Annali del Dipartimento di Metodi e Modelli per l'Economia il Territorio e la Finanza
2025, DOI: 10.13133/2611-6634/1865

E-ISSN 2611-6634
ISSN 2385-0825

Research paper

First published online: December 7, 2025

Marianna Belloc*

SIZE DISTRIBUTION OF ALL ITALIAN CITIES

Abstract

By revisiting well-established empirical regularities, this paper examines the distributional properties of city sizes in Italy over the period 2001–2024. Three main findings emerge. First, when considering the full population of approximately 7,900 municipalities, city sizes are well approximated by a lognormal distribution. Second, consistent with Gibrat's law, city population growth rates appear to be independent of initial city size. Third, the upper tail of the distribution follows Zipf's law, although the results are sensitive to the choice of the truncation threshold. The implications of these findings are discussed in relation to urban dynamics and the organization of the Italian urban system.

Keywords: City size distribution, Zipf's Law, Gibrat's Law, log-normality.

*Sapienza University of Rome, CEPR, and CESifo (email: marianna.belloc@uniroma1.it)

1 Introduction

According to a well-established line of research in urban geography, the distribution of city sizes exhibits two key empirical regularities: the upper tail follows a Pareto distribution with a shape parameter equal to one (Zipf's law; Zipf, 1949) and city growth is proportionate and independent of initial size (Gibrat's law; Gibrat, 1931). While these patterns have been largely—though not unanimously—confirmed by the literature (Gabaix, 2009; Gabaix and Ioannides, 2004), their empirical validity depends on the threshold used to define the upper tail, as shown by Eeckhout (2004). This sensitivity to the truncation threshold highlights the importance of examining the full distribution of city sizes. In this context, Eeckhout (2004) finds that the non-truncated distribution of US cities follows a log-normal pattern—a result that can coexist with evidence supporting both Zipf's and Gibrat's laws. In this paper, we empirically assess the validity of all three regularities—log-normality, proportionate growth, and the Pareto distribution of the upper tail—using data on approximately 7,900 Italian municipalities over the period 2001–2024.

Our work relates to the extensive literature on urban structure, city growth, and their economic implications. A rich body of research has examined the role of economic forces in shaping city structure and diffusion—such as increasing returns, trade and congestion externalities, and both market and non-market dynamics (Gabaix and Ioannides, 2004). The inherent tension between increasing returns to agglomeration at the local level and constant returns to scale at the aggregate level (Rossi-Hansberg and Wright, 2007) provides a theoretical foundation for both the emergence of cities and the empirical evidence of balanced growth. These dynamics are consistent with a substantial body of empirical work showing that city size data conform to two robust regularities: Zipf's law and Gibrat's law.

Zipf's law, in the context of urban economics, states that the city size distribution can be approximated by a Pareto distribution with a coefficient equal to one. Its empirical validity has been confirmed by numerous studies, including Ioannides and Overman (2003), Rosen and Resnick (1980), and Soo (2005) (see Gabaix, 2009 and Gabaix, 2016, for reviews). The robustness and apparent spatial invariance of Zipf's law highlight a remarkable regularity in the distribution of city sizes, offering predictive insights for policymakers and motivating a range of theoretical explanations over time (for a review, see Gabaix and Ioannides, 2004). Formally, this empirical regularity suggests that the probability that a city population is greater than x is proportional to $1/x$. This statement approximately implies that the size of the city of rank k is proportional to $1/k$, that is the population of the largest city is twice that of the second largest city and three times that of the third largest city.

While Zipf's law has received strong empirical support by previous research, several studies have documented systematic deviations from it (Gabaix and Ioannides, 2004; Eeckhout, 2004; Rossi-Hansberg and Wright, 2007). In particular, the log-rank versus log-size relationship often exhibits a local concavity in the left tail, indicating an underrepresentation of small cities relative to Zipf's prediction, and a local convexity in the right tail, suggesting the absence of extremely large cities compared to what Zipf's law would imply. These patterns have been observed across many countries (Soo, 2005), including the evidence presented here for Italy. A widely accepted explanation for Zipf's law is Gibrat's law of proportional city growth (Gabaix, 1999).

Building on the intuition of Champernowne (1953)—that random multiplicative growth processes can generate power-law distributions—Gabaix (2009) develops a formal model showing that, in a steady state with a fixed number of cities whose sizes evolve stochastically with common mean and variance, the resulting distribution converges to Zipf's law (with a power coefficient equal to one).

The present work contributes to the existing literature by exploring—to the best of our knowledge for the first time—the empirical validity of Zipf's and Gibrat's laws together with the log-normality of the city size distribution for Italian cities over a long time horizon, spanning from 2001 to 2024. As in many advanced economies, both population and economic activity in Italy are highly concentrated in urban areas. These dynamics exhibited a sustained growth trend between 1980 and 2010, but have largely stabilized over the past fifteen years, following distinct trajectories across cities of different sizes (Accetturo and Mocetti, 2019; andrea and andrea Petrella, 2019; Alaimo et al., 2023; Salvati et al., 2013).

Investigating empirical regularities and identifying distinctive characteristics in the distribution and growth of the urban population is relevant from a policy standpoint, given their implications for cost efficiency, quality of life, and investment in infrastructure. The long-term perspective adopted here allows us to isolate the stable properties of the city size distribution from more transient phenomena associated with urban dynamics, such as agglomeration gains and losses, innovation shocks, and migration flows (Duranton, 2007). A deeper understanding of these structural regularities is essential for designing effective incentive schemes and models of urban development.

