied Economics* 57.21 (2025), pp. 2665–2678. DOI: 10.1080/00036846.2024.2331027. eprint: https://doi.org/10.1080/00036846.2024.2331027. URL: https://doi.org/10.1080/00036846.2024.2331027.

Clark, J. Stephen and Jack C. Stabler (1991). "Gibrat's Law and the Growth of Canadian Cities". *Urban Studies* 28.4 (1991), pp. 635–639. DOI: 10.1080/00420989120080701. eprint: https://doi.org/10.1080/00420989120080701. URL: https://doi.org/10.1080/00420989120080701.

Clauset, Aaron, Cosma Rohilla Shalizi, and M. E. J. Newman (2009). "Power-Law Distributions in Empirical Data". *SIAM Review* 51.4 (2009), pp. 661–703. DOI: 10.1137/070710111. eprint: https://doi.org/10.1137/070710111. URL: https://doi.org/10.1137/070710111.

Doan, Nguyen, Huong Doan, Canh Phuc Nguyen, and Binh Quang Nguyen (2024). "From Kyoto to Paris and beyond: A deep dive into the green shift". *Renewable Energy* 228 (2024), p. 120675. ISSN: 0960-1481. DOI: https://doi.org/10.1016/j.renene.2024.120675. URL: https://www.sciencedirect.com/science/article/pii/S0960148124007432.

Eeckhout, Jan (2004). "Gibrat's Law for (All) Cities". *American Economic Review* 94.5 (Dec. 2004), pp. 1429–1451. DOI: 10.1257/0002828043052303. URL: https://www.aeaweb.org/articles?id=10.1257/0002828043052303.

Fang, Chuanglin, Bo Pang, and Haimeng Liu (2017). "Global city size hierarchy: Spatial patterns, regional features, and implications for China". *Habitat International* 66 (2017), pp. 149–162. ISSN: 0197-3975. DOI: https://doi.org/10.1016/j.habitatint.2017.06.002. URL: https://www.sciencedirect.com/science/article/pii/S0197397517301091.

Gabaix, Xavier (1999a). "Zipf's Law and the Growth of Cities". *The American Economic Review* 89.2 (1999), pp. 129–132. ISSN: 00028282. URL: http://www.jstor.org/stable/117093.

— (1999b). "Zipf's Law for Cities: An Explanation". *The Quarterly Journal of Economics* 114.3 (Aug. 1999), pp. 739–767. ISSN: 0033-5533. DOI: 10.1162/003355399556133. eprint: https://academic.oup.com/qje/article-pdf/114/3/739/5228563/114-3-739.pdf. URL: https://doi.org/10.1162/003355399556133.

Gabaix, Xavier and Rustam Ibragimov (2011). "Rank − 1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents". *Journal of Business Economic Statistics* 29.1 (2011), pp. 24–39. ISSN: 07350015. URL: http://www.jstor.org/stable/25800776 (visited on Oct. 7, 2025).

Gabaix, Xavier and Yannis M. Ioannides (2004). "Chapter 53 - The Evolution of City Size Distributions". In: *Cities and Geography.* Ed. by J. Vernon Henderson and Jacques-FranÃ§ois Thisse. Vol. 4. Handbook of Regional and Urban Economics. Elsevier, 2004, pp. 2341–2378. DOI: https://doi.org/10.1016/S1574-0080(04)80010-5. URL: https://www.sciencedirect.com/science/article/pii/S1574008004800105.

Ganapathy, Kumar and Subbarayan Arumugam (2014). "The temporal dynamics of regional city size distribution: Andhra pradesh (1951-2001)". *Journal of Mathematics and Statistics* 10 (Apr. 2014), pp. 221–230. DOI: 10.3844/jmssp.2014.221.230.

Giesen, Kristian and Jens Súdekum (2010). "Zipf's law for cities in the regions and the country". *Journal of Economic Geography* 11.4 (May 2010), pp. 667–686.

González-Val, Rafael (2023). "Parametric, semiparametric and nonparametric models of urban growth". *Cities* 132 (2023), p. 104079. ISSN: 0264-2751. DOI: https://doi.org/10.1016/j.cities.2022.104079. URL: https://www.sciencedirect.com/science/article/pii/S0264275122005182.

González-Val, Rafael, Luis Lanaspa, and Fernando Sanz-Gracia (2014). "New Evidence on Gibrat's Law for Cities". *Urban Studies* 51.1 (2014), pp. 93–115. DOI: 10.1177/0042098013484528. eprint: https://doi.org/10.1177/0042098013484528. URL: https://doi.org/10.1177/0042098013484528.

González-Val, Rafael and Fernando Sanz-Gracia (2024). "A test of the relationship between the Pareto exponent and sample size". *Investigaciones Regionales - Journal of Regional Research* (Nov. 2024). DOI: 10.38191/iirr-jorr.24.064.

González-Val, Rafael, Domingo P. Ximénez-de-Embún, and Fernando Sanz-Gracia (2024). "A long-term, regional-level analysis of Zipf's and Gibrat's laws in the United States". *Cities* 149 (2024), p. 104946. ISSN: 0264–2751. DOI: https://doi.org/10.1016/j.cities.2024.104946. URL: https://www.sciencedirect.com/science/article/pii/S0264275124001604.

Kundak, Seda and Vedia Dókmeci (2018). "A Rank-Size Rule Analysis of The City System at The Country and Province Level in Turkey". *ICONARP International Journal of Architecture and Planning* 6.1 (June 2018), pp. 77–98. DOI: 10.15320/ICONARP.2018.39. URL: https://iconarp.ktun.edu.tr/index.php/iconarp/article/view/202.

Nitsch, Volker (2005). "Zipf zipped". *Journal of Urban Economics* 57.1 (2005), pp. 86–100. ISSN: 0094-1190. DOI: https://doi.org/10.1016/j.jue.2004.09.002. URL: https://www.sciencedirect.com/science/article/pii/S0094119004000981.

Peña, Guillermo (2025). "Wildfire impacts on Spanish municipal population". *Journal of Environmental Management* 377 (2025), p. 124504. ISSN: 0301-4797. DOI: https://doi.org/10.1016/j.jenvman.2025.124504. URL: https://www.sciencedirect.com/science/article/pii/S0301479725004803.

Peña, Guillermo, Miguel Puente-Ajovín, and Fernando Sanz-Gracia (2025). "Gibrat's law for CO2 emissions revisited". *Applied Economics Letters* 32.6 (2025), pp. 825–829. DOI: 10.1080/13504851.2023.2289893.

