Modeling Rank Distribution and the Relative Importance Factor Index in Power-Law Models: Application to Social Resilience Using Scopus Databases

**OTRO POSIBLE TITULO: Modeling Rank Distribution and Introducing the Relative Importance Factor Index in Heavy-Tailed Distributions: Application to Social Resilience Using Scopus Databases**

Brian Llinas1,2[0009-0002-1344-9168], Erika Frydenlund2[[0000-0002-7694-7845](https://orcid.org/0000-0002-7694-7845)], Katherine Palacio3[[0000-0001-6134-7700](https://orcid.org/0000-0001-6134-7700)], Humberto Llinas4[[0000-0002-2976-5109](https://orcid.org/0000-0002-2976-5109)], Jose J. Padilla2[0000-0003-0720-4148]

1 Computer Science Department, Old Dominion University, Norfolk, USA

2 Virginia Modeling, Analysis, and Simulation Center, Old Dominion University, Suffolk, USA

3 Department of Industrial Engineering, Universidad del Norte, Barranquilla, Colombia

4 Department of Mathematics and Statistics, Universidad del Norte, Barranquilla, Colombia

bllin001@odu.edu

**Abstract.** Heavy-tailed distributions, such as the power-law distribution, are prevalent in both natural and human-made phenomena, characterized by their slow-decaying tails. Previous studies have primarily focused on modeling these distributions based on frequencies, often neglecting the estimation of rank distributions. This paper addresses this gap by introducing a novel approach to model the rank distribution of frequent factors and by defining the Relative Importance Factor (RIF) index. The first objective is to develop and validate a formula that accurately models the rank distribution of frequent factors. The second objective is to introduce the RIF index, comparing the probability of a factor occupying the first position with its probability of being in any other position within the range. We apply our methodology to databases downloaded from Scopus, focusing on social resilience and analyzing the number of citations of articles. Using maximum likelihood estimation (MLE) to determine power-law parameters and the Kolmogorov-Smirnov (KS) statistic to estimate the optimal threshold, we ensure robust modeling. Additionally, we employ bootstrapping techniques to assess the uncertainty of our estimates. By modeling the rank distribution and introducing the RIF index, this study enhances the understanding of heavy-tailed distributions and provides a valuable tool for analyzing complex systems.

**Keywords:** Power-law distribution; Rank distribution; Relative Importance Factor (RIF); Heavy-tailed distributions; Social resilience; Citation analysis

1. Introduction

Heavy-tailed distributions are common in both natural and human-made phenomena. These probability distributions are characterized by a slower decay in their tails compared to the normal or exponential distribution. Several distributions that fall into this category include the log-normal and stretched exponential distributions, among others. In addition to these, the most well-known heavy-tailed distribution is the power-law distribution (Jiang et al., 2023). Besides its ability to describe phenomena with heavy-tailed distributions, the power-law model has captured the interest of researchers in various disciplines because it can also be applied to analyze the behavior of many complex systems. This makes it ideal for describing phenomena in diverse areas (Corral and Gonzalez, 2019; Blasius, 2020; Chen, 2021; Jiang, and de Rijke, 2021; Sethna, 2022). However, limitations in data collection often result in empirical datasets that cover a narrow observation range, complicating the clear identification of power-law behavior (Navas-Portella et al., 2019).

Power-law distributions have several characteristics. One of them is their scale-invariance property, which means that the shape of the distribution remains constant regardless of the scale at which it is observed (Corral and Gonzalez, 2019; Banshal, 2020). Another is that when the entire range is plotted on a linear axis, the curve takes the shape of a perfect "L". Moreover, when represented on a logarithmic scale, the curve always appears straight (Chen et al., 2020; Banshal, 2020).

In computer science, the power law has been applied to the study of networks, where it has been observed that a few nodes have an extremely high number of connections while most have few (Artico et al., 2020; Devlin, 2021; Somin et al., 2022). This pattern has also been observed in social networks, where a few users have a large number of followers, and most have few (Xu et al., 2019; Arthur and Williams, 2019; Rajput et al., 2020). Additionally, in risk management, the power law is used to model the occurrence of extreme events, such as natural disasters (Pisarenko and Rodkin, 2019; Yum, 2023; Sohn et al., 2023) or financial crises (Dufrénot and Paret, 2019; Taleghani, 2020; Ghosh et al., 2021; Ben Yaala and Henchiri, 2023).

In economics, it has been used to describe the distribution of wealth, where a small fraction of the population holds most of the financial resources while another suffers from extreme poverty (Masseran, 2019; Cardoso et al., 2020; Safari et al., 2020; Puttanapong et al., 2022; Kumer, 2024). Similarly, power laws have been used to study the distribution of crime, where a few criminals commit most of the crimes while the majority commit only a few (He et al., 2023; Ng et al., 2023). In ecology, it has been employed to model the distribution of forest types (Atkins et al., 2022), freshwater fishes (Baumgartner and Peláez, 2024), hot-spring microbiomes (Li and Ma, 2019), among others. In physics, power laws are fundamental for studying the behaviors of the decay rate as a link between dissociation energy and temperature (Fischer and Schweikhard, 2020), fully developed turbulent flows in a smooth pipe (Afzal et al., 2023), and nonlinear phonon hydrodynamics (Sciacca and Jou, 2024).

In the health sector, power laws have been used to model the distribution of epidemics, where a few outbreaks can infect a large number of people, while most outbreaks affect a much smaller number (Blasius, 2020; Neipel et al., 2020; Jha, 2020). This model has also been applied to analyze the distribution of health resources, such as the availability of hospital beds or the allocation of medications, where a few hospitals or health centers receive the majority of resources (Srivastav et al., 2021; Dong et al., 2021).

In the field of academic publications, the academic influence of articles, journals, authors, etc., has been studied for several decades. Currently, one of the most widely used metrics and long recognized as an important indicator for evaluating the impact of a journal or author is the number of citations (Zhao et al., 2019). Numerous studies have revealed that the citation distribution of scientific articles follows a power law (Thelwall and Nevill, 2018; Arroyo-Machado et al., 2020; Benatti et al., 2023). In particular, the number of articles with a specific number of citations \(x\) is proportional to \(x\) raised to a negative scale exponent (Popescu, 2003; Banshal, 2020).

It is emphasized that previous studies model the power-law distribution based solely on frequencies but do not estimate the distribution of the rank. Therefore, firstly, in this work, we present and demonstrate a formula that allows us to model the distribution of the rank of the most frequent factors. Secondly, we introduce a parameter, which we will call

**(BRIAN. SON DOS OPCIONES DE NOMBRE. Para continuar, sigo con el de preferencia, el primero):**

1. Relative Importance Factor index (RIF index).

2. Relative Prominence Factor index (RPF index).

The RIF index compares the probability of a factor occupying the first position with the probability of it being in any other position within the range. This innovation provides a new perspective for evaluating the relative prominence and importance of factors, an aspect that has not been considered in previous studies.

The proposed theory will be applied using databases downloaded from Scopus in the area of social resilience. In particular, we will focus on the number of citations of articles published in this area. This approach allows us to identify citation distribution patterns and evaluate how academic attention is concentrated on certain articles and specific topics but based on the RIF index. The choice of social resilience as a study area responds to the growing importance of this topic in the context of social sciences and its relevance for public policy formulation and the implementation of resilient practices in communities and organizations. According to Dagdeviren et al. (2020), social resilience depends on power relations, rules/institutions, and resource distribution, which are interconnected. Without favorable conditions in these three elements, individuals may be overwhelmed by crises or survive through harmful mechanisms.

The article is organized as follows. Section 1 contains the introduction. In Section 2, we explain the power-law model and the different methods for estimating the scale parameter and the optimal threshold of the model. In Section 3, we introduce the distribution of the rank and the so-called Relative Importance Factor index. In Section 4, we present an application. Finally, in Section 5, we present the conclusions.

1. Power-law for data frequencies
   1. The power law model

Power-law distributions can appear in two forms: continuous distributions, which govern continuous real numbers, and discrete distributions, where the quantity of interest can only take a discrete set of values, typically positive integers (Clauset et al., 2009). Variables of the problem are defined as follows: (1) denotes the variable of interest, and (2) is the frequency of occurrence of the factor within . Considering that our data capture the frequency with which F occurs, our study deals with a discrete problem. Based on that, the discrete power-law model to estimate the probability that the factor within appears with frequency can be defined as

where > 0 is the exponent or scaling parameter for the type of frustration in the country , indicating the steepness of the distribution, and is a normalization constant to ensure the probabilities sum up to 1 for the type of frustration in the country . Both and depend on the distribution and can be found in Clauset et al. (2009).

This distribution diverges to zero when , so there must be a lower bound > 0 for the behavior of the power-law. To calculate the normalizing constant, can be defined as

where is the generalized or Hurwitz zeta function.

* 1. The power law model

To ensure accurate estimation of the power-law exponent and determine the range over which the power-law behavior is applicable, it is necessary to decide on the lower bound . This estimation helps identify the specific part of the media coverage distribution where the power-law model is valid. Moreover, obtaining an estimate is crucial for deriving an unbiased estimate of the power-law exponent .