The rest of this manuscript is organized as follows. Section 2 describes the data. Section 3 presents the methodologies and main findings. Section 4 concludes with some final remarks.

2 Data description

Data on the population size of each Italian municipality are provided by the Italian National Institute of Statistics (ISTAT). The municipality—the basic entity of local government in Italy—is used as the unit of analysis. In 2016, the traditional census, previously based on field data collection, was replaced by an integrated system of statistical registers that combines administrative sources with survey data. This methodological shift has enhanced data quality and ensured the granularity required for our study. Since 2018, the census has become permanent and is conducted annually (Italian Institute of Statistics, 2025a).

We collected data on approximately 7,900 municipalities for the period 2001–2024, although the exact number of municipalities varies slightly by year due to administrative mergers and consolidations. The dataset was retrieved from ISTAT's open data platform (Italian Institute of Statistics, 2025b). Table 1, presents descriptive statistics for the population size and the (natural) logarithm of population size at five-year intervals, as well as for 2024—the final year of our analysis. Data for 2025 are excluded, as only projections are currently available. Because of data availability, we are forced to adopt an administrative rather than a functional definition of cities. Such an aggregation might imply a Modifiable Areal Unit Problem, which is a potential source of errors due to scale and zonal effects (Wong, 2009).

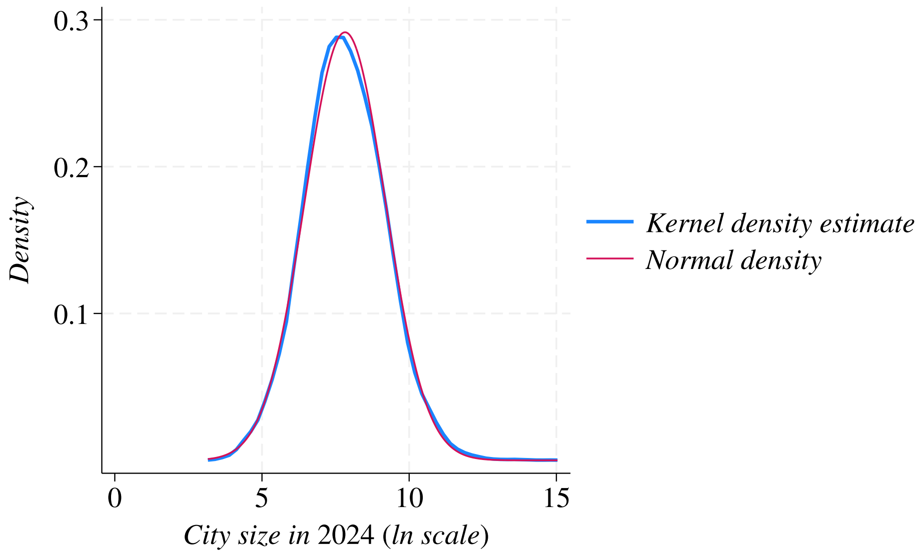
Table 1. Data description

Year:	2001	2006	2011	2016	2021	2024
Panel A: Summary statistics of city population size						
Obs	7,912	7,913	7,913	7,913	7,902	7,899
Mean (ln)	7,204	7,366	7,576	7,603	7,491	7,460
	(7.841)	(7.862)	(7.877)	(7.859)	(7.823)	(7.813)
St. dev. (ln)	39,787	39,927	40,730	42,226	41,868	41,606
	(1.300)	(1.310)	(1.332)	(1.350)	(1.368)	(1.376)
Panel B: Percentiles						
10th percentile (ln)	512	513	512	490	462	451
	(6.238)	(6.240)	(6.238)	(6.194)	(6.138)	(6.111)
50th percentile (ln)	2,431	2,491	2,539	2,498	2,411	2,391
	(7.796)	(7.820)	(7.840)	(7.823)	(7.789)	(7.779)
90th percentile (ln)	13,057	13,493	14,187	14,244	14,074	14,162
	(9.477)	(9.510)	(9.560)	(9.564)	(9.552)	(9.558)
Panel B: Rome-municipality i population ratio						
Milan	2.027	2.180	2.108	2.049	2.002	2.006
Naples	2.535	2.812	2.759	2.883	2.984	3.012
Torino	2.943	3.143	2.986	3.142	3.206	3.322
Palermo	3.709	4.088	4.026	4.202	4.314	4.365
Genova	4.173	4.540	4.455	4.777	4.858	4.893
Bologna	6.861	7.489	7.112	7.158	7.025	7.054
Florence	7.152	7.720	7.266	7.340	7.469	7.589
Bari	8.046	8.615	8.340	8.705	8.675	8.702
Catania	8.134	9.001	8.956	9.273	9.162	9.213

Note: The table presents descriptive statistics of city population size computed at five-year intervals and for 2024. Panel A summarizes the corresponding number of observations, mean and standard deviation, with their natural log in brackets. Panel B reports the 10th, 50th, and 90th percentiles. Panel C displays the ratio between the population of Rome and that of each of the subsequent nine medium-to-large cities in the distribution.