Peña, Guillermo and Fernando Sanz-Gracia (2021). "Zipf's exponent and Zipf's law in the BRICS: a rolling sample regressions approach". *Economics Bulletin* 41.4 (None 2021), pp. 2543–2549. DOI: None. URL: https://ideas.repec.org/a/ebl/ecbull/eb-21-00417.html.

Peng, Guohua (2010). "Zipf's law for Chinese cities: Rolling sample regressions". *Physica A: Statistical Mechanics and its Applications* 389.18 (2010), pp. 3804–3813. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2010.05.004. URL: https://www.sciencedirect.com/science/article/pii/S037843711000381X.

Pérez Valbuena, Gerson Javier and Adolfo Meisel Roca (2014). "Ley de Zipf y de Gibrat para Colombia y sus regiones: 1835-2005". *Revista de Historia Economica - Journal of Iberian and Latin American Economic History* 32.2 (2014), pp. 247–286. DOI: 10.1017/S021261091400007X. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84927560528&doi=10.1017%2fS021261091400007X&partnerID=40&md5=68c9882f2b40be279e30c777ea7d0ab1.

Pérez-Campuzano, Enrique, Lev Guzmán-Vargas, and F. Angulo-Brown (2015). "Distributions of city sizes in Mexico during the 20th century". *Chaos, Solitons Fractals* 73 (2015), pp. 64–70. ISSN: 0960-0779. DOI: https://doi.org/10.1016/j.chaos.2014.12.015. URL: https://www.sciencedirect.com/science/article/pii/S0960077914002355.

Sanso-Navarro, Marcos and María Vera-Cabello (2023). *The impact of terrorism at the municipality level in Colombia: A new difference-in-differences approach.* Tech. rep. Social Sciences Research Network, 2023. DOI: 10.2139/ssrn.4130613.

Soo, Kwok Tong (2005). "Zipf's Law for cities: a cross-country investigation". *Regional Science and Urban Economics* 35.3 (2005), pp. 239–263. ISSN: 0166-0462. DOI: https://doi.org/10.1016/j.regsciurbeco.2004.04.004. URL: https://www.sciencedirect.com/science/article/pii/S016604620400033X.

Subbarayan, A. (2009). "The size distribution of cities in Tamilnadu (1901-2001)". *International Journal of Agricultural and Statistical Sciences* 5.2 (2009). Cited by: 7, pp. 373–382. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-78049344274&partnerID=40&md5=366848f35fefdfb23c829f4b62d6b1b5.

Taylor, L. R., I. P. Woiwod, and J. N. Perry (1978). "The Density-Dependence of Spatial Behaviour and the Rarity of Randomness". *Journal of Animal Ecology* 47.2 (1978), pp. 383–406. ISSN: 00218790, 13652656. URL: http://www.jstor.org/stable/3790 (visited on Sept. 19, 2025).

Taylor, Leroy R. (1961). "Aggregation, Variance and the Mean". *Nature* 189 (1961), pp. 732–735. URL: https://api.semanticscholar.org/CorpusID:4263093.

Wan, Guanghua, Dongqing Zhu, Chen Wang, and Xun Zhang (2020). "The size distribution of cities in China: Evolution of urban system and deviations from Zipf's law". *Ecological Indicators* 111 (2020), p. 106003. ISSN: 1470-160X. DOI: https://doi.org/10.1016/

j.ecolind.2019.106003. URL: https://www.sciencedirect.com/science/article/pii/S1470160X19309987.

Wang, Qiuping A. (2021). "Principle of least effort vs. maximum efficiency: deriving Zipf-Pareto's laws". *Chaos, Solitons Fractals* 153 (2021), p. 111489. ISSN: 0960-0779. DOI: https://doi.org/10.1016/j.chaos.2021.111489. URL: https://www.sciencedirect.com/science/article/pii/S0960077921008432.

Wei, Jing, Jianjun Zhang, Bofeng Cai, Ke Wang, Sen Liang, and Yuhuan Geng (2021). "Characteristics of carbon dioxide emissions in response to local development: Empirical explanation of Zipf's law in Chinese cities". *Science of The Total Environment* 757 (2021), p. 143912. ISSN: 0048-9697. DOI: https://doi.org/10.1016/j.scitotenv.2020.143912. URL: https://www.sciencedirect.com/science/article/pii/S004896972037443X.

Ye, Xinyue (2011). "Re-examination of Zipf's law and urban dynamic in China: A regional approach". *Annals of Regional Science* 49 (Aug. 2011), pp. 1–22. DOI: 10.1007/s00168-011-0442-8.

Ziqin, Wen (2016). "Zipf law analysis of urban scale in China". *Asian Journal of Social Science Studies* 1 (Mar. 2016), p. 53. DOI: 10.20849/ajsss.v1i1.21.

Tables and figures

Table 1: Descriptive statistics of the used variables

Variable	Obs	Mean	Std. dev.	Min	Max
Population	650	637869.9	736975.70	89057	6747781
lnPopulation		13.04	0.69	11.42	15.68
nuts1		3.94	1.73	1	7
Code_Community		7.66	4.58	1	17
Income_Quartiles		2.52	1.12	1	4
Zipf_treatment		7.62	4.11	0	13
Gibrat_treatment		4.3	1.95	0	9
Cimate_code		2.62	1.31	1	5

Table 2: Pareto Coefficient per NUTS 3 level for Zipf's Law when coefficient equals 1