According to Clauset et al. (2009), if we assume that our data are sampled from a power-law distribution for values of greater than or equal to , the maximum likelihood estimator (MLE) for in the discrete case is defined as

where is the number of occurrences by *i*-th factor and are the observed values of *x* such that , where is the estimated value of . We estimate using the maximum likelihood (ML) method. The corresponding log-likelihood function for this estimation process is derived from the data for media coverage factors related to frustrations in Colombia and Greece.

Note in the above formula that, to calculate we must first estimate . To estimate the lower bound on the power-law , we used a metric known as the Kolmogorov-Smirnov (KS) statistic (Ramos et al., 2024), which is defined as the maximum difference between , the cumulative distribution function (CDF) of the observations, and , the CDF of the power law that optimally fits the data . The statistic is defined as

The fitting process is sometimes performed by linear regression using logarithmically transformed variables. This approach is used because applying the logarithm to the power law function results in

Thus, a power law appears as a straight line with slope in a logarithmic plot. It is important to note that changes in the scale parameter can affect the slope of the curve in the log-log plot, resulting in changes in the shape and behavior of the distribution represented by the power law. As an example can be seen in Figure 1.

The bootstrapping procedure is used to analyze the uncertainty associated with exponent estimation. It consists of randomly selecting data samples with replacement and then applying MLE procedure with a KS cutoff on that sample. This process is repeated several times to evaluate the uncertainty. In this study, 1000 iterations were performed on all data sets. In addition, we perform a particular hypothesis test and provide the corresponding p-value to test : one power-law distribution fits adequately versus : another distribution might fit better. We will use the R 4.3.1 package poweRlaw (Gillespie, 2014) to perform all the above analyses.

1. Power-law for rank and RIF index
   1. Relation between the frecuencies and the rank

For a type of frustration in the country L, we will suppose that the frequency of occurrences is a descending ordered series. That is,

According to Popescu (2003), Zipf’s law of rank-frequency states that, in a generally ordered set of data, the frequency of occurrence of an element is inversely proportional to its rank :

Where is the rank of the element (1 for the most frequent, 2 for the second most frequent, etc.); is the frequency of the element in the rank for a type of frustration in the country L; is the power law exponent, and is a constant of proportionality. This implies that

That is, to estimate the parameters and , we only need to estimate the previous regression model with response variable and explanatory variable . Once these estimates have been found, we can write .

* 1. Estimation of the distribution of the rank

Let be the variable representing the rank of the occurrence frequency for a type of frustration in the country L. Under the condition that , we know that and are inversely proportional for all ordered frequencies . Then, for all , the probability that, in the country L, this factor is in the rank is:

where is the power law exponent for the rank ; is a proportionality constant and is the normalising constant, defined by

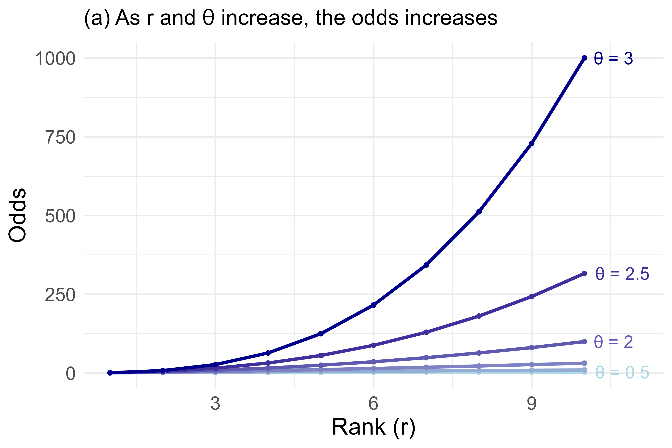
* 1. The Relative Importance Factor (RIF) Index

We can define the odds ratio (estimated) of probabilities with respect to the rank , as

This estimate will help us explain how the probability of observing the first element in the rank (rank 1) compares to the probability of observing the element in the ranking . We know that . Then, in this case, is given by

This ratio tells us how many times more likely it is to observe the first element in the rank (rank 1) compared to the element in rank . When , then . This is trivially 1, since we are comparing the probability of the first rank with itself. Suppose that . Then, the probability of observing the first element in the rank is times higher than the probability of observing the element in the rank . In other words,

For different values of , the Figure 1(a) and Figure 1(b) show how the rank r relates to both the odds ratio and the probability Given that is positive, the graph illustrates how the odds and the probability change as both r and are increased.

A graph with lines and lines on a black background

Description automatically generated

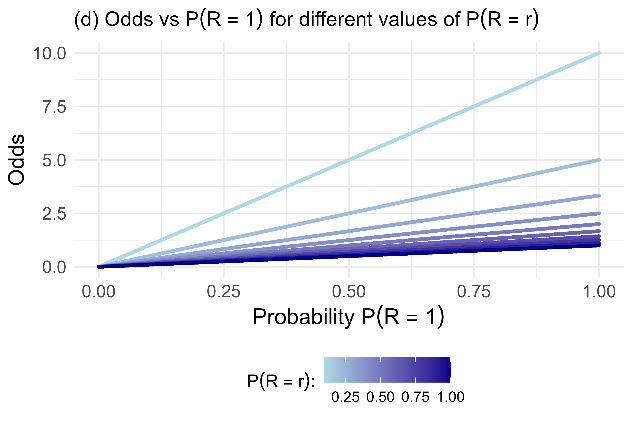
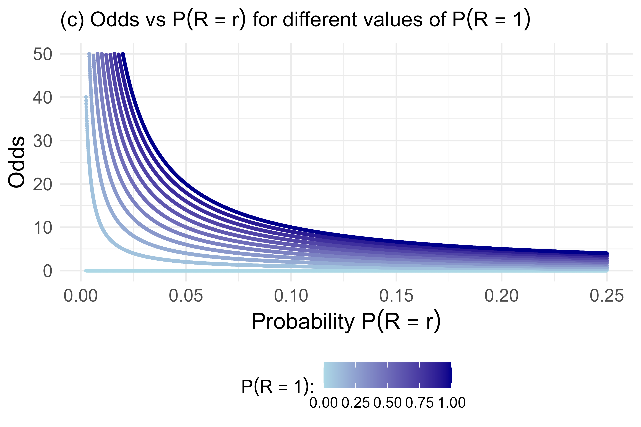


Figure 1. Distribution of: (a) odds and (b) probability vs rank, for different values of odds vs (c) P(R=r); (d) P(R=1)

Each line, color-coded according to the legend, represents the distribution for a particular value of . The dark line shows a steep increase (decrease) in odds ratio (probability) with higher ranks, indicating rapid growth (decline), while other lines exhibit more gradual or stable trends in the odds. This suggests significant variability in how odds ratio (probability) changes with rank across different values of , highlighting the unique impact of rank on odds ratio (probability) for each type. As r increase and becomes more positive, the odds increase (the probability decrease) exponentially. The higher the value of , the steeper the increase (decrease) in the odds (probability) with respect to r. This indicates that for larger values of , the disparity between the probability of the highest-ranked event and lower-ranked events becomes much more pronounced. As r increases, the odds diverge significantly and the probability converges to 0, especially for higher values, highlighting an exponential growth pattern.

The Figure 1(c) explores the relationship between the probability and the odds for different values of . As increases, the odds decrease sharply initially and then level off. The gradient color represents different values of , indicating that higher values of correspond to higher odds for a given The Figure 1(d) examines the relationship between the probability , and the odds for different values of . As , increases, the odds increase linearly. The gradient color represents different values of , showing that higher values of , lead to higher odds for a given .

1. Application to Scopus database
   1. Log-log scale graphical representation for each type of frustration

Figure 1 shows the percentage frequency distribution of occurrences from various data sources. Notably, the distributions are only defined for integer values in the ranges, and the connecting lines do not imply continuity in the paths. For both the Colombia and Greece plots, the *x-axis* represents the rank of the factors, while the *y-axis* represents the proportion of occurrences.

Figure 2(a) shows the plot of the percentage frequencies  as a function of the appropriate *x*-range for Colombia. The distributions for geopolitics, government, infrastructure, and migrants demonstrate a clear decline in proportion as the rank increases, indicating that a few factors dominate media coverage while many factors receive less attention. On the other hand, Figure 2(b) shows the plot of the percentage frequencies as a function of the appropriate *x*-range for Greece. Like Colombia, the distributions for geopolitics, government, infrastructure, and migrants in Greece show a decline, indicating that a few factors are more frequently covered in the media.