Panel A reports the mean, the standard deviation, and the number of observations at five year intervals between 2001 and 2021, as well as for 2024. As shown, the number of Italian municipalities has decreased over the past two periods, largely due to the consolidation of very small municipalities that have experienced population decline. Panel B presents the 10th, the 50th, and the 90th percentiles. It reveals that, on the left tail of the distribution, 10% of the municipalities is very small, registering a population below 500 units; 25% of the cities is below 1,000 inhabitants (not reported in the table), and 50% is below 2,500. On the right tail of the distribution, there are few very large cities (Rome, Milan, Naples) and a number of medium-large ones (Torino, Palermo, Genova, Bologna,

Figure 1. Density function of city size



Note: The graph plots the city size density function and the theoretical normal density for 2024.

Florence, Bari, and Catania). Rome is the largest, followed by Milan and Naples: in 2024, their populations were 2,751,747, 1,371,499, and 913,704, respectively. Panel C reports the ratios between the population of Rome and those of other major Italian cities, measured at five-year intervals and for 2024. Denoting by s_i the ratio between the population of Rome and the population of city i, we observe a striking regularity—previously documented in other countries as well (Gabaix, 2009): $s_{Milan} \approx 2$ and $s_{Naples} \approx 3$. In other words, the population of the largest city is approximately inversely proportional to its rank relative to the second- and third-largest cities, a pattern consistent with Zipf's law. This regularity, as shown in the table, also roughly extends to the next seven medium-to-large cities in the distribution.

3 Methods and results

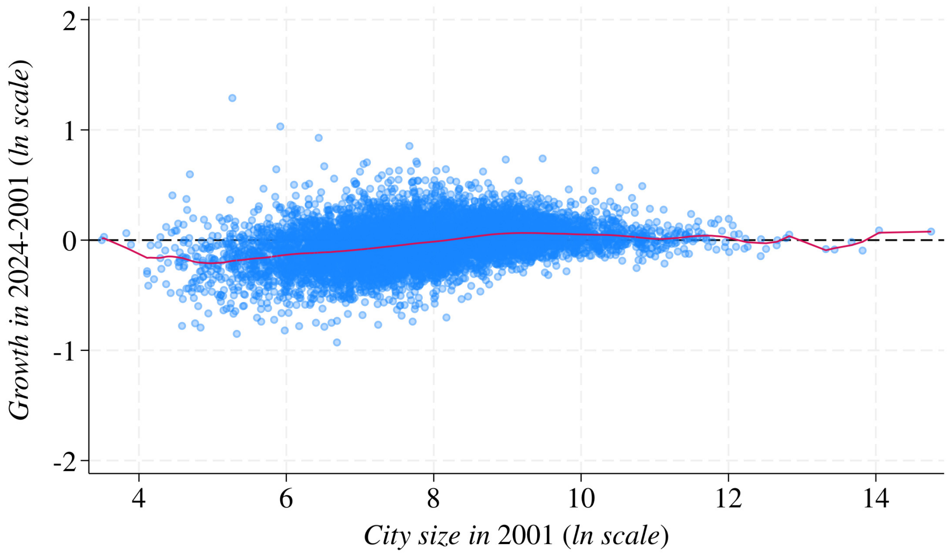
3.1 Log-normality

In this subsection, we examine the properties of the entire size distribution of Italian municipalities. Specifically, relying on Eeckhout (2004), we investigate whether the distribution follows a log-normal pattern, that is:

$$\phi(\hat{\mu}, \hat{\sigma}) = \frac{1}{\hat{\sigma}\sqrt{2\pi}} e^{[\ln(Size_i) - \hat{\mu}]^2 / 2\hat{\sigma}^2}, \tag{1}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are the mean and the standard deviation of $\ln(Size)$, respectively. Figure 1 compares the empirical distribution of municipality sizes in 2024 with the corresponding

Figure 2. Growth rate, city size, and kernel conditional mean estimate



Note: The graph plots the (natural) log of growth rates (2024-2001) against the (natural) log of initial city size (2001), and the kernel conditional mean estimate.

theoretical log-normal distribution, showing a very close approximation. Results for other years are visually similar and are reported in Appendix A.1. The goodness of fit is further supported by a Kolmogorov–Smirnov (K–S) test following the approach of Berger and Zhou (2014). Using the sample mean $\hat{\mu}$=7.813 and standard deviation $\hat{\sigma}$=1.376 of the log-transformed city sizes in 2024, the test yields a D statistic of 0.013 and a p-value of 0.141. These results indicate that we cannot reject the null hypothesis of log-normality at conventional significance levels, supporting the hypothesis that the full distribution of city sizes follows a log-normal pattern.

3.2 Proportionate growth

Gibrat's law posits that the growth rate of an economic unit is independent of its initial size. A substantial body of empirical research has shown that this relationship holds not only for firms but also for cities (Gabaix, 2009). In this section, we test the applicability of Gibrat's law to Italian municipalities by computing their population growth rates between 2001 and 2024.