Year					Province/Pareto coefficient					
Province	Araba/Álava	Badajoz	Cádiz	Cuenca	Huelva	Rioja, La	Navarra	Pontevedra	Sevilla	Valencia/València
1900	1.12867992	0.82662638	0.8454609	1.1568493	0.84690302	0.97721655	1.05079574	1.18011014	0.84359693	0.8597014
1910	1.08186717	0.82315799	0.86047194	1.16997524	0.82888414	0.97714457	1.04131014	1.25340189	0.86245202	0.84111905
1920	1.07059921	0.84757163	0.86546239	1.15245504	0.93370107	0.96169227	1.00680351	1.27668166	0.89567236	0.83298221
1930	1.05693467	0.85399677	0.87655002	1.10899173	0.89314909	0.92411846	0.98460968	1.26311265	0.92862647	0.81884926
1940	1.0636757	0.86482632	0.83215918	1.1047087	0.88802497	0.9051626	0.96642297	1.22466416	0.94681255	0.8076737
1950	1.06245816	0.94511333	0.80614398	1.07859015	0.86503378	0.89156825	0.95253777	1.18498984	0.91821729	0.79888656
1960	0.99058889	1.00006797	0.77962102	1.07570586	0.83568157	0.85236977	0.90647977	1.18482149	0.91684349	0.78455187
1970	0.84024029	0.97341854	0.70607642	0.93067095	0.76366399	0.70299277	0.77519691	1.1451021	0.87937599	0.69809804
1981	0.7402857	0.88223347	0.65902129	0.81348749	0.6910965	0.61888616	0.70477259	1.07413272	0.82803843	0.62775777
1991	0.73442123	0.84649397	0.62542386	0.78209838	0.65897866	0.61769687	0.68023587	1.03715049	0.81671416	0.60746625
2001	0.73971863	0.84000222	0.6076592	0.75250978	0.63684542	0.60572915	0.644104	0.97036074	0.8199938	0.60086618
2011	0.72023379	0.81000513	0.59068948	0.70427876	0.62065892	0.57039468	0.60111146	0.90311409	0.77969567	0.59282389
2020	0.69980772	0.79144997	0.58113704	0.67086744	0.60421018	0.55283347	0.5670365	0.86866098	0.75920468	0.57288518
Province	Albacete	Balears, Illes	Castellón/Castelló	Girona	Huesca	Lugo	Ourense	Salamanca	Soria	Valladolid
1900	1.16998447	1.03748979	1.02220905	1.1581711	1.01495109	1.24404785	2.08737888	1.41328399	1.11951289	1.15366309
1910	1.15356951	1.02831423	0.99648757	1.11830449	1.06731705	1.26976413	2.07460518	1.40526708	1.12105286	1.1493048
1920	1.16228605	1.00258996	0.98137833	1.10479925	1.09600556	1.2946088	2.06852374	1.40331835	1.10339419	1.15659445
1930	1.12369822	0.98759343	0.96348465	1.06058902	1.07207386	1.34549775	1.9173426	1.36042727	1.09654649	1.11143401
1940	1.09038048	0.93105451	0.93812716	1.04303828	1.07926737	1.40815224	1.86932464	1.33879653	1.08017559	1.08330486
1950	1.07381318	0.92205197	0.89471137	1.02516916	1.04486591	1.41297461	1.80586747	1.29565949	1.04995641	1.07631112
1960	1.03873934	0.87829336	0.81134119	0.95730613	1.00479792	1.39757868	1.72097018	1.27695943	1.02154899	1.03312019
1970	0.92042154	0.78277737	0.6906767	0.82561283	0.91677985	1.37655371	1.566638	1.1942031	0.9077037	0.92043606
1981	0.83571316	0.74871757	0.60290986	0.73998034	0.84650371	1.27714335	1.44743731	1.10236548	0.82320829	0.81945171
1991	0.77511364	0.72539569	0.58688927	0.70429713	0.82771455	1.18804692	1.37147472	1.01296428	0.78893422	0.77547333
2001	0.74296704	0.70179102	0.56741476	0.69283608	0.83678962	1.10550446	1.24636095	0.94764609	0.76015397	0.72711777
2011	0.70626359	0.66627021	0.54796013	0.66567846	0.81298739	1.01969196	1.1418222	0.90091281	0.7225678	0.6719141
2020	0.6823548	0.6472371	0.53295224	0.65120093	0.78676136	0.98726214	1.06779353	0.87173357	0.69645747	0.63998501
Province	Alicante/Alacant	Barcelona	Ciudad Real	Granada	Jaén	Madrid	Asturias	Santa Cruz de Tenerife	Tarragona	Bizkaia
1900	0.79965546	0.95833967	0.76286518	1.06065362	0.97797548	0.90892582	0.97107462	1.45458076	1.09205203	1.01787206
1910	0.78615262	0.95508039	0.76982134	1.08302677	1.03998804	0.90268908	0.9527433	1.37279608	1.07923663	0.98363596
1920	0.78077103	0.94138101	0.84043733	1.09621327	1.09511436	0.88665082	0.92245739	1.43575148	1.0383749	0.94548579
1930	0.78803614	0.87137586	0.8699748	1.10080699	1.13658851	0.86778956	0.90431207	1.41527808	0.99066023	0.92354337
1940	0.78610739	0.83662785	0.89647902	1.085851	1.12678224	0.84733288	0.89696928	1.34670805	0.97162957	0.91437069
1950	0.75757005	0.79844365	0.89916164	1.05801297	1.12487959	0.81850332	0.87721337	1.25808517	0.90590424	0.88713813