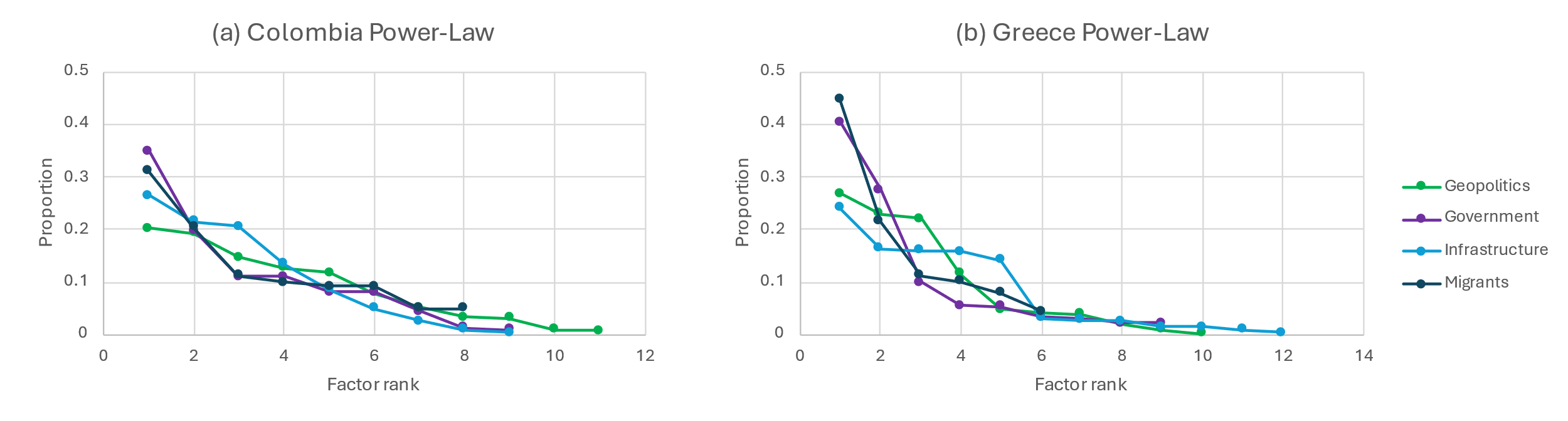


Figure 2. Representation of percentage frequencies of different types of frustrations as a function of the range of occurrences in Colombia and Greece. Subfigure 1(a) shows the percentage frequencies on a linear scale for Colombia, and subfigure 1(b) shows the percentage frequencies on a linear scale for Greece.

Figure 3 shows the percentage frequency distribution of occurrences from various data sources on a log-log scale. For both the Colombia and Greece plots, the *x-axis* represents the logarithmic rank of the factors, while the *y-axis* represents the logarithmic proportion of occurrences. Figure 3(a) shows the plot of the percentage frequencies as a function of the appropriate *x*-range is for Colombia on a log-log scale. The log-log plot reveals a linear trend for each type of frustration (geopolitics, government, infrastructure, and migrants), indicating that the media coverage could follow a power-law distribution. The linearity suggests that a few factors dominate media coverage while many factors are less frequently covered. Also, Figure 3(b) presents the same data for Greece on a log-log scale. Similar to Colombia, the log-log plot for Greece shows a linear trend for each type of frustration. This consistency indicates that the media coverage in Greece could also follow a power-law distribution, with a few dominant factors and many less frequent ones.

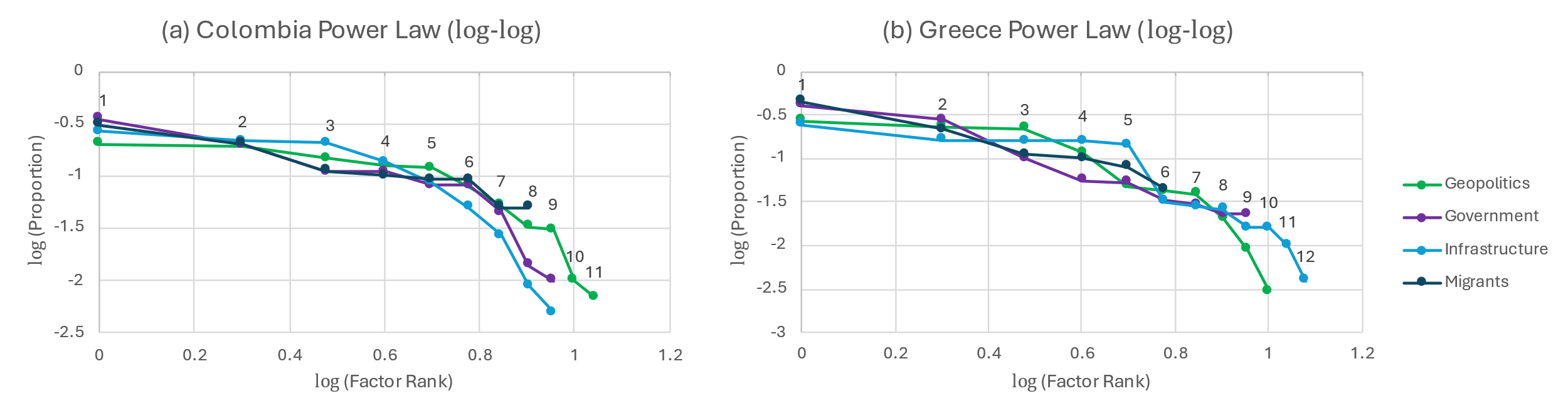


Figure 3. Log-log representation of percentage frequencies of different types of frustrations as a function of the range of occurrences in Colombia and Greece. Subfigure 2(a) displays the log-log plot for Colombia, and subfigure 2(b) displays the log-log plot for Greece.

The observation of linearity in the log-log plots (Figure 3) suggests that both Colombia and Greece exhibit characteristics that indicate a power-law distribution, which is recognized by its linear behavior when plotted on a logarithmic scale. However, this observation alone does not provide conclusive evidence. As explained in Section 3.3, it is common to set a minimum value below which a power-law distribution cannot hold. This value, ​, can be visually identified as the point where the linear trend begins to disappear.

In Figure 2 and Figure 3, a consistent linear trend is observed for different types of frustrations, starting with the most frequent factor (rank 1) within the type of frustration in country and continuing up to aspects with significantly lower frequencies. However, this linear trend fades away at the lower end of the graphs, indicating that the graphs exhibit characteristics suggestive of a power-law distribution. The minimum value ​ or final range where the linear trend is lost in Figure 3 is crucial for confirming the presence of a power-law distribution. The Parameters and will be estimated in Section 4.2.

* 1. Estimation results with MLE+KS vs. bootstrapping methods

In addition to the analysis carried out in Section 4.1, it is possible to perform a quantitative analysis of the corresponding distributions. Section 3.3 explained a technique called the MLE step combined with the KS cutoff criterion. This approach was used to optimally fit a power-law distribution to the distribution of observed perceptions in each frustration-related dataset. Both Table 1 and Table 2 summarize the most relevant results for the four datasets (for Colombia and Greece). The R package called poweRlaw (Gillespie, 2014) was employed for fitting purposes to perform this analysis. Table 1 shows the results obtained using the MLE+KS cutoff for both Colombia and Greece. In contrast, Table 2 shows the results of the uncertainty assessment using the bootstrapping method for Colombia and Greece.

### PoweRlaw estimations (MLE + KS) for Colombia and Greece

The following table presents the results[[1]](#footnote-2) of the estimations using the Maximum Likelihood Estimation (MLE) and Kolmogorov-Smirnov (KS) statistics for different types of frustrations in Colombia and Greece. For each type of frustration (F) in the country (L), the table provides a comparative analysis of the estimated scale parameter , the estimated minimum value , the cumulative percentage CP % (frequency ) of data in the tail[[2]](#footnote-3), and the KS statistic. These results offer insights into the distribution patterns of various frustrations in the media coverage of both countries.

Table 1. Results of the estimations (MLE + KS) for Colombia and Greece

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| = Colombia (222 observations) | | | | | = Greece (499 observations) | | | | |
| Frustration F (#*n*) |  |  | CP % () |  | Frustration F (#*n*) |  |  | CP % () |  |
| Infrastructure (9) | 1.70 | 6 | 77.8 (7) | 0.273 | Infrastructure (12) | 6.33 | 71 | 41.7 (5) | 0.236 |
| Migrants (8) | 3.50 | 24 | 75.0 (6) | 0.241 | Migrants (6) | 2.49 | 61 | 83.3 (5) | 0.188 |
| Government (9) | 2.98 | 24 | 66.7 (6) | 0.254 | Government (9) | 1.96 | 17 | 100.0 (9) | 0.167 |
| Geopolitics (11) | 4.77 | 69 | 45.5 (5) | 0.240 | Geopolitics (10) | 1.68 | 7 | 80.0 (8) | 0.234 |

**Note:**

Number of factors in each type of frustration.

= Estimated scale parameter.

= Estimated cutoff value (it has very different values due to the different weights used in the data).

Percentage (frequency) accumulated of data .

KS= Kolmogorov-Smirnov statistic.

For infrastructure, Colombia has an value of 1.7, significantly lower than Greece’s 6.33. This suggests that the distribution declines less steeply in Colombia, meaning there is a wider spread of occurrences across different factors. Conversely, Greece’s high indicates a steep decline, with a few factors dominating media coverage. Colombia’s is 6, while Greece's is 71, indicating that the power-law behavior for infrastructure in Greece starts at a higher threshold. Additionally, Colombia has a higher CP % (77.8%) than Greece (41.7%), suggesting a larger proportion of the data fits the power-law distribution in Colombia. Both countries have similar KS values, indicating comparable fit quality.