Figure 2 plots the distribution of growth rates against initial city size. The result provides visual evidence on the independence between size and growth. With two exceptions, the logarithmic growth rate remains within the range of (–1, +1). On average, smaller municipalities—particularly in Southern Italy—tend to experience population decline over time, while medium-sized cities generally exhibit positive growth. The largest cities (Rome, Milan, and Naples) show growth rates that are close to zero. Overall, the scatter plot forms

Table 2. Estimated Pareto coefficients

Truncation threshold:	250	1,000	2,500	5,000	7,000
$\hat{\alpha}$	-1.426***	-1.398***	-1.215***	-1.019***	-0.822***
	(0.010)	(0.003)	(0.003)	(0.003)	(0.003)
$\ln(\hat{\beta})$	20.36***	20.05***	18.18***	16.30***	14.52***
	(0.114)	(0.029)	(0.024)	(0.025)	(0.024)
Obs	250	1,000	2,500	5,000	7,000
R^2	0.987	0.996	0.987	0.963	0.920
Cutoff size	34,478	11,803	3,669	1,493	495

Note: The table reports OLS estimated coefficients from equation (4) for 2024 across the five different population thresholds: 250, 1,000, 2,500, 5,000, and 7,000.

a horizontally compressed cloud, visually consistent with the prediction of Gibrat's law, which implies no systematic relationship between city size and growth.

To examine this relationship more formally, we follow the methodology of Eeckhout (2004) and compute a non-parametric kernel estimate of the conditional mean of (natural log) population growth using the Nadaraya–Watson estimator (Nadaraya, 1965; Watson, 1964). In line with the recommendations of Cattaneo and Jansson (2018), we employ the Epanechnikov kernel and apply the optimal bandwidth (0.223), with 500 bootstrap replications for inference. The resulting smooth curve, also shown in Figure 2, indicates that the estimated conditional mean remains relatively flat across the distribution of city sizes—further supporting Gibrat's law.

In Appendix A.2, we replicate this analysis using five-year intervals. The scatter plots appear even more compressed over shorter-time horizons, providing stronger visual evidence of proportionate growth. This suggests that the result is not sensitive to the specific time interval chosen, and that Gibrat's law holds over both the short and long run for Italian municipalities.

3.3 Zipf's law

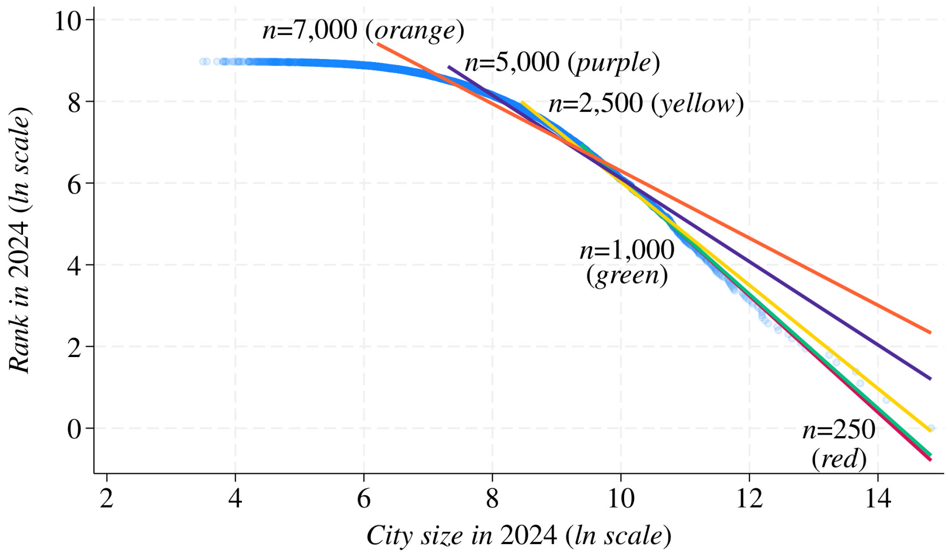
Zipf's law for city size states that the upper tail of the city size distribution follows a power law with a shape parameter equal to one. Formally, the probability that a city has population size greater than or equal to x is given by:

$$P(Size_i \geqslant x) \propto x^{-\alpha} \tag{2}$$

where α (the shape parameter) is equal to one under Zipf's law, greater (smaller) values being associated with more (less) equality in the distribution. This implies that the distribution of the largest cities is scale-invariant, and the rank of a city is inversely proportional to its size, that is:

$$Size_i = \frac{\beta}{Rank_i^{\alpha}}, \text{ with } \alpha = 1. \tag{3}$$

Figure 3. Rank versus city size



Note: The graph plots the (natural) log of city rank against the (natural) log of city size with linear regression lines for 2024 across five different truncation thresholds.

Gabaix (1999) demonstrates that if the city size distribution evolves according to a stochastic process consistent with Gibrat's law, it will converge in the steady state to a power law with a shape parameter of one—that is, Zipf's law. Empirical studies have largely confirmed the validity of Zipf's law for several large countries (Gabaix, 2009; Gabaix and Ioannides, 2004). The evidence is more mixed for smaller economies (Soo, 2005). For example, while Knudsen (2001) finds support for Zipf's law in Denmark, Brakman et al. (1999) reject it for Sweden. Finally, very few studies have considered regions or countries' territorial partitions (Postigliola and Salvati, 2025; Muolo et al., 2025).

To the best of our knowledge, Italy has not been the subject of a dedicated empirical investigation on this front, except for its inclusion in the multi-country analysis conducted by Soo (2005). In that study, a common threshold was applied across countries to define the upper tail, resulting in a sample of 228 Italian municipalities. The OLS estimated parameter for Italy in 1999 was $\hat{\alpha}$=1.381. Notably, Eeckhout (2004) finds for the United States that the absolute value of the estimated shape parameter tends to decrease as the truncation threshold for the sample population increases—highlighting the sensitivity of power law estimates to the definition of the upper tail.