1960	0.70654569	0.69892197	0.90310688	1.04785104	1.12802968	0.74681814	0.81954425	1.16787338	0.83851563	0.78589666
1970	0.62655368	0.57126108	0.8430578	0.986495	1.03996872	0.60812932	0.75291835	1.01057565	0.72678612	0.668457
1981	0.55203195	0.50456879	0.75351153	0.90330238	0.95812174	0.5044045	0.68052131	0.92886149	0.65992332	0.59390663
1991	0.5177278	0.49289822	0.73189049	0.84958009	0.90953452	0.48051524	0.65247213	0.8855033	0.66343214	0.58217568
2001	0.50275715	0.48397917	0.70052064	0.7820187	0.87089862	0.46212633	0.63240753	0.85507435	0.65146245	0.58436298
2011	0.47163027	0.47178127	0.66839188	0.72727531	0.84308924	0.45019077	0.60905278	0.78878346	0.61137543	0.59589955
2020	0.45990921	0.46238693	0.66124969	0.70112633	0.82016755	0.44151009	0.59421464	0.77241673	0.5908752	0.59365714
Province	Almería	Burgos	Córdoba	Guadalajara	León	Málaga	Palencia	Cantabria	Teruel	Zamora
1900	1.00982399	1.1543712	0.97938546	1.25874075	1.38725594	1.02941494	1.18557659	1.30718689	1.26232234	1.46482396
1910	0.98609157	1.15613507	1.02609343	1.25621184	1.35949279	1.02309482	1.18169075	1.27087903	1.23485611	1.48132394
1920	1.01645206	1.14381492	1.03435005	1.26723018	1.31043723	0.99124925	1.15696862	1.23431284	1.23008128	1.47133956
1930	1.05077353	1.12939832	1.01644469	1.23796848	1.33225477	0.96365152	1.13289443	1.19273655	1.2325468	1.4421217
1940	1.01402578	1.10487671	0.96423801	1.22290455	1.30664154	0.92887358	1.09422798	1.16391884	1.19845281	1.40512228
1950	0.95927165	1.09551766	0.9817206	1.20942012	1.28174837	0.89094095	1.07528085	1.13445373	1.17832941	1.37160822
1960	0.91558047	1.06388375	0.97981318	1.18031602	1.19867546	0.87758619	1.03971004	1.04056012	1.10334572	1.34207729
1970	0.83574169	0.95867238	0.89984746	1.00158033	1.097328	0.80025753	0.95051012	0.92440552	0.91138508	1.28091929
1981	0.7466238	0.86639831	0.83038047	0.82220169	1.00605999	0.7078469	0.88097851	0.82233831	0.7872473	1.23047245
1991	0.67284138	0.84511081	0.7872048	0.81436148	0.94389861	0.65381553	0.84099431	0.77669759	0.78372625	1.17679523
2001	0.62448224	0.83279729	0.76319444	0.7670498	0.8822256	0.61526458	0.79382991	0.72518218	0.76328043	1.11020846
2011	0.59086643	0.7686975	0.74655574	0.65903133	0.83152823	0.58363733	0.75497276	0.68455163	0.7611078	1.04455538
2020	0.57101363	0.76366832	0.76059592	0.65976853	0.80562906	0.5596861	0.74945198	0.65837157	0.75291361	0.98352415
Province	Ávila	Cáceres	Coruña, A	Gipuzkoa	Lleida	Murcia	Palmas, Las	Segovia	Toledo	Zaragoza
1900	1.29516556	0.79479504	1.81692778	0.8701642	1.07881521	0.67702345	1.15508842	1.32299421	0.99627333	1.11295122
1910	1.28467091	0.79107944	1.80912472	0.83455648	1.06762905	0.74719821	1.10013777	1.33243531	1.01029552	1.12258526
1920	1.26672988	0.80664134	1.76335037	0.8060382	1.13319189	0.7744742	1.08060594	1.31094062	1.01557257	1.10224836
1930	1.24243723	0.81522079	1.69353707	0.77475245	1.18890247	0.82484402	1.05758919	1.28309167	1.00853212	1.08066203
1940	1.25122251	0.82513621	1.61947881	0.75236143	1.19542703	0.84293807	0.97427265	1.25851847	1.01938308	1.06842553
1950	1.24701906	0.84825354	1.50965666	0.72103021	1.12541734	0.84229761	0.9586251	1.22785133	1.02845718	1.0428374
1960	1.22197445	1.06255828	1.42344747	0.67169509	1.04719285	0.83095874	0.93868918	1.2116257	1.00565122	0.97780767
1970	1.13229284	1.08276163	1.31816582	0.58545022	0.90663326	0.79595663	0.81732775	1.04812831	0.90230836	0.80569582
1981	1.00058831	0.97940819	1.17999478	0.51860875	0.83576471	0.74015096	0.75587229	0.89339136	0.81513757	0.68011109
1991	0.93960909	0.92720266	1.08796575	0.50944636	0.82357958	0.71667435	0.76031639	0.85921013	0.81823797	0.67908974
2001	0.88924454	0.88696976	1.02315745	0.51015063	0.80659214	0.67421793	0.76912263	0.81289512	0.72129333	0.68330893
2011	0.82890941	0.85969846	0.94926856	0.52872343	0.76727716	0.63809287	0.71885537	0.74543709	0.63825813	0.64231954
2020	0.78899848	0.84890837	0.91043592	0.523382	0.7466131	0.58814815	0.54018908	0.6981764	0.63280942	0.6231306