For migrants, Colombia has an value of 3.5, higher than Greece’s 2.49. This indicates a steeper distribution in Colombia, where the occurrences drop off more sharply across factors. Greece’s is much higher at 61 compared to Colombia’s 24, suggesting the power-law behavior in Greece begins at a higher number of occurrences. Greece also has a slightly higher CP % of 83.3% compared to Colombia’s 75%, indicating a greater proportion of the data fits the power-law distribution in Greece. The KS values are lower for Greece (0.188) than Colombia (0.241), suggesting a better fit in Greece.

For government, Colombia’s value is 2.98, higher than Greece’s 1.96, indicating a steeper distribution in Colombia with a more rapid decline in occurrences across different factors. Greece’s ​ is lower at 17 compared to Colombia’s 24, suggesting that the power-law behavior starts at a lower threshold in Greece. Greece has a perfect CP % of 100%, suggesting all data points fit the power-law distribution, whereas Colombia’s CP % is 66.7%. The KS values show that Greece has a better fit (0.167) than Colombia (0.254).

For geopolitics, Colombia has an value of 4.77, significantly higher than Greece’s 1.68, suggesting a much steeper distribution in Colombia. Greece’s is 7, much lower than Colombia’s 69, indicating that the power-law behavior in Greece starts at a lower number of occurrences. Greece has a higher CP % of 80% compared to Colombia's 45.5%, indicating a larger proportion of data fits the power-law distribution in Greece. The KS values are similar, indicating comparable fit quality.

In summary, the comparative analysis reveals distinct patterns in the distribution of different frustrations between Colombia and Greece. Greece generally exhibits flatter distributions (lower values) and higher thresholds for power-law behavior (), indicating a more uniform distribution of occurrences across factors. In contrast, Colombia shows steeper distributions (higher values) with lower thresholds, suggesting a concentration of occurrences around a few factors. The KS values indicate the fit quality is relatively comparable between the two countries, though Greece often shows a slightly better fit. These differences highlight the varying nature of media coverage of frustrations between Colombia and Greece.

### Bootstrapping uncertainty evaluation for Colombia and Greece

The following table presents the results[[3]](#footnote-4) of the bootstrapping uncertainty evaluation for different types of frustrations in Colombia and Greece. Bootstrapping is a statistical method used to estimate the distribution of a statistic by resampling with replacement from the original data. For each type of frustration (F) in the country (L), this table includes the mean and standard deviation (SD) of both the estimated scale parameter and the minimum value ​; the tail length ​; and the to test : one power-law distribution fits adequately versus : another distribution might fit better. The results provide insights into the variability and reliability of the power-law parameter estimates. Figure 4 at the end of this section provides a visual summary of the main findings.

Table 2. Results of the uncertainty assessment using the bootstrapping method for Colombia and Greece

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| = Colombia (\*) | | | | | = Greece (\*) | | | | |
| Frustration (#*n*) | Mean (SD) | | | | Frustration (#*n*) | Mean (SD) | | | |
|  |  |  |  |  |  |  |  |
| Infrastructure (9) | 2.07 (1.09) | 8.46 (9.37) | 7.10 (1.63) | 0.052 | Infrastructure (12) | 3.93 (2.66) | 35.4 (34.2) | 8.73 (2.99) | 0.394 |
| Migrants (8) | 4.62 (2.10) | 25.0 (7.14) | 6.00 (1.49) | 0.287 | Migrants (6) | 3.43 (1.67) | 68.2 (30.7) | 4.87 (0.99) | 0.895 |
| Government (9) | 3.87 (1.99) | 24.4 (10.48) | 6.10 (1.76) | 0.248 | Government (9) | 2.39 (0.92) | 24.2 (12.1) | 7.94 (1.57) | 0.665 |
| Geopolitics (11) | 4.27 (2.31) | 48.7  (30.71) | 7.10 (2.61) | 0.529 | Geopolitics (10) | 2.05 (1.09) | 10.15 (9.45) | 7.74 (1.83) | 0.153 |

**Note:**

(\*) Bootstrapping based on 1000 iterations; >0.05 implies that one power-law distribution fits adequately.

Number of factors in each type of frustration.

SD= standard deviation.

= Mean of the estimated scale parameter.

= Mean of estimated cutoff value.

Mean of frequency accumulated of data .

For Infrastructure, Colombia has a mean value of 2.07 (SD 1.09), which is lower than Greece’s mean value of 3.93 (SD 2.66). This suggests that in Colombia, the distribution of occurrences declines less steeply, indicating a broader spread of occurrences across different factors. Greece’s mean ​ is significantly higher at 35.4 (SD 34.2) compared to Colombia’s 8.46 (SD 9.37), suggesting that the power-law behavior in Greece starts at a much higher threshold. The tail lengths are similar, with Colombia at 7.10 (SD 1.63) and Greece at 8.73 (SD 2.99). for Greece (0.394) indicates a better fit to a power-law model than Colombia (0.052).

For Migrant, Colombia has a mean value of 4.62 (SD 2.10), higher than Greece’s 3.43 (SD 1.67). This suggests a steeper distribution in Colombia, with occurrences dropping off more sharply across factors. Greece’s mean ​ is significantly higher at 68.2 (SD 30.7) compared to Colombia’s 25.0 (SD 7.14), indicating that the power-law behavior in Greece begins at a higher number of occurrences. Greece’s tail length is shorter at 4.87 (SD 0.99) compared to Colombia’s 6.00 (SD 1.49), with a higher (0.895) suggesting a better fit to a power-law model than Colombia (0.287).

For Government, Colombia’s mean value is 3.87 (SD 1.99), higher than Greece’s 2.39 (SD 0.92), indicating a steeper distribution in Colombia with a more rapid decline in occurrences across different factors. The mean ​ values are similar, with Colombia at 24.4 (SD 10.48) and Greece at 24.2 (SD 12.1). Greece’s tail length is longer at 7.94 (SD 1.57) compared to Colombia’s 6.10 (SD 1.76). Greece’s higher (0.665) suggests a better fit to a power-law model than Colombia (0.248).

For Geopolitics, Colombia’s mean value is 4.27 (SD 2.31), significantly higher than Greece’s 2.05 (SD 1.09), suggesting a much steeper distribution in Colombia. Greece’s mean is 10.15 (SD 9.45), much lower than Colombia’s 48.70 (SD 30.71), indicating that the power-law behavior in Greece starts at a lower number of occurrences. The tail length is similar, with Colombia at 7.10 (SD 2.61) and Greece at 7.74 (SD 1.83). Greece’s lower (0.153) compared to Colombia (0.529) suggests a slightly better fit to a power-law model.

In summary, the comparative analysis of bootstrapping uncertainty evaluation reveals distinct patterns in the distribution of different types of frustrations between Colombia and Greece. Greece generally exhibits higher mean values and thresholds, indicating steeper distributions and higher starting points for power-law behavior. In contrast, Colombia shows more varied and spread-out distributions. The tail lengths and suggest that Greece’s data generally fits a power-law model better than Colombia’s, highlighting differences in media coverage and public perception of frustrations between the two countries.

* 1. Frustration factors probability estimations

For each type of frustration (F) in the country (L), Table 3 to Table 6 presents the frequency  (percentage %) distribution[[4]](#footnote-5) of the respective ranged factor f of F for Colombia and Greece, where 1 denotes the factor with the highest frequency or proportion of occurrence and the last rank with the lowest frequency . Each table also displays , based in the probabilities estimated through the power-law model based on the estimated values and according to MLE+KS method (as shown in Table 1). Additionally, the ratio between the estimated probabilities of the first rank and those of the other ranks is included. The notion of each table is as follows: r =Rank; =Frecuency (count) of each factor f of F in country L; % =Percentage; Ratio=; Prob.= ; =Estimations of scaling parameters; =Estimated cutoff value; =Estimations of the regression parameters; =Normalization constants; =Proportionality constant; KS=MLE + KS method; and Boos.=Bootstrapping method. Interpretations will focus on results obtained using the MLE + KS method and the top ranked factors. The interpretations for the results obtained with the bootstrapping method are analogous. The factors for each type of frustration and their corresponding descriptions can be found in the FSM Codebook (version 0.5) by Cleveland et al. (2023).