Motivated by these insights, we estimate the log-linear form of equation (1) by OLS:

$$\ln(Rank_i) = \ln(\beta) + \alpha \ln(Size_i) + \varepsilon_i, \quad (4)$$

adopting alternative truncation thresholds: 250, 1000, 2,500, 5,000, and 7,000. Results are reported in Table 2.

As one can notice, the estimated coefficients are not too far from, but not even exactly equal to, one. When the sample size is more reduced, the estimate is larger than one, meaning a more even distribution than that implied by Zipf's law. As the sample size increases, the estimated coefficient gets smaller and is very close to one when the cutoff population size is around 5,000.

A graphical interpretation of these findings is provided in Figure 3, which plots the (natural) logarithm of city rank against the (natural) logarithm of city size, using 2024 as the reference year. If the data were perfectly consistent with Zipf's law, the points would lie along a straight line with a slope of –1. The extent to which the observed distribution deviates from this benchmark provides a visual assessment of the law's empirical validity for the upper tail of the Italian city size distribution. In Appendix A.3, we replicate this figure at five-year intervals, obtaining very similar results.

An additional empirical question is how these distributional properties evolve over time. As discussed in Subsection 3.2, Italian smaller cities tend to experience population decline on average, larger cities tend to grow, and the very largest cities remain relatively stable. From this pattern, we can infer that inequality at the lower end of the distribution is likely increasing, while at the upper end it is decreasing. As a result, we expect the estimated Pareto coefficients (in absolute value) to decline over time when using a broad sample that includes smaller cities, and to increase when the sample is more severely truncated to include only the largest municipalities. This predicted pattern is confirmed in Figure 4, which plots the evolution of the coefficient estimates (as well as the corresponding 95% confidence intervals) over time across the five different population thresholds.

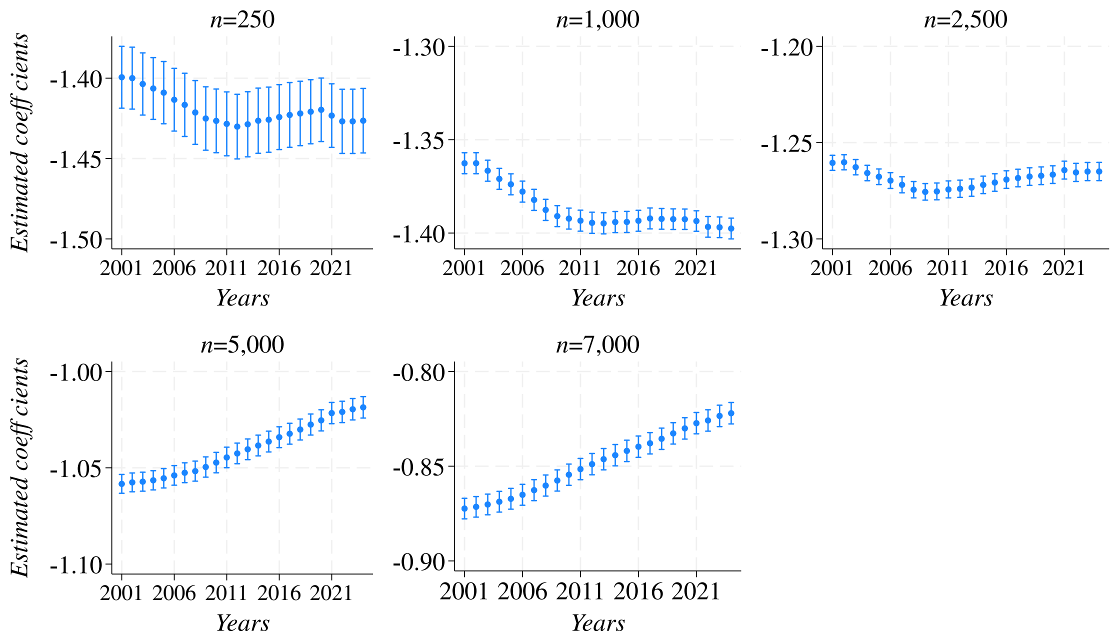
4 Concluding remarks

Understanding the drivers and patterns of urban development is a central theme in policy agendas. The structure of economic incentives—those that, on the one hand, push populations to concentrate in cities and towns, and those that, on the other hand, lead cities to differentiate and organize into hierarchical systems—is crucial for predicting and designing urban systems. Yet, our understanding of these push and pull forces remains incomplete, particularly when it comes to reconciling empirical evidence with theoretically grounded, microfounded growth models. In particular, while push and pull factors can explain the emergence of cities, they are not sufficient to account for the persistence of urban hierarchies in equilibrium (for a discussion, see, for instance, Brakman et al., 1999). In the absence of a widely accepted theoretical framework, and given the presence of robust and recurrent empirical regularities, gathering evidence across as many countries as possible is essential to deepen our understanding and to inform policymakers about long-run dynamics.

This study contributes to the literature on the distributional properties of city size by examining the case of Italy. To the best of our knowledge, it is the first to empirically test three foundational regularities—Gibrat's law, Zipf's law, and the log-normality of the nontruncated city size distribution—using data for the entire population of Italian municipalities. The analysis provides valuable insights for urban geography by shedding light on the agglomeration process in an advanced European country, an area where empirical research has been relatively scarce compared to other economies.