Table 3: P-value per NUTS 3 level for testing Gibrat Law

Year	Province/P-value									
Province	Araba/Álava	Badajoz	Cádiz	Cuenca	Huelva	Rioja, La	Navarra	Pontevedra	Sevilla	Valencia/València
1910	0.549	0.811	0.059	0.044	0.415	0.094	0.536	0.56	0.276	0.973
1920	0.642	0.045	0.579	0.077	0.773	0.002	0.054	0.18	0.366	0.643
1930	0.108	0.001	0.031	0.012	0.83	0.007	0.039	0.025	0.366	0.109
1940	0.101	0.683	0.172	0.051	0.451	0.001	0	0.024	0.281	0.002
1950	0.856	0.534	0.163	0.021	0.491	0.093	0.001	0.087	0.272	0.418
1960	0.007	0.979	0.106	0.006	0.117	0.004	0	0.33	0.541	0.777
1970	0	0.01	0.023	0	0.002	0	0.025	0.01	0.295	0.046
1981	0.007	0	0.017	0	0.001	0.001	0.011	0.009	0.369	0.07
1991	0.163	0	0.238	0	0.033	0.1	0.165	0.135	0.968	0.437
2001	0.657	0.008	0.898	0.009	0.161	0.147	0.23	0.256	0.717	0.807
2011	0.768	0	0.113	0	0.159	0.202	0.461	0.329	0.688	0.81
2020	0.187	0.001	0.115	0.004	0.142	0.157	0.645	0.08	0.518	0.306
Province	Albacete	Balears, Illes	Castellón/Castelló	Girona	Huesca	Lugo	Ourense	Salamanca	Soria	Valladolid
1910	0.028	0.532	0.827	0.004	0.648	0.037	0.869	0.711	0.978	0.906
1920	0.483	0.254	0.057	0.627	0.894	0.002	0.405	0.596	0.818	0.635
1930	0.001	0.58	0.011	0	0.197	0.944	0.08	0	0.264	0.018
1940	0	0	0	0.001	0.328	0.236	0.332	0	0.123	0.018
1950	0.11	0.095	0	0.003	0	0.018	0	0.017	0.007	0.267
1960	0.17	0.005	0	0	0	0.153	0.088	0.087	0.009	0.005
1970	0	0.026	0	0	0	0.258	0.01	0	0	0
1981	0	0.211	0	0	0	0.031	0.036	0	0	0.001
1991	0.004	0.872	0.031	0.009	0.002	0.006	0.068	0.092	0.002	0.229
2001	0.001	0.899	0.053	0.995	0.613	0	0.033	0.433	0.037	0.633
2011	0.015	0.781	0.144	0.041	0.156	0	0.027	0.465	0.015	0.547
2020	0.013	0.703	0.012	0.902	0.025	0.001	0.003	0.23	0.189	0.63
Province	Alicante/Alacant	Barcelona	Ciudad Real	Granada	Jaén	Madrid	Asturias	Santa Cruz de Tenerife	Tarragona	Bizkaia
1910	0.114	0.762	0.261	0.367	0	0.728	0.6	0	0.464	0.426
1920	0.145	0.368	0.479	0.801	0.88	0.026	0.001	0.002	0	0.071
1930	0.453	0.26	0.138	0.747	0.017	0.18	0.019	0.655	0.005	0.511
1940	0.266	0.348	0.978	0.106	0.27	0.206	0.001	0.268	0.002	0.057
1950	0.008	0.027	0.185	0.685	0.06	0.588	0.011	0	0	0.288
1960	0.002	0.383	0.249	0.817	0.564	0.306	0.048	0.075	0.005	0.045
1970	0	0.543	0	0.013	0.001	0.518	0	0.23	0	0.055
1981	0.001	0.639	0	0.001	0	0.97	0	0.102	0	0.198
1991	0.124	0.288	0	0.33	0.001	0.569	0.002	0.88	0.594	0.667
2001	0.751	0.066	0.001	0.668	0.005	0.125	0.004	0.617	0.778	0.156
2011	0.925	0.171	0	0.349	0.102	0.083	0.005	0.531	0.278	0.028
2020	0.086	0.444	0	0.172	0.002	0.969	0.017	0.283	0.025	0.785
Province	Almería	Burgos	Córdoba	Guadalajara	León	Málaga	Palencia	Cantabria	Teruel	Zamora
1910	0.892	0.378	0.082	0.509	0.835	0.966	0.487	0.483	0.019	0.327
1920	0.22	0.32	0.703	0.925	0.003	0.281	0.028	0.545	0.973	0.491
1930	0.084	0.007	0.026	0.113	0.449	0.037	0.098	0.634	0.557	0.517
1940	0.006	0	0.067	0.538	0	0.014	0.001	0.021	0.177	0
1950	0.039	0.004	0.412	0.392	0.002	0.087	0.046	0.292	0.013	0
1960	0.036	0	0.528	0.047	0	0.332	0.002	0.013	0	0.004
1970	0.058	0	0	0	0	0.075	0	0.001	0	0
1981	0.004	0	0	0	0	0.009	0	0.003	0	0
1991	0.038	0.036	0.066	0.134	0	0.197	0.003	0.058	0.001	0.002
2001	0.041	0.514	0.166	0.362	0.041	0.392	0.02	0.667	0.007	0.007
2011	0.14	0.328	0.357	0.104	0.006	0.58	0.212	0.83	0.049	0.019
2020	0.005	0.397	0.06	0.012	0.232	0.038	0.397	0.293	0.046	0.593
Province	Avila	Cáceres	Coruña, A	Gipuzkoa	Lleida	Murcia	Palmas, Las	Segovia	Toledo	Zaragoza
1910	0.384	0.175	0.713	0.006	0.634	0.228	0.017	0.384	0.004	0.748
1920	0.042	0.025	0.072	0	0.061	0.464	0.781	0.2	0.928	0.066
1930	0.062	0.892	0.015	0.126	0.596	0.31	0.369	0.082	0.879	0.099
1940	0.401	0.293	0.002	0.259	0.603	0.848	0.518	0.162	0.32	0.007
1950	0.172	0.478	0	0.066	0	0.972	0.069	0.037	0.09	0.024
1960	0.001	0.352	0.001	0.194	0.001	0.343	0.616	0.02	0.014	0.002
1970	0	0.183	0.007	0.116	0	0.713	0.057	0	0	0.002
1981	0	0	0.004	0.024	0.002	0.233	0.182	0	0	0.043
1991	0	0	0.04	0.126	0.148	0.712	0.271	0.118	0.103	0.371
2001	0.002	0.001	0.434	0.787	0.675	0.856	0.193	0.436	0.682	0.495
2011	0.001	0.001	0.054	0.051	0.061	0.936	0.497	0.597	0.831	0.512
2020	0.023	0.085	0.004	0.78	0.064	0.964	0.705	0.06	0.012	0.238

Table 4: ATTs of Zipf rejection on population Unconditional and Conditional with Temporal Lags and Geographical factors (Geo)

		Unconditional			Lags			Geo			Lags & Geo		
		Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value
Average	Pre_event	-2324.781	8941.8	0.795	5426.919	11169.85	0.627	-2353.346	7851.986	0.764	6137.123	7330.084	0.402
	Post_event	**948992.1**	263431.3	0	**851221**	436067.9	0.051	**756361.6**	290229.5	0.009	**451829.7**	248294.6	0.069
Event	T	19670.07	26841.12	0.464	40725.79	15021.83	0.007	32518.58	19909.64	0.102	40562.1	17404.56	0.02
Post-event	T+1	61218.71	54504.17	0.261	98910.48	29786.19	0.001	77071.54	41929.69	0.066	85771.09	32022.56	0.007
	T+2	146347.8	96811.82	0.131	214739.3	67768.41	0.002	183545.2	80670.36	0.023	195482.9	62058.3	0.002
	T+3	257060.3	157766.8	0.103	374753.4	127729.2	0.003	330370.9	133309.9	0.013	347761.9	101027.5	0.001
	T+4	328634.7	198972.1	0.099	459563.6	166046.9	0.006	354200.1	179088.7	0.048	358891.8	124482.6	0.004
	T+5	712345.2	358161.1	0.047	724622.4	349994.5	0.038	617888	374699.8	0.099	449111.3	248489.7	0.071
	T+6	1801736	778192	0.021	1770215	1015099	0.081	1478361	811582.4	0.069	857433.9	648676.3	0.186
	T+7	2438111	973625.3	0.012	2546676	1447872	0.079	2076476	999588.6	0.038	1294645	883996.7	0.143
	T+8	1685626	892043.8	0.059	1430782	1273842	0.261	1296847	933711.8	0.165	436806.6	578181.4	0.45
	T+9	1147279	117534.9	0				818469.4	231490.6	0			
	T+10	1418649	156953.3	0				944426.5	324552.1	0.004			
	T+11	1371229	163770.4	0				866165	335468.3	0.01			

Table 5: ATTs on population of Zipf and Gibrat rejection by climate regions subsamples for Geo & Lags