### Blaming towards Infrastructure

The results obtained with the two methods in Colombia are very similar to each other. However, compared to the bootstrapping method, which found only 6 factors, the MLE+KS method found 7 more factors. For this reason, our interpretation of the results for this country will be the MLE+KS results only. In Colombia, the most frequently reported factor related to infrastructure is about *locals’ perception of healthcare*, with 59 occurrences (26.6%) and an estimated probability approximate of 0.616. This indicates that healthcare issues are the most prominent in media coverage. The ratio shows that *healthcare* is approximately 3.5 times more likely to be reported than *housing*, the second most frequent factor, which has a probability of 0.175 and 48 occurrences (21.6%). *Economics* is the third most frequently mentioned factor, with 46 occurrences (20.7%) and probability of 0.084, making it 7.3 times less likely to be reported than *healthcare*. The factors *Other locals on infrastructure* (rank 4), *Sanitation* (rank 5), *Environment* (rank 6)and *Education* (rank 7) are also significant factors, with respective probabilities between 0.01 and 0.05, highlighting that these issues, while less frequent, are still of notable concern. They have and odds of between 12 and 35. These values show that *Healthcare* is perceived as a significantly more frequent problem compared to *Infrastructure*, *Sanitation*, *Environment* and *Education*, with the greatest disparity compared to *Education*.

In contrast, in Greece, the results obtained with both methods show some significant differences. In general, the factor with the highest frequency and probability is *locals’ perceptions of housing*, with 121 reports (24.2%) and an estimated probability of 0.663 (MLE+KS) y 0.486 (Bootstrapping). This factor is considerably more dominant in Greece than Colombia’s leading factor. Recall that the values in the "Ratio" column indicate how many times lower the probability of perception of each factor is compared to *Housing*. The second most frequent factor in Greece is *Environmental* issues, with 82 occurrences (16.4%) and a lower probability of 0.177 y 0.214 (Bootstrapping). This indicates that while environmental concerns are significant, they are reported far less frequently than *Housing*, with a ratio of 8.1 (MLE+KS) y 3.7 (Bootstrapping). *Locals’ perceptions of sanitation*, *economics*, and *NGOs* also have a minimal role, with probabilities between 0.03 and 0.14, respectively. These factors have significantly lower probabilities than housing, highlighting the disparity in the media coverage of different infrastructure-related issues in Greece. To be more precise, the likelihood of perceiving 'sanitation' is 8.1 times lower (MLE+KS) and 3.7 times lower (Bootstrapping) than that of perceiving 'housing'. The probability for "Economy" is 14.0 times (MLE+KS) and 5.1 times (Bootstrapping) lower. The probability for "NGOs" is 21.4 times (MLE+KS) and 6.7 times (Bootstrapping) lower. To sum up, the higher the value of "Ratio", the less often this factor is perceived compared to "Housing". Furthermore, the (MLE+KS) scores are consistently higher than the Bootstrapping scores, suggesting that all factors are perceived less frequently when using MLE+KS as opposed to Bootstrapping.

Overall, Colombia and Greece reveal significant differences in the prominence and distribution of how topics related to infrastructure are reported in the media. In Colombia, healthcare and housing are the most reported issues, while in Greece, housing dominates the media coverage. The probabilities and ratios further emphasize the stark differences in how these issues are perceived and prioritized in both countries. Table 3 summarizes the results of blaming towards infrastructure for both Colombia and Greece.

Table 3. Distribution of frequencies and proportions for Blaming towards Infrastructure (F) for Colombia (based on 222 observations) and Greece (based on 499 observations)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | L = Colombia (\*) | | | | | L = Greece (\*\*) | | | | | | | | |
| r | Factor f (locals’ perceptions of…) |  | % | Ratio (Prob.) | | | Factor f (locals’ perceptions of …) | |  | | % | Ratio (Prob.) | | |
| KS | Boos. | | KS | Boos. | |
| 1 | Healthcare | 59 | 26.6 | 1.0 (0.616) | 1.0 (0.613) | | | Housing | | 121 | 24.2 | 1.0 (0.663) | | 1.0 (0.486) |
| 2 | Housing | 48 | 21.6 | 3.5 (0.175) | 3.4 (0.182) | | | Environment | | 82 | 16.4 | 3.7 (0.177) | | 2.3 (0.214) |
| 3 | Economics | 46 | 20.7 | 7.3 (0.084) | 6.9 (0.089) | | | Sanitation | | 80 | 16.0 | 8.1 (0.082) | | 3.7 (0.133) |
| 4 | Other Locals on Infrastructure | 30 | 13.5 | 12.4 (0.050) | 11.4 (0.054) | | | Economics | | 79 | 15.8 | 14.0 (0.047) | | 5.1 (0.094) |
| 5 | Sanitation | 19 | 8.6 | 18.6 (0.033) | 16.9 (0.036) | | | NGOs | | 71 | 14.2 | 21.4 (0.031) | | 6.7 (0.073) |
| 6 | Environment | 11 | 5.0 | 25.9 (0.024) | 23.2 (0.026) | | | NA | |  |  |  | |  |
| 7 | Education | 6 | 2.7 | 34.2 (0.018) | NA | | | NA | |  |  |  | |  |

**NOTE:** Bootstrapping (Boots.) based on 1000 iterations; KS=MLE + KS method; Ratio = = probability 1st rank/prob. in rank r. In brackets appear the probabilities Prob.= .

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Country** | **Method** |  |  |  |  |  |  |  |  |
| (\*) | MLE+KS: | 1.70 | 6.0 | 90.61 | -1.083 | 1.815 | 0.616 | 0.004 | 12.2 |
| (\*) | Boots.: | 2.07 | 8.5 | 77.50 | -0.848 | 1.756 | 0.613 | 0.007 | 90.7 |
| (\*\*) | MLE+KS: | 6.33 | 71.0 | 113.45 | -0.301 | 1.904 | 0.663 | 0.014 | 217526018935.7 |
| (\*\*) | Boots.: | 3.93 | 35.4 | 113.45 | -0.301 | 1.182 | 0.486 | 0.026 | 6340811.9 |

### Blaming towards Migrants

The most frequent factor in Colombia is the *locals’ perception of crime*, accounting for 31.2% of the reports with the highest probability of 0.784 (MLE+KS) y 0.928 (Bootstrapping). This indicates a significant concern about crime associated with migrants. The second most frequent factor is *the perception of migrant arrivals*, making up 20.4% of the reports with probability of 0.130 (MLE+KS) and 0.057 (Bootstrapping), and odds ratios of. 6.0 (MLE+KS) and 16.1 (Bootstrapping) (Boots.). Other factors have significantly greater odds ratios: *Migrants' presence* at 11.2% with odds ratios of 17.3 (MLE+KS) and 82.2 (Bootstrapping); *Disease spread* at 10.0% with odds ratios of 36.6 (MLE+KS) and 260.8 2 (Bootstrapping); *Fear or threat* and *Prostituion* at 9.2% with an odds ratio of 65.3 (MLE+KS). With the MLE method we found 6 more relevant factors, but with the Bootstrapping method we found only 4. These factors highlight the main concerns among the Colombian population regarding migrants. The Bootstrapping method generally shows higher odds ratios, indicating a greater disparity in perception compared to the MLE+KS method.

In Greece, the *locals’ perception of crime* is also the most frequent factor, comprising 44.7% of the reports with a high probability of 0.808 (MLE+KS) and 0.920 (Bootstrapping). The second most common factor is the perception of *fear or threat* from migrants, accounting for 21.6% of the reports with a probability of 0.123 (MLE+KS) and 0.063 (Bootstrapping). It has odds ratios of 6.6 (MLE+KS) and 14.7 (Bootstrapping). Other factors have lower probabilities: *Migrant presence* reported by 11.2%, odds ratios 19.9 (MLE+KS) and 70.9 (Bootstrapping); odds ratios of 43.4 (MLE+KS) and 216.2 (bootstrapping) for *Arrivals* with 10.1% reporting. Finally, *Disease spread* with a reporting rate of 7.9% and odds ratio of 79.7 (MLE+KS). Similarly, the Bootstrapping method shows higher odds ratios compared to the MLE+KS method. These results show that, like Colombia, crime is the most significant concern.

Overall, both Colombia and Greece identify crime as the most significant factor related to migrants. However, the distribution of other factors varies between the two countries. In Colombia, migrant arrivals and the spread of disease are more frequently reported, whereas in Greece, the fear or threat perception and the presence of migrants are more prominent. This comparison indicates that while crime is a universal concern, other issues such as fear, threat, and disease spread vary in importance between the two countries. Table 4 summarizes the results for blaming towards migrants from both Colombia and Greece.

Table 4. Distribution of frequencies and proportions for Blaming towards Migrants (F) for Colombia (based on 260 observations) and Greece (based on 770 observations)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | L = Colombia (\*) | | | | | | L = Greece (\*\*) | | | | | | | | |
| r | Factor f (locals’ perceptions of…) |  | % | Ratio (Prob.) | | | Factor f (locals’ perceptions of…) | |  | | % | Ratio (Prob.) | | | |
| KS | Boos. | | KS | | Boots. | |
| 1 | Crime | 81 | 31.2 | 1.0 (0.784) | | 1.0 (0.928) | | Crime | | 344 | 44.7 | | 1.0 (0.808) | | 1.0 (0.920) | |
| 2 | Arrivals | 53 | 20.4 | 6.0 (0.130) | | 16.1 (0.057) | | Fear or Threat | | 166 | 21.6 | | 6.6 (0.123) | | 14.7 (0.063) | |
| 3 | Migrants’ presence | 29 | 11.2 | 17.3 (0.045) | | 82.2 (0.011) | | Migrants’ presence | | 86 | 11.2 | | 19.9 (0.041) | | 70.9 (0.013) | |
| 4 | Disease spread | 26 | 10.0 | 36.6 (0.021) | | 260.8 (0.004) | | Arrivals | | 78 | 10.1 | | 43.4 (0.019) | | 216.2 (0.004) | |
| 5 | Fear or Threat | 24 | 9.2 | 65.3 (0.012) | | NA | | Disease spread | | 61 | 7.9 | | 79.7 (0.010) | | NA | |
| 5 | Prostitution | 24 | 9.2 | 65.3 (0.012) | | NA | | NA | |  |  | |  | |  |

**NOTE:** Bootstrapping (Boots.) based on 1000 iterations; KS=MLE + KS method; Ratio = = probability 1st rank/prob. in rank r. In brackets appear the probabilities Prob.= .