Our findings show that the size distribution of Italy's approximately 7,900 municipali-

Figure 4. Pareto coefficients over time



Note: The graphs plot OLS estimated coefficients from equation (4) over time across the five different population thresholds: 250, 1,000, 2,500, 5,000, and 7,000. Confidence intervals at the 95% level are also reported.

ties closely follows a log-normal distribution, that population growth rates are largely independent of city size in line with Gibrat's law, and that conformity with Zipf's law is highly sensitive to the choice of truncation thresholds.

We further explore how these empirical properties evolve over time. While the log-normality of the full distribution and the proportionate growth property have proven to remain relatively stable, the power-law behavior of the upper tail—captured by the estimated Pareto coefficient—changes significantly. Notably, the Pareto coefficient can be interpreted as an indicator of inequality in city size. Our results suggest that inequality in the upper tail of the distribution has declined over time, while inequality in the overall distribution has increased. This finding aligns with broader trends identified in the literature on Italian urban dynamics (e.g., Salvati et al., 2013).

The persistence of urban hierarchies over time in Italy—characterized by a few very large centers, several medium-sized cities, and a multitude of small towns—is a crucial feature with significant implications for a range of policy issues, from sustainable territorial development to regional and local disparities. A distinctive feature of Italy's urban system—common to other Southern European countries—is its pronounced spatial heterogeneity, particularly the persistent migration gradient from Southern to Northern regions (Buonomo et al., 2024). While this aspect was not addressed in the present analysis, examining how the three empirical regularities vary across different regions within Italy and similar European countries represents a promising direction for future research.

References

Accetturo, A. and S. Mocetti (2019). Historical Origins and Developments of Italian Cities. *Italian Economic Journal 5*, 205–222.

Alaimo, L. S., C. Ciaschini, F. Mariani, E. Cudlinova, M. Postigliola, D. Strangio, and L. Salvati (2023). Unraveling Population Trends in Italy (1921–2021) with Spatial Econometrics. *Scientific Reports 13*, article 20358.

andrea, L. and andrea Petrella (2019). Stylized Facts on Italian Cities. *Italian Economic Journal 5*, 223–249.

Berger, V. and Y. Zhou (2014). Kolmogorov–Smirnov Test: Overview. In B. N., T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, and J. Teugels (Eds.), *Wiley StatsRef: Statistics Reference Online*, pp. 2341–2378. Wiley Online Library.

Brakman, S., H. G., C. van Marrewijk, and M. van den Berg (1999). The Return of Zipf: Towards a Further Underst anding of the Rank-Size Distribution. *Journal of Regional Science 39*(1), 182–213.

Buonomo, A., F. Benassi, G. Gallo, L. Salvati, and S. Strozza (2024). In-between Centers and Suburbs? Increasing Differentials in Recent Demographic Dynamics of Italian Metropolitan Cities. *Genus, Journal of Population Sciences 80*(1), 1–19.

Cattaneo, M. D. and M. Jansson (2018). Kernel-based Semiparametric Estimators: Small Bandwidth Asymptotics and Bootstrap Consistency. *Econometrica 86*, 955–995.

Champernowne, D. G. (1953). A Model of Income Distribution. *Economic Journal 63*(250), 318–351.

Duranton, G. (2007). Urban Evolutions:The Fast, the Slow and the Still. *American Economic Review 97*(1), 197–221.

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

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

Gabaix, X. (2009). Power Laws in Economics and Finance. *Annual Review of Economics 1*, 255–294.

Gabaix, X. (2016). Power Laws in Economics: An Introduction. *Journal of Economic Perspectives 30*(1), 185–294.

Gabaix, X. and Y. M. Ioannides (2004). The Evolution of the City Size Distributions. In J. V. Henderson and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4, pp. 2341–2378. Elsevier.

Gibrat, R. (1931). *Les inégalités économiques. Paris: Librairie du Recueil Sirey.* Paris: Librairie du Recueil Sirey.

Ioannides, Y. M. and H. G. Overman (2003). Zipf's Law for Cities: An Empirical Examination. *Regional Science and Urban Economics 33*, 127–137.

Italian Institute of Statistics (2025a). *Censuses, Population and Households, Permanent Census of Population and Housing*. https://www.istat.it/en/statistical-themes/population/ (last access: June 26, 2025).

Italian Institute of Statistics (2025b). *Demo: Demografia in Cifre*. https://demo.istat.it/ (last access: June 26, 2025).

Knudsen, T. (2001). Zipf's Law for Cities and Beyond: The Case of Denmark. *American Journal of Economics and Sociology 60*(1), 123–146.

Muolo, A. a., M. Postigliola, L. S. Alaimo, and L. Salvati (2025). Zipf and the MAUP: Empirical Evidence of Non-parametric Estimates from a Mediterranean Country. *Quality and Quantity*.

Nadaraya, E. A. (1965). On Nonparametric Estimates of Density Functions and Regression Curves. *Theory of Probability and its Applications 10*(1), 186–1905.

Postigliola, M. and L. Salvati (2025). Land-use Change and Zipf's Law: A Novel Tool to Investigate Regional Transformations. *Papers in Regional Science 104*(1), article 100074.

Rosen, K. T. and M. Resnick (1980). The Size Distribution of Cities: An Examination of the Pareto Law and Primacy. *Journal of Urban Economics 8*(2), 165–186.