		Zipf bsk			Zipf csb			Zipf csa			Gibrat csa		
		Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value
Average	Pre-event	-26280.21	18608.77	0.158	15987.37	15422.43	0.3	14818.15	10693.89	0.166	-46773.2	29665.49	0.115
	Post-event	1013840	359581.3	0.005	150162	66110.5	0.023	1569887	141665.4	0	-361645.1	145200.6	0.013
Event	T	-102643.4	58238.45	0.078	22976.7	24317.73	0.345	79715.04	19181.44	117310	-50309.5	22809.64	0.027
Post-event	T+1	-160113.9	157382.6	0.309	40001.87	38109.97	0.294	182331.7	32928.15	246869.6	-97481	66237.78	0.141
	T+2	-90571	306830.5	0.768	119849.7	39225.62	0.002	366613.6	105802.5	573982.6	-112458.7	151758	0.459
	T+3	-37838.52	586350.8	0.949	269513.1	130129.1	0.038	625014.8	232863.9	1081420	-118208.3	205097.2	0.564
	T+4	-85802.86	622658.6	0.89	298468.6	153111.7	0.051	800625.9	345182.5	1477171	-190256.8	207710.4	0.36
	T+5	1082496	1352524	0.424				1360060	654877	2643596	-262725.1	217982.2	0.228
	T+6	3603503	7.24E-10	0				3045871	6.10E-10	3045871	-221141.5	315246.4	0.483
	T+7	3901688	7.24E-10	0				3677163	2.55E-10	3677163	-242847.8	346373.7	0.483
	T+8							3991585	3.81E-10	3991585	-638853.1	152863.9	0
	T+9										-868424.4	147335	0
	T+10										-1175390	120913.9	0

Table 6: ATTs on population of Zipf rejection by NUTS 1 macroregions subsamples for non-treated and not-yet treated (NY) control groups with Geo & Lags

		Center			Center NY			Center+Madrid Autonomous Community			Center+Madrid Autonomous Community NY		
		Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value
Average	Pre-event	4711.02	3718.748	0.205	-332.0271	2636.437	0.9	5707.785	3820.744	0.135	3331.707	4585.957	0.468
	Post-event	74664.53	18222.38	0	68659.2	17808.89	0	1926696	122928.4	0	1854465	126128.6	0
Event	T	19929.22	7405.79	0.007	13672.64	6960.207	0.049	21919.31	7174.478	0.002	17595.16	7436.547	0.018
Post-event	T+1	34126.61	10213.3	0.001	23479.4	9224.881	0.011	77684.21	43031.24	0.071	66255.33	42297.01	0.117
	T+2	62097.09	13711.7	0	54738.17	14461.95	0	255384.8	184713.7	0.167	251178.3	187015.7	0.179
	T+3	124460.3	32621.15	0	118696.4	33103.94	0	514180.8	365665.9	0.16	512559.5	369902.1	0.166
	T+4	132709.4	35632.29	0	132709.4	35632.29	0	552497.1	393911.3	0.161	550759.2	392702.2	0.161
	T+5										4014525	189189.2	0
	T+6										4210346	2.19E-10	0
	T+7										5212501	7.01E-10	0

Table 7: ATTs on population of Zipf rejection by income quartile regions subsamples with Geo & Lags

		First Income Quartil (Q1)			Second Income Quartil (Q2)			Third Income Quartil (Q3)			Fourth Income Quartil (Q4)		
		Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value	Coefficient	Std. err.	p-value
Average	Pre-event	6042.67	10887.62	0.58	2670.92	14838.43	0.86	776742.30	71868.72	0.00	14440.50	9871.09	0.14
	Post-event	207965.40	91558.70	0.02	210873.40	103068.60	0.04	1129405.00	201305.90	0.00	1183796.00	492552.00	0.02
Event	T	23986.50	39658.13	0.55	41973.63	72429.48	0.56	401386.00	117564.40	0.00	48618.12	43125.73	0.26
Post-event	T+1	69354.25	67008.31	0.30	211798.90	223333.50	0.34	687538.70	202259.70	0.00	160854.70	71925.80	0.03
	T+2	188530.80	82340.72	0.02	289257.20	300089.00	0.34	835405.00	219257.20	0.00	593057.20	451636.00	0.19
	T+3	353612.20	142991.90	0.01	319758.50	96114.20	0.00	1070915.00	262907.90	0.00	1656791.00	1012291.00	0.10
	T+4	404343.40	160777.70	0.01	308492.70	117020.10	0.01	1611616.00	320157.00	0.00	2064688.00	843300.40	0.01
	T+5				93959.37	0.00	0.00	2169567.00	294248.50	0.00	2215951.00	845547.10	0.01
	T+6										1546611.00	954576.50	0.11