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Country** | **Method** |  |  |  |  |  |  |  |  |
| (\*) | MLE+KS: | 3.50 | 24.0 | 79.18 | -0.742 | 2.597 | 0.784 | 0.004 | 20097.5 |
| (\*) | Boots.: | 4.62 | 25.0 | 84.58 | -0.869 | 4.013 | 0.928 | 0.002 | 2092517.0 |
| (\*\*) | MLE+KS: | 2.49 | 61.0 | 337.72 | -1.094 | 2.720 | 0.808 | 0.006 | 13343.3 |
| (\*\*) | Boots.: | 3.43 | 68.2 | 343.52 | -1.131 | 3.878 | 0.920 | 0.003 | 1719677.1 |

### Blaming towards Government

In Colombia, *Policy-making* is the most frequently reported factor, accounting for 34.9% of occurrences, with an estimated probability of 0.772 (MLE+KS) and 0.895 (Bootstrapping). The odds ratios for other factors compared to *Policy-making* show significantly lower probabilities: *Governance of infrastructure* is at 19.9% of occurrences with odds ratios of 5.8 (MLE+KS) and 11.3 (Bootstrapping), *Border control* is at 11.0% of ocurrences with odds ratios of 16.0 (MLE+KS) and 46.9 (Bootstrapping), and *Rhetoric* is at 11.0% with the same odds ratios of 16.0 (MLE+KS) and 46.9 (Bootstrapping). *Funding* and *other Locals' views about government* are both at 8.2% with odds ratios of 33.1 (MLE+KS). The Bootstrapping method shows higher odds ratios, indicating a greater disparity in perception compared to the MLE+KS method.

In Greece, *Governance of infrastructure* is perceived as the most significant factor by locals, with a proportion of 40.5% of reports. Other factors have lower probabilities: *Policy-making* is at 27.7% of occurrences with odds ratios of 7.1 (MLE+KS) and 10.3 (Bootstrapping); *Migratory flows* are at 10.0% with odds ratios of 22.3 (MLE+KS) and 40.0 (Bootstrapping), and *Rhetoric* is at 5.6% with odds ratios of 50.2 (MLE+KS) and 105.2 (Bootstrapping). *Law enforcement* is at 5.3% with odds ratios of 94.2 (MLE+KS) and 222.5 (Bootstrapping), *Transparency* is at 3.3% with odds ratios of 157.6 (MLE+KS), *Funding* is at 3.0% with odds ratios of 243.6 (MLE+KS), and *Both border control* and *Political corruption* are at 2.3% with odds ratios of 355.2 (MLE+KS).

In conclusion, the ratios indicate that the likelihood of the first factor compared to others is significantly higher in Greece than in Colombia. For instance, in Colombia, the ratio for *Policy-making* to the second factor is 5.8, whereas in Greece, the ratio for *Governance of infrastructure* to the second factor is 7.1. This suggests that while *Policy-making* and *Governance of infrastructure* are significant concerns in both countries, the concentration of reports on these issues is higher in Greece. Additionally, in Colombia, *Border control* is the third most frequent factor at 11% with a probability of 0.048 (MLE+KS), while in Greece, it is migratory flows at 10% with a probability of 0.036 (MLE+KS). *Rhetoric* is similarly perceived in both countries but with different ranks and ratios. Finally, funding issues in Colombia and law enforcement in Greece have lower frequencies and probabilities, indicating less focus on these factors than others. Table 5 shows the results of blaming towards government in Colombia and Greece

Table 5. Distribution of frequencies and proportions for Blaming towards Government (F) for Colombia (based on 292 observations) and Greece (based on 334 observations)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | L = Colombia (\*) | | | | | | L = Greece (\*\*) | | | | | | | | | |
| r | Factor f (locals’ perceptions of…) |  | % | Ratio (prob.) | | | | Factor f (locals’ perceptions of…) |  | | | % | | Ratio (prob.) | |
| KS | Boots. | | | KS | Boots. | | |
| 1 | Policy-making | 102 | 34.9 | 1.0 (0.772) | | 1.0 (0.895) | | Governance of infrastructure | | 297 | 40.5 | | 1.0 (0.812) | | 1.0 (0.880) | | | |
| 2 | Governance of infrastructure | 58 | 19.9 | 5.8 (0.134) | | 11.3 (0.079) | | Policy-making | | 203 | 27.7 | | 7.1 (0.115) | | 10.3 (0.086) | | | |
| 3 | Border Control | 32 | 11.0 | 16.0 (0.048) | | 46.9 (0.019) | | Migratory Flows | | 73 | 10.0 | | 22.3 (0.036) | | 40.0 (0.022) | | | |
| 4 | Rhetoric | 32 | 11.0 | 16.0 (0.048) | | 46.9 (0.019) | | Rhetoric | | 41 | 5.6 | | 50.2 (0.016) | | 105.2 (0.008) | | | |
| 5 | Funding | 24 | 8.2 | 33.1 (0.023) | | NA | | Law Enforcement | | 39 | 5.3 | | 94.2 (0.009) | | 222.5 (0.004) | | | |
| 6 | Other Locals about Govt | 24 | 8.2 | 33.1 (0.023) | | NA | | Transparency | | 24 | 3.3 | | 157.6 (0.005) | | NA | | | |
| 7 | NA |  |  |  |  | | | Funding | | 22 | 3.0 | | 243.6 (0.003) | | NA | | | |
| 8 | NA |  |  |  |  | | | Border Control | | 17 | 2.3 | | 355.2 (0.002) | | NA | | | |
| 8 | NA |  |  |  |  | | | Political Corruption | | 17 | 2.3 | | 355.2 (0.002) | | NA | | | |

**NOTE:** Bootstrapping (Boots.) based on 1000 iterations; KS=MLE + KS method; Ratio = = probability 1st rank/prob. in rank r. In brackets appear the probabilities Prob.= .

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Country** | **Method** |  |  |  |  |  |  |  |  |
| (\*) | MLE+KS: | 2.98 | 24.0 | 98.54 | -0.847 | 2.524 | 0.772 | 0.004 | 4435.9 |
| (\*) | Boots.: | 3.87 | 24.4 | 101.84 | -0.905 | 3.503 | 0.895 | 0.005 | 316494.5 |
| (\*\*) | MLE+KS: | 1.96 | 17.0 | 363.28 | -1.441 | 2.824 | 0.812 | 0.001 | 72.6 |
| (\*\*) | Boots.: | 2.39 | 24.2 | 357.94 | -1.405 | 3.358 | 0.880 | 0.002 | 2969.1 |

### Blaming towards Geopolitics

In Colombia, the most frequently reported factor is *locals’ perception of Migratory flows*, accounting for 20.2% of observations with an estimated probability of 0.615 (MLE+KS) and 0.572 (Bootstrapping). This is followed by *locals’ perception of border control or closures*, representing 19.3% with a probability of 0.190 (MLE+KS) and 0.200 (Bootstrapping) and odds ratios of 3.2 (MLE+KS) and 2.9 (Bootstrapping). The odds ratios for other factors compared to Migratory flows show significantly lower probabilities: *Economics* is at 14.7% of occurrences with odds ratios of 6.4 (MLE+KS) and 5.3 (Bootstrapping); and *Crime perpetuated by foreign nationals* is at 12.8% of the reports with odds ratios of 10.5 (MLE+KS) and 8.2 (Bootstrapping). *Policy-making* is present in 11.9% of the events, with odds ratios of 15.3 (MLE+KS) and 11.5 (Bootstrapping). The Bootstrapping method shows higher odds ratios, indicating a greater disparity in perception compared to the MLE+KS method.