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

Salvati, L., V. G. Morelli, K. Rontos, and A. Sabbi (2013). Latent Exurban Development: City Expansion along the Rural-to-Urban Gradient in Growing and Declining Regions of Southern Europe. *Urban Geography 34*(3), 376–394.

Soo, K. T. (2005). Zipf's Law for Cities: A Cross-country Investigation. *Regional Science and Urban Economics 35*, 239–263.

Watson, G. S. (1964). Smooth Regression Analysis. *Sankhyā: The Indian Journal of Statistics 26*(4), 359–372.

Wong, D. W. (2009). Modifiable Areal Unit Problem. In R. Kitchin and N. Thrift (Eds.), *International Encyclopedia of Human Geography*, pp. 169–174. Elsevier.

Zipf, G. K. (1949). *Human Behavior and the Principleof Least Effort*. Cambridge, MA: Addison-Wesley Press.

Acknowledgements

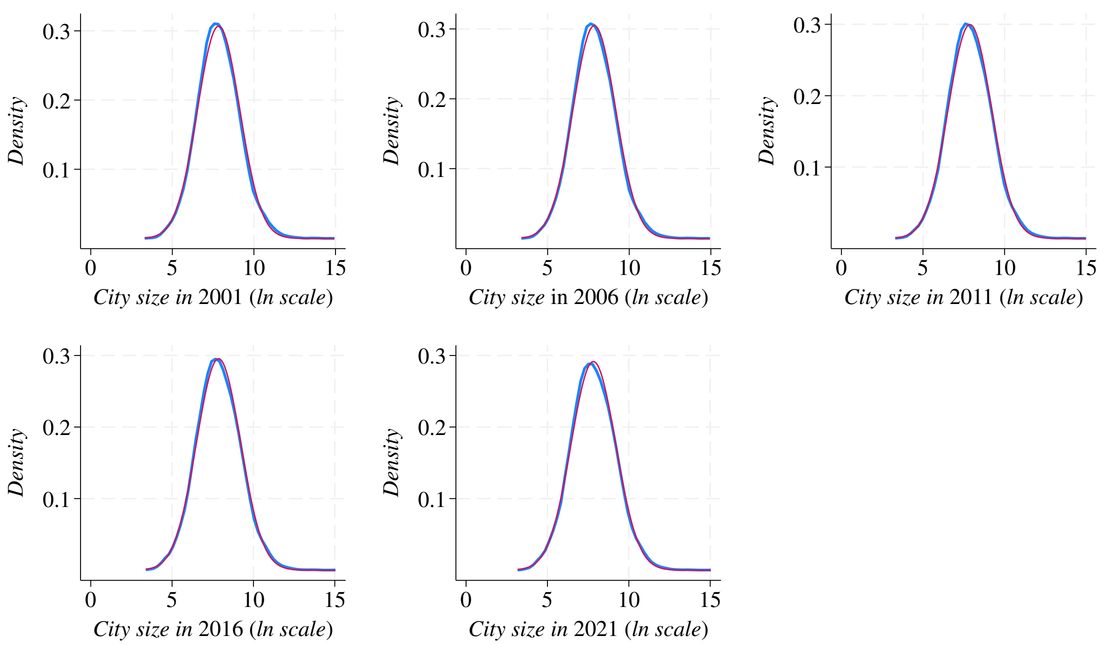
I would like to thank two anonymous referees for valuable comments and suggestions. All remaining errors are my own.

A ONLINE APPENDIX

A.1 Log-normality over time

Figure A.1 compares the empirical distribution of municipality sizes with the corresponding theoretical log-normal distribution every five years within the sample period: 2001, 2006, 2011, 2016, and 2021.

Figure A.1: Density functions of city size



Note: The graphs plot the city size density functions and the theoretical normal density at five-year intervals.

Corresponding results from a Kolmogorov–Smirnov (K–S) test are reported in Table A.1, suggesting not to reject the null hypothesis of log-normality with one exception only.