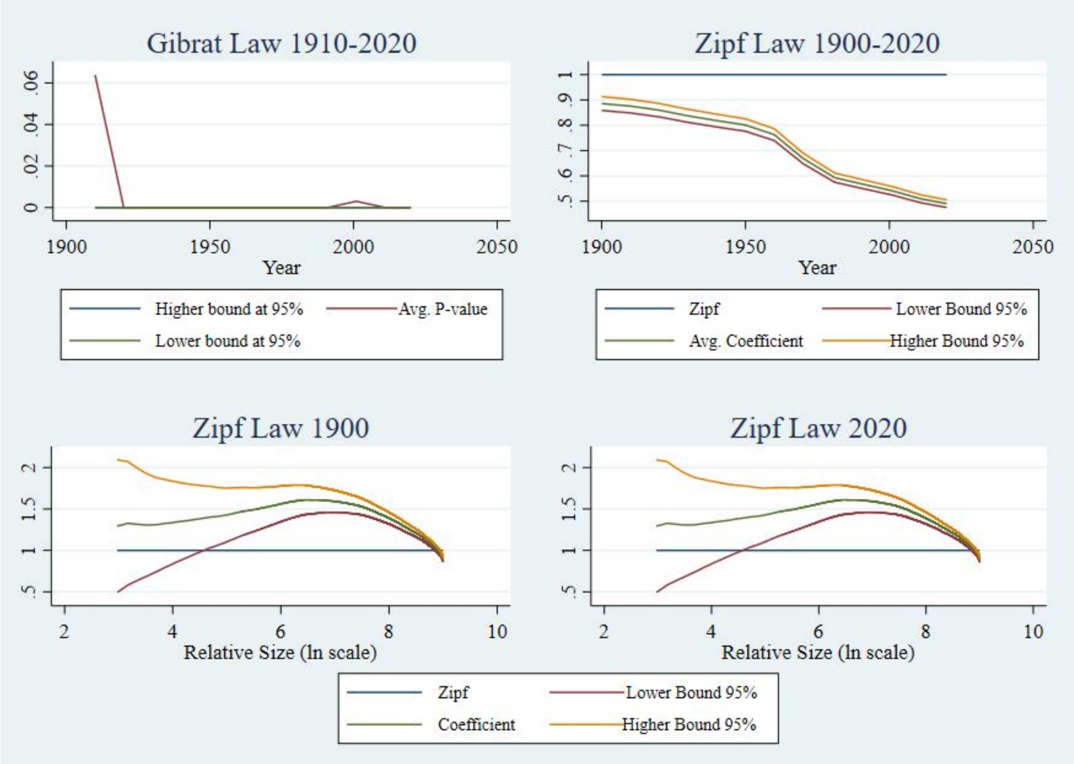

Figure 1: Average (above) Zipf (right) and Gibrat (left) coefficients per year and rolling sample regressions for Zipf law (bottom) on 1900 (left) and 2020 (right)

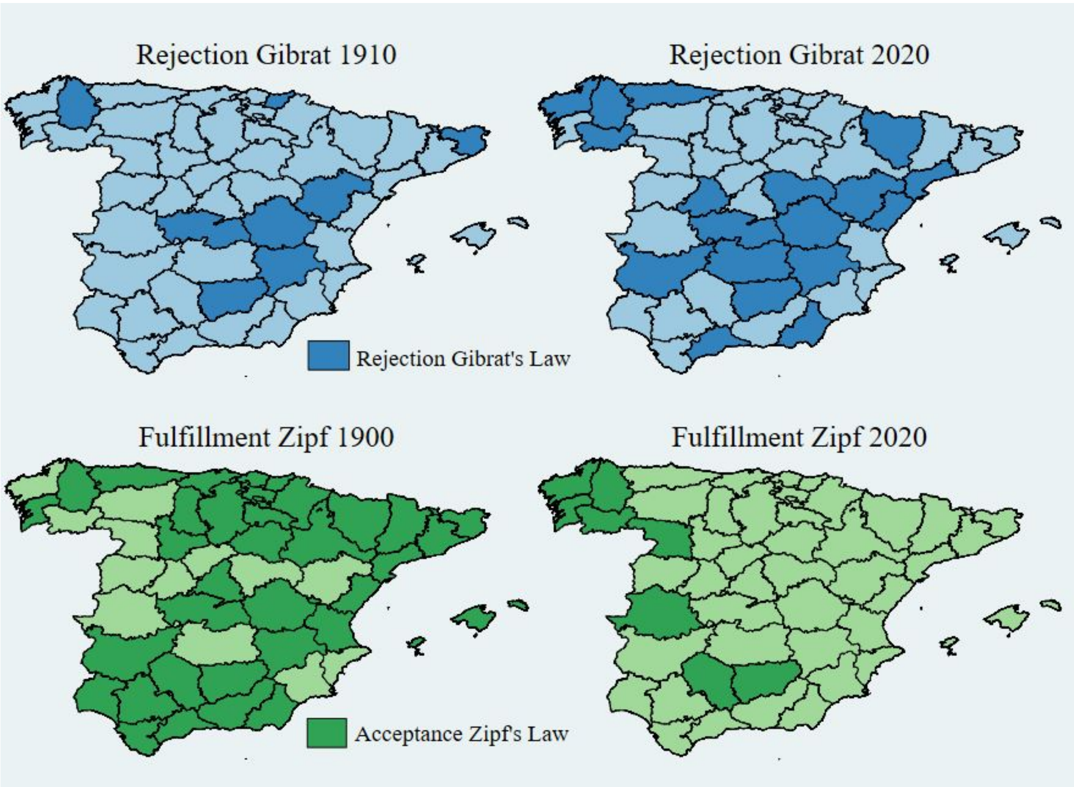

Figure 2: Maps representing the NUTS 3 level administrations (provinces) where Gibrat (up) and Zipf (down) Laws were, respectively, accepted and rejected in 1900 (left) and 2020 (right)

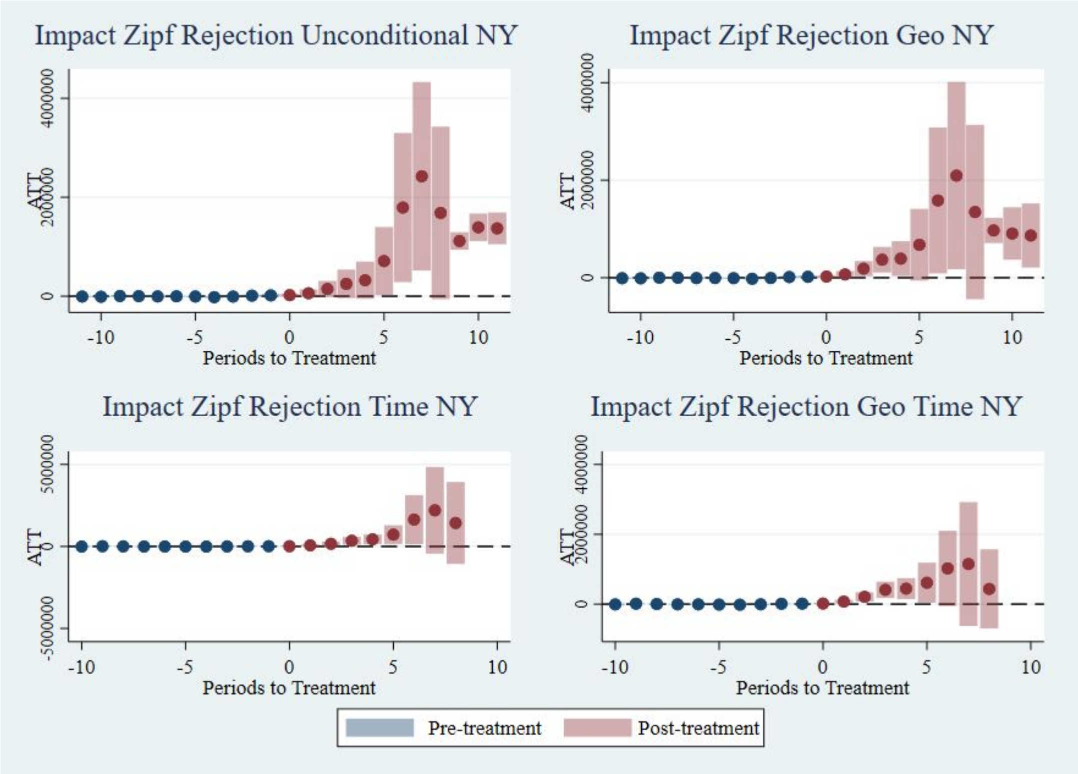

Figure 3: ATT effects of first rejection of Zipf law after compliance on population for not-yet treated (NY) control group, Unconditional (above) and with climate conditions