In Greece, *Migratory flows* are also perceived as the most significant factor in blaming geopolitics, with a proportion of 26.9% and a probability of 0.686 (MLE+KS) and 0.731 (Bootstrapping). The second most common factor is *Locals’ perception of policy-making*, which accounts for 23.1% of the observations and has a probability of 0.156 (MLE+KS) and 0.146 (Bootstrapping). Other factors have lower probabilities: *Crime perpetrated by foreigners* is reported in 22.2% of cases, with odds ratios of 10.4 (MLE+KS) and 12.8 (Bootstrapping), and *Rhetoric* is reported in 11.7% of cases, with odds ratios of 19.2 (MLE+KS) and 25.0 (bootstrapping). *Border control or closure* occurs in 4.8% of reports with odds ratios of 30.9 (MLE+KS) and 42.0 (Boostrapping), economics in 4.2% of reports with odds ratios of 45.5 (MLE+KS) and 64.2 (Boostrapping), funding for humanitarian response at 3.9% with odds ratios of 63.3 (MLE+KS) and 91.8 (Bootstrapping), and UN agencies at 2.1% with an odds ratio of 84.1 (MLE+KS).

The comparison reveals that both countries share similar concerns regarding *Migratory flows*, although the proportion is higher in Greece (26.9%) compared to Colombia (20.2%). This factor has the highest probability in both countries, indicating its significant presence in public discourse. However, the ratios of probabilities between the most cited factor and others vary greatly. In Colombia, the probability of *Migratory flows* is substantially higher than other factors, with a ratio of 27.2 times that of the second factor. Although the factor about *Migratory flows* has the highest probability in Greece, the disparity with other factors is less pronounced, with a ratio of 3.2 to the second factor, about *Policy-making*. Overall, the data suggests that while both countries prioritize concerns about *Migratory flows*, the focus and intensity of other geopolitical issues vary significantly, reflecting different national contexts and public sentiments. Table 6 provides the results for geopolitics in Colombia and Greece.

Table 6. Distributions of frequencies and proportions for Blaming towards Geopolitics (F) for Colombia (based on 580 observations) and Greece (based on 733 observations)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | L = Colombia (\*) | | | | | L = Greece (\*\*) | | | | |
| r | Factor f (locals’ perceptions of…) |  | % | Ratio | | Factor f (locals’ perceptions of…) |  | % | Ratio | |
| KS | Boos. | KS | Boots. |
| 1 | Migratory Flows | 117 | 20.2 | 1.0 (0.615) | 1.0 (0.572) | Migratory Flows | 90 | 26.9 | 1.0 (0.686) | 1.0 (0.731) |
| 2 | Border Control or Closures | 112 | 19.3 | 3.2 (0.190) | 2.9 (0.200) | Policy-making | 77 | 23.1 | 4.4 (0.156) | 5.0 (0.146) |
| 3 | Economics | 85 | 14.7 | 6.4 (0.096) | 5.3 (0.108) | Crime perpetuated by foreign nationals | 74 | 22.2 | 10.4 (0.066) | 12.8 (0.057) |
| 4 | Crime perpetuated by foreign nationals | 74 | 12.8 | 10.5 (0.059) | 8.2 (0.070) | Rhetoric | 39 | 11.7 | 19.2 (0.036) | 25.0 (0.029) |
| 5 | Policy-making | 69 | 11.9 | 15.3 (0.040) | 11.5 (0.050) | Border Control or Closures | 16 | 4.8 | 30.9 (0.022) | 42.0 (0.017) |
| 6 | NA |  |  |  |  | Economics | 13 | 4.2 | 45.5 (0.015) | 64.2 (0.011) |
| 7 | NA |  |  |  |  | Funding for humanitarian response | 14 | 3.9 | 63.3 (0.011) | 91.8 (0.008) |
| 8 | NA |  |  |  |  | UN Agency Actors | 7 | 2.1 | 84.1 (0.008) | NA |

**NOTE:** Bootstrapping (Boots.) based on 1000 iterations; KS=MLE + KS method; Ratio = = probability 1st rank/prob. in rank r. In brackets appear the probabilities Prob.= .

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Country** | **Method** |  |  |  |  |  |  |  |  |
| (\*) | MLE+KS: | 4.77 | 69.0 | 125.55 | -0.355 | 1.695 | 0.615 | 0.016 | 258416665.4 |
| (\*) | Boots.: | 4.27 | 48.7 | 125.55 | -0.355 | 1.518 | 0.572 | 0.019 | 29831133.5 |
| (\*\*) | MLE+KS: | 1.68 | 7.0 | 151.77 | -1.268 | 2.131 | 0.686 | 0.002 | 13.1 |
| (\*\*) | Boots.: | 2.05 | 10.15 | 137.16 | -1.133 | 2.323 | 0.731 | 0.002 | 71.4 |

### Graphical summary of the most important results

The figure 4 illustrates the relationship between rank (r) and odds ratio for different types of frustration, methods of estimation, and countries. The plot is organized into a 2x2 grid of charts. Each row represents a different method (Bootstrapping and MLE-KS), and each column represents a different country (Colombia and Greece). The x-axis represents the rank (r), while the y-axis represents the odds (ratio). Each colored line corresponds to a different type of frustration, as indicated by the legend on the right.

**PENDIENTE MEJORAR LA GRAFICA**

A graph of different colored lines

Description automatically generated

Figure 4. Distribution of the odds ratio for different values of the parameter , by type of frustration, methods of estimation, and countries.

In both Colombia and Greece, the bootstrapping method shows a notable increase in odds ratio with higher ranks. Specifically, the Migrant type (purple line) shows a steep increase in odds ratio in both countries. Geopolitics (red line) and Infrastructure (blue line) show a more gradual increase in odds ratio. Similar to Bootstrapping, the MLE-KS method in both countries shows increasing odds ratio with higher ranks. The Government type (green line) exhibits the steepest increase in both countries, especially in Greece. For both methods, in Colombia, the Migrant type shows a significant increase in odds ratio with higher ranks. This trend is more pronounced in the Bootstrapping method. The Government type in the Greek data shows the highest increase in odds ratio with higher ranks for both methods, indicating a strong relationship between rank and odds in governmental issues. With respect to type-specific trends, Geopolitics shows moderate growth in odds ratio with higher ranks in both countries and methods. Government has the steepest increase in the Greece MLE-KS method. Infrastructure, generally, exhibits a gradual increase in odds ratio across all plots. Migrant, hows a steep increase, particularly in the Bootstrapping method for both countries.

1. Conclusions

ALGUNAS IDEAS GENERALES QUE PUEDEN SERVIR COMO BASE.

HAY QUE LEER Y REVISAR CON CUIDADO

"We have developed the Relative Importance Factor (RIF) Index, which compares the probability of a factor occupying the first position with the probability of it being in any other position within the range. This innovation provides a new perspective for evaluating the relative prominence and importance of factors, an aspect that has not been considered in previous studies."

By using $\widehat{RPC}(r)$, researchers can gain deeper insights into the dynamics of factor importance and prominence, enhancing the analysis of complex datasets.

Este estudio se basa en el análisis cuantitativo de la distribución de citaciones mediante técnicas de minería de datos y análisis de contenido. Utilizamos modelos estadísticos para estimar los parámetros de la ley de potencias y evaluar la bondad de ajuste del modelo [Referencia 18, Referencia 19]. Los resultados de este análisis pueden proporcionar una comprensión más profunda de los patrones de citaciones en el ámbito de la resiliencia social y contribuir a la literatura existente sobre leyes de potencias y distribución de la atención académica [Referencia 20].

Además, este trabajo explora las implicaciones de estos patrones de citaciones para la evaluación del impacto académico y la formulación de políticas científicas. Al identificar los factores más frecuentemente citados y su distribución, podemos ofrecer recomendaciones para mejorar la visibilidad y el impacto de la investigación en resiliencia social [Referencia 21, Referencia 22].

Finalmente, comparamos los resultados obtenidos con estudios previos en otras áreas de investigación para identificar similitudes y diferencias en la distribución de citaciones y la dinámica de la atención académica. Este enfoque comparativo permite contextualizar los hallazgos y ofrecer una perspectiva más amplia sobre la aplicación del modelo de ley de potencias en diferentes disciplinas [Referencia 23, Referencia 24].

Los hallazgos pueden tener importantes implicaciones para la evaluación del impacto académico y la formulación de políticas científicas [Referencia 25].

**FUTURO TRABAJOS**

1. Acknowledgment

This research is funded by grant number N000141912624 by the Office of Naval Research through the Minerva Research Initiative and grant number P116S210003 by the US Department of Education. Liss Romero and Lia Castillo (Universidad del Norte, Colombia) and Lydia Cleveland and Madison Gonzalez (VMASC) created the codebooks and qualitatively coded the news articles.

1. References

Afzal, N., Seena, A., & Bushra, A. (2023). Friction factor power law with equivalent log law, of a turbulent fully developed flow, in a fully smooth pipe. *Zeitschrift für angewandte Mathematik und Physik*, *74*(4), 144.

Arroyo-Machado, W., Torres-Salinas, D., Herrera-Viedma, E., & Romero-Frías, E. (2020). Science through Wikipedia: A novel representation of open knowledge through co-citation networks. *PloS one*, *15*(2), e0228713.

Arthur, R., & Williams, H. T. (2019). Scaling laws in geo-located Twitter data. *PloS one*, *14*(7), e0218454.

Artico, I., Smolyarenko, I., Vinciotti, V., & Wit, E. C. (2020). How rare are power-law networks really?. Proceedings of the Royal Society A, 476(2241), 20190742.