Table A.1: Kolmogorov–Smirnov test

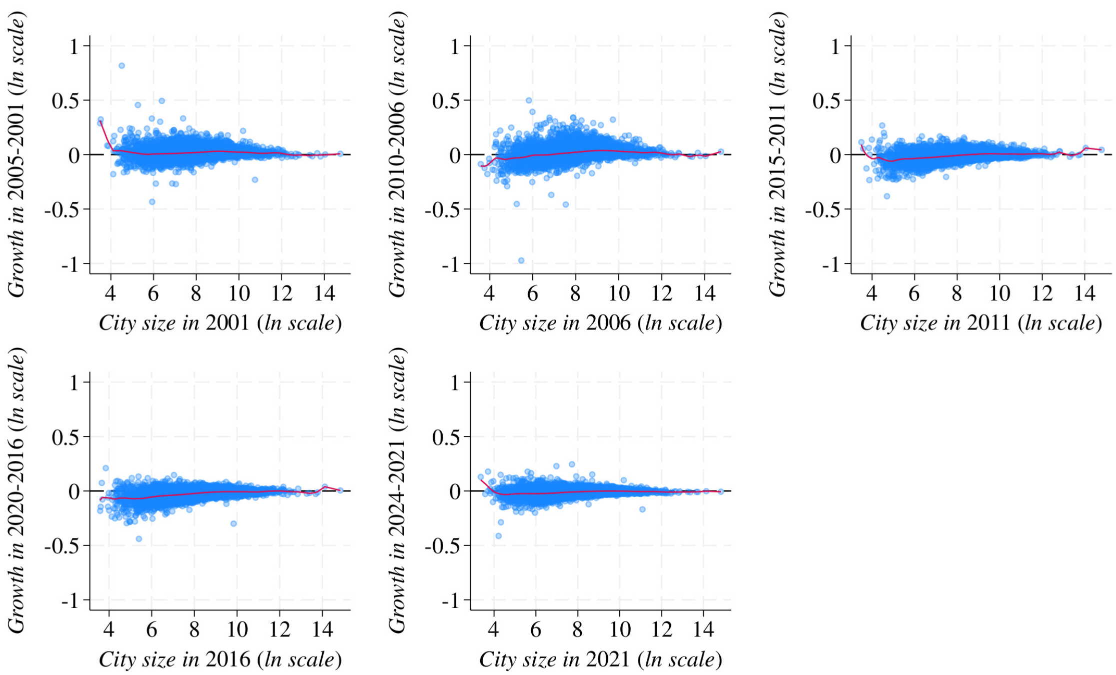
Year	2001	2006	2011	2016	2021
Mean of (natural) log distribution	7.841	7.862	7.877	7.859	7.823
St. dev. of (natural) log distribution	1.330	1.310	1.332	1.350	1.368
$\hat{D}$	0.019	0.016	0.014	0.014	0.014
p-value	0.005	0.037	0.101	0.091	0.103

Note: The table reports results from Kolmogorov–Smirnov test (Berger and Zhou, 2014).

A.2 Proportionate growth over time

Figure A.2 plots the distribution of municipality-level growth rates against initial city size (log scale) over five-year intervals: 2001–2005, 2006–2010, 2011–2015, 2016–2020, and 2021–2024 (with the final interval covering four years). The red lines represent non-parametric kernel estimates of the conditional mean of (log) population growth, computed using the Nadaraya–Watson estimator (Epanechnikov kernel with optimal bandwidth and 500 bootstrap replications). As can be observed, the scatter plots for each interval appear more compressed compared to the full-period analysis, and consistently support the hypothesis of proportionate growth. This result supports the stability of Gibrat's law across different time windows.

Figure A.2: Growth rate, city size, and kernel conditional mean estimate

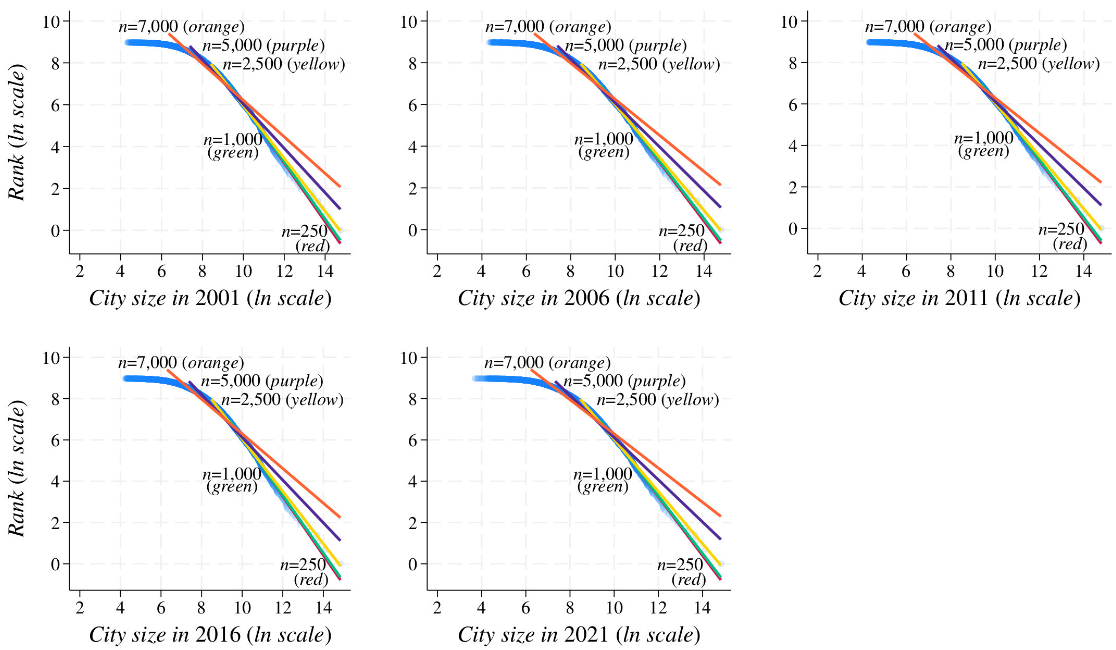


Note: The graphs plot the (natural) log of growth rates against the (natural) log of initial city size, and the kernel conditional mean estimate at five-year intervals.

A.3 Zipf's coefficients estimates over time

Figure A.3 presents log-log plots of city rank versus city size at five year intervals. The deviations of the empirical distributions from the fitted regression lines remain consistent across years and closely resemble the pattern observed in 2024.

Figure A.3: City size versus rank



Note: The graphs plot the (natural) log of city size against the (natural) log of rank with linear regression lines at five-year intervals across five different truncation thresholds.