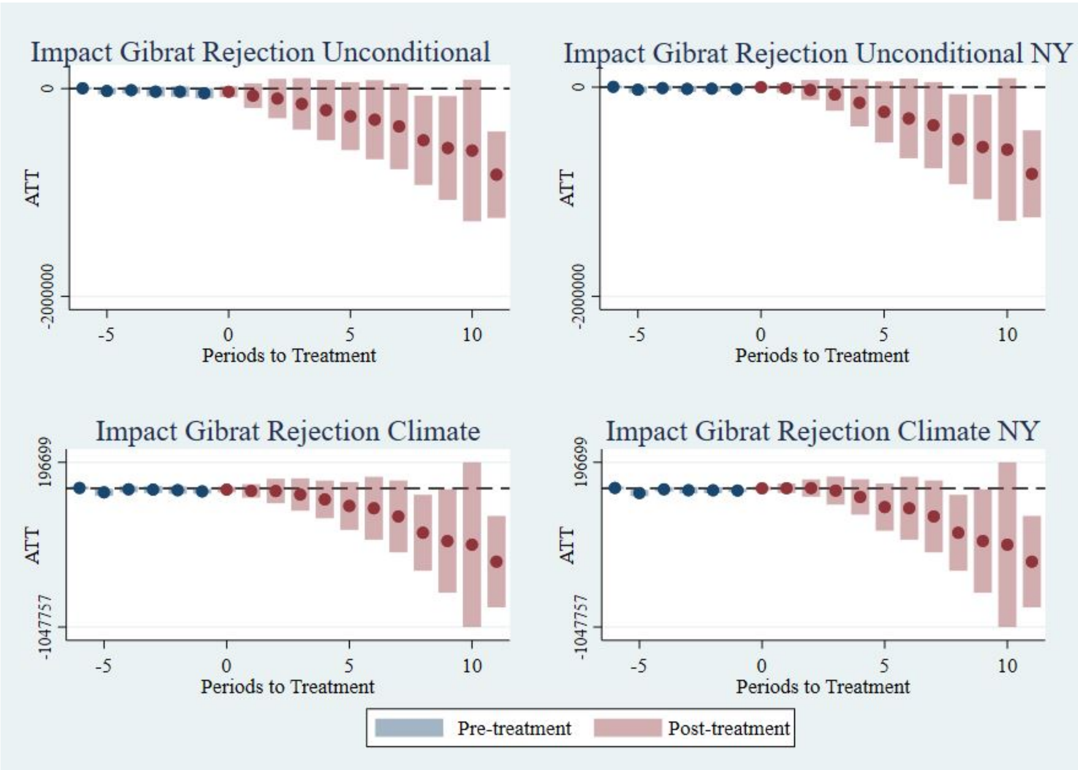

Figure 4: ATT effects of first rejection of Gibrat law after compliance on population with unconditional and conditional (climate zone variables) for not treated and not-yet treated (NY) control groups

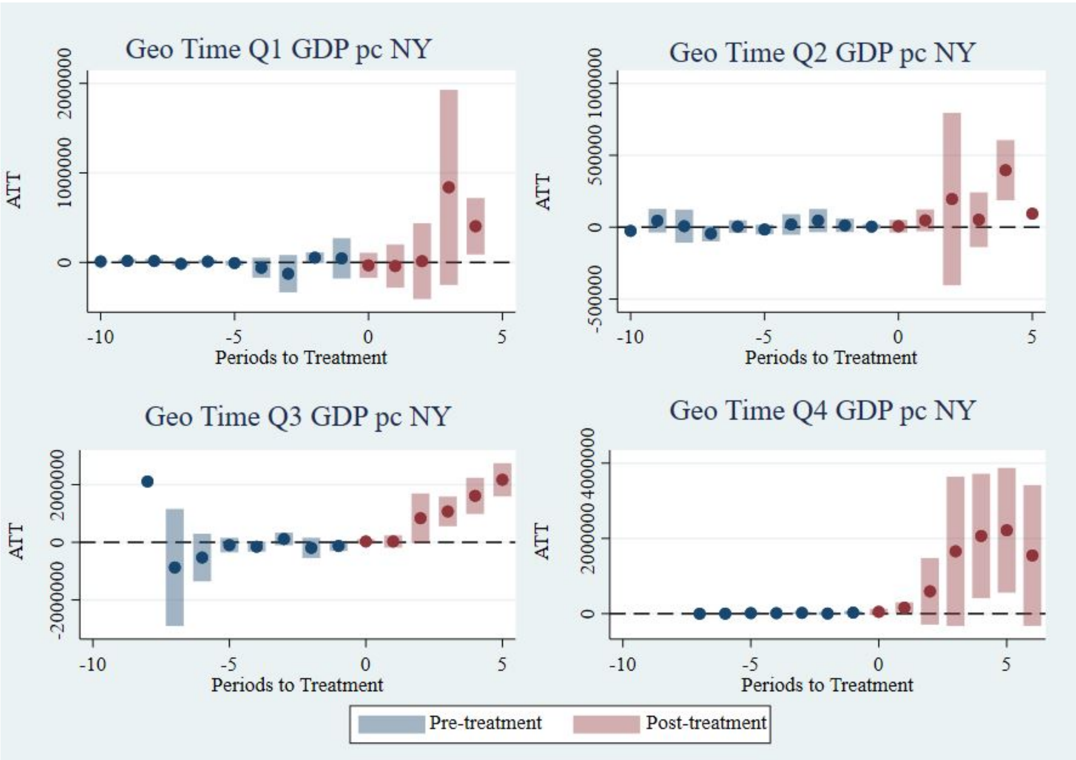

Figure 5: ATTs on population of Zipf rejection by income (GDP pc) quartile regions subsamples for not-yet treated (NY) control group for Geo and Time (lags)