Atkins, J. W., Walter, J. A., Stovall, A. E., Fahey, R. T., & Gough, C. M. (2022). Power law scaling relationships link canopy structural complexity and height across forest types. *Functional Ecology*, *36*(3), 713-726.

Ben Yaala, S., & Henchiri, J. E. (2023). Predicting stock market crashes on the African stock markets: evidence from log-periodic power law model. *African Journal of Economic and Management Studies*.

Benatti, A., de Arruda, H. F., Silva, F. N., Comin, C. H., & da Fontoura Costa, L. (2023). On the stability of citation networks. *Physica A: Statistical Mechanics and its Applications*, *610*, 128399.

Banshal, S. K., Gupta, S., Lathabai, H. H., & Singh, V. K. (2022). Power Laws in altmetrics: An empirical analysis. *Journal of Informetrics*, *16*(3), 101309.

Baumgartner, M. T., & Peláez Zapata, O. E. (2024). Taylor's power law for freshwater fishes: Functional traits beyond statistical inevitability. *Journal of Animal Ecology*.

Blasius, B. (2020). Power-law distribution in the number of confirmed COVID-19 cases. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *30*(9).

Cardoso, B. H. F., Gonçalves, S., & Iglesias, J. R. (2020). Wealth distribution models with regulations: Dynamics and equilibria. *Physica A: Statistical Mechanics and its Applications*, *551*, 124201.

Corral, Á., & González, Á. (2019). Power law size distributions in geoscience revisited. *Earth and Space Science*, *6*(5), 673-697.

Chen, B., Ma, C., Krajewski, W. F., Wang, P., & Ren, F. (2020). Logarithmic transformation and peak-discharge power-law analysis. Hydrology Research, 51(1), 65-76.

Chen, X., Pan, Y., & Luo, B. (2021). Research on power-law distribution of long-tail data and its application to tourism recommendation. Industrial Management & Data Systems, 121(6), 1268-1286.

Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, *51*(4), 661-703.

Dagdeviren, H., Capucha, L., Calado, A., Donoghue, M., & Estêvão, P. (2020). Structural foundations of social resilience. Social Policy and Society, 19(4), 539-552.

Devlin, P., Kepner, J., Luo, A., & Meger, E. (2021, June). Hybrid power-law models of network traffic. In *2021 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)* (pp. 280-287). IEEE.

Dufrénot, G., & Paret, A. C. (2019). Power-law distribution in the external debt-to-fiscal revenue ratios: Empirical evidence and a theoretical model. *Journal of Macroeconomics*, *60*, 341-359.

Dong, X., Tang, D., & Tang, C. (2021). Social network analysis in China’s hospital healthcare. *Physica A: Statistical Mechanics and its Applications*, *565*, 125546.

Fischer, P., & Schweikhard, L. (2020). Decay-rate power-law exponent as a link between dissociation energy and temperature. Physical Review Research, 2(4), 043177.

Ghosh, B., Kenourgios, D., Francis, A., & Bhattacharyya, S. (2021). How well the log periodic power law works in an emerging stock market?. *Applied Economics Letters*, *28*(14), 1174-1180.

Gillespie, C. S. (2014). Fitting heavy tailed distributions: the poweRlaw package. *arXiv preprint arXiv:1407.3492*.

He, Z., Wang, Z., Gu, Y., & An, X. (2023). Measuring the Influence of Multiscale Geographic Space on the Heterogeneity of Crime Distribution. ISPRS International Journal of Geo-Information, 12(10), 437.

Jha, P. K. (2020). What does simple power law kinetics tell about our response to coronavirus pandemic?. medRxiv, 2020-04.

Jiang, B., & de Rijke, C. (2021). A power-law-based approach to mapping COVID-19 cases in the United States. *Geo-Spatial Information Science*, *24*(3), 333–339. https://doi.org/10.1080/10095020.2020.1871306

Jiang, J.J., Yamada, K., Takayasu, H. et al. Scale-dependent power law properties in hashtag usage time series of Weibo. Sci Rep 13, 22298 (2023). <https://doi.org/10.1038/s41598-023-49572-6>

Kumar, R. (2024). Power-law Behaviour and Inequality in the Upper Tail of Wealth, Income and Consumption: Evidence from India. *The Indian Economic Journal*, 00194662241238471.

Li, L., & Ma, Z. (2019). Comparative power law analysis for the spatial heterogeneity scaling of the hot‐spring microbiomes. Molecular Ecology, 28(11), 2932-2943.

Masseran, N., Yee, L. H., Safari, M. A. M., & Ibrahim, K. (2019). Power law behavior and tail modeling on low income distribution. *Mathematics and Statistics*, *7*(3), 70-77.

Navas-Portella, V., González, Á., Serra, I., Vives, E., & Corral, Á. (2019). Universality of power-law exponents by means of maximum-likelihood estimation. *Physical Review E*, *100*(6), 062106.

Neipel, J., Bauermann, J., Bo, S., Harmon, T., & Jülicher, F. (2020). Power-law population heterogeneity governs epidemic waves. *PloS one*, *15*(10), e0239678.

Ng, Y. M., Ariel, B., & Harinam, V. (2023). Crime on the mass transit system in Hong Kong: a hotspots and harmspots trajectory approach. *Policing: An International Journal*, *46*(5/6), 908-921.

Pisarenko, V., & Rodkin, M. V. (2019). A Collection of Statistical Methods for Analysis of the Disaster Damages and the Seismic Regime. *Earthquake risk perception, communication and mitigation strategies across Europe, Il Sileno Edizioni, Geographies of the Anthropocene book series, Rende*, *2*(2), 43-63.

Popescu, I. I. (2003). On a Zipf's law extension to impact factors. Glottometrics, 6, 83-93.

Puttanapong, N., Martinez Jr, A., Bulan, J. A. N., Addawe, M., Durante, R. L., & Martillan, M. (2022). Predicting poverty using geospatial data in Thailand. ISPRS International Journal of Geo-Information, 11(5), 293.

Rajput, N. K., Grover, B. A., & Rathi, V. K. (2020). Word frequency and sentiment analysis of twitter messages during coronavirus pandemic. arXiv preprint arXiv:2004.03925.

Ramos, P. L., Jerez-Lillo, N., Segovia, F. A., Egbon, O. A., & Louzada, F. (2024). Power-law distribution in pieces: a semi-parametric approach with change point detection. *Statistics and Computing*, *34*(1), 16.

Safari, M. A. M., Masseran, N., Ibrahim, K., & AL-Dhurafi, N. A. (2020). The power-law distribution for the income of poor households. Physica A: Statistical Mechanics and its Applications, 557, 124893.

Sciacca, M., & Jou, D. (2024). A power-law model for nonlinear phonon hydrodynamics. *Zeitschrift für angewandte Mathematik und Physik*, *75*(2), 70.

Sethna, J. P. (2022). Power laws in physics. Nature Reviews Physics, 4(8), 501-503.

Sohn, H. G., Kim, Y. K., & Kwon, Y. (2023). Studying Disasters Through Complexity Theory. In *International Handbook of Disaster Research* (pp. 1077-1088). Singapore: Springer Nature Singapore.

Somin, S., Altshuler, Y., & Shmueli, E. (2022). Remaining popular: Power-law regularities in network dynamics. EPJ Data Science, 11(1), 61.

Srivastav, A. K., Ghosh, M., & Bandekar, S. R. (2021). Modeling of COVID-19 with limited public health resources: a comparative study of three most affected countries. The European Physical Journal Plus, 136, 1-26.

Taleghani, F., Shakibaie, A., Salehi, M., Jalaee, S. A., & Nejati, M. (2020). An analysis of the crisis transmission in the global financial network to Iran. Quarterly Journal of Quantitative Economics, 17(1), 155-183.

Thelwall, M., & Nevill, T. (2018). Could scientists use Altmetric. com scores to predict longer term citation counts?. Journal of informetrics, 12(1), 237-248.

Xu, X. T., Wang, N., Bian, J., & Zhou, B. (2019). Understanding the diversity on power-law-like degree distribution in social networks. *Physica A: Statistical Mechanics and its Applications*, *525*, 576-581.

Yum, S. (2023). Spatial response and power law distribution according to Winter storm Jonas. *Quality & Quantity*, *57*(6), 5241-5255.

Zhao, F., Zhang, Y., Lu, J. et al. Measuring academic influence using heterogeneous author-citation networks. Scientometrics 118, 1119–1140 (2019). https://doi.org/10.1007/s11192-019-03010-5

Zaghloul, G. (2019). A note on the zeros of generalized Hurwitz zeta functions. Journal of Number Theory, 196, 197-204.

1. has very different values due to the different weights used in the data. [↑](#footnote-ref-2)
2. Percentage (frequency ) accumulated data [↑](#footnote-ref-3)
3. Based on 1000 bootstrapping iterations. [↑](#footnote-ref-4)
4. We present only the factors f of the type of frustration F in the country L such that [↑](#footnote-ref-5)