What do analyses of city size distributions have in common?

Cottineau, C.

Citation (APA)
Cottineau, C. (2022). What do analyses of city size distributions have in common? *Scientometrics: an international journal for all quantitative aspects of the science of science, communication in science and science policy*, *127*(3), 1439-1463. https://doi.org/10.1007/s11192-021-04256-8

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

What do analyses of city size distributions have in common?

Clémentine Cottineau[1,2]

Abstract

In this article, I conduct a textual and contextual meta-analysis of the empirical literature on Zipf's law for cities. Combining citation network analysis and bibliometrics, this meta-analysis explores the link between publication bias and reporting bias in the multidisciplinary field of quantitative urban studies. To complement a set of metadata already available, I collect the full-texts and reference lists of 66 scientific articles published in English and construct similarity networks of the terms they use as well as of the references and disciplines they cite. I use these networks as explanatory variables in a model of the similarity network of the distribution of Zipf estimates reported in the 66 articles. I find that the proximity in words frequently used by authors correlates positively with their tendency to report similar values and dispersion of Zipf estimates. The reference framework of articles also plays a role, as articles which cite similar references tend to report similar average values of Zipf estimates. As a complement to previous meta-analyses, the present approach sheds light on the scientific text and context mobilized to report on city size distributions. It allows to identified gaps in the corpus and potentially overlooked articles. It confirms the relationship between publication and reporting biases.

Keywords City · Size · Zipf · Text · Context · Meta-analysis

Introduction

The parallel development of ever-increasing amounts of data accessible at the city level on the one hand and of non-linear regularities being found as a marker of complexity (in networks, in organisms, in cities) on the other hand have reignited interest in the study of city size distributions in recent years, more specifically a renewal of the appraisal of Zipf's law for cities. Of particular importance in this research stream are the definition of the objects studied (i.e. the limits, thresholds and composition of cities which affect the population counted or not as urban), the model to summarize the distribution (power-law, lognormal, polynomial) and its fit

✉ Clémentine Cottineau
  c.cottineau@tudelft.nl

1 Centre National de la Recherche Scientifique (CNRS), UMR 8097 Centre Maurice Halbwachs, Paris, France

2 Department of Urbanism, Technical University of Delft (TU-Delft), Urban Studies Section, Delft, Netherlands

to the data (fitting procedure, value of the power exponent, uncertainty). It is usually agreed upon that city populations follow a heavy-tail distribution in most countries or regions and at most times, although the precise form of the distribution and the estimation of its main parameters tend to vary. The universality claim of Zipf's law (1949) can thus be accepted with respect to the general trend, but it is rejected in its strictest form (i.e. a power law of exponent −1 between city sizes and their ranks when ordered by size). Previous meta-analyses of studies providing an empirical estimation of Zipf's exponent have shown indeed that on average, empirical estimations tend to deviate from the strict value of −1 (Cottineau, 2017; Nitsch, 2005). For simplicity in the rest of the paper, I consider the absolute value of the exponent, called "alpha" and its deviation from 1. A share of such deviations can be attributed to differences in the technical specifications of the studies (their total number of estimates, the range of countries and periods analysed) and of the empirical estimation (delineation of cities, population thresholds, estimation procedure, etc.). A smaller share of the variance can be attributed to territorial characteristics of the city system (its level of urbanisation for instance) and no share has been found to vary significantly with planning actions (Cottineau, 2017). Therefore, empirical deviations to Zipf's law remain for the most part unexplained or unexplored. To address this research gap, this contribution turns to publication bias in a meta–meta-analysis.

Publication biases as well as differences in reference frameworks and disciplinary traditions might generate systematic differences in the measurement and reporting of empirical distributions of city sizes, but they are unobservable with a traditional meta-analysis. For instance, despite addressing the same empirical estimation of Zipf's law (same country, same set of cities, same date, same estimation method), there can be strong differences in the way the papers from a given meta-analysis frame, exploit and report on this result, depending on the aim of their research (such as: "proving that Zipf's law is a universal feature of urban systems", "showing that the lognormal form is better suited", "looking for national differences in urban hierarchy", etc.). The empirical results of such studies could therefore appear clustered along the lines of schools of thoughts. The present article asks the question: what do analyses of city size distributions have in common? It goes a step further in the secondary analysis of Zipf's law for cities, by exploiting similarity networks drawn by the studies included in a meta-analysis. Building on the open-source corpus of MetaZipf (Cottineau, 2017), which contains 1962 empirical estimations of Zipf's exponent alongside their technical and territorial specifications from 86 studies, it characterises the pairwise similarities of studies based on their reference lists, textual content and disciplinary exposure—elements which, incidentally, loop us back to Zipf's original research in linguistics. Combined with the pairwise similarity network of the study content, the analysis aims to reveal new insights about the observed deviation of urban Zipf estimates in recent publications. I find evidence that pairs of articles with similar wording and similar reference lists tend to report similar average values of estimates. Similar wording also correlates positively with a similarity level of dispersion of values reported. The data and code of the present study has been made fully available on Github, including an R notebook with all visualisations: https://clementinecttn.github.io/MetaZipf/metametazipf_notebook.nb.html .

Bibliometrics, networks and systematic literature reviews: towards a meta–meta-analysis

Meta-analyses are convenient tools to reflect on the collective production of an established field of inquiry, especially when it produces quantitative estimations and prediction statements. In that respect, city size distributions and their modelling with power laws date back more than a century (Auerbach, 2013), and still inspire dozens of dedicated articles every year. This scientific production originates from a diversity of disciplines and research domains such as economics, geography, statistics, physics, regional science, planning and mathematics. Consequently, authors of studies included in a Zipf meta-analysis tend to publish in a diversity of journals which all have different formal and theoretical requirements: the amount of text, the type of proofs received as valid, different evaluations of the necessary, legitimate and superfluous references. For instance, economics journals usually require econometric models with particular control variables and specific ways of presenting results in standardised tables. Physics journals tend to publish shorter articles with large supplementary materials. "*Planning papers tend to cite eclectically. […] This will be a feature of social science in general compared with science journals but, within the social sciences, one might expect certain broader applied subjects such as planning to be especially unfocused in the literature they cite. […] Planning papers are also eclectic in the type of references cited reports and plans as well as academic papers and this may lower impact statistics.*" (Webster, 2006, p. 488). Journals in geography will tend to favour analyses of spatial variations of a given phenomenon rather than its universality. Could such meta-properties signal differences in the definition of the aim of the research, the design of the experiment and ultimately the value of the results reported? Could they offer a new angle to explain the remaining variability of published results? The hypothesis driving this research is that they could. Indeed, science is a social practice performed by actors embedded within institutions, disciplinary frameworks and legacies (Latour and Woolgar, 1986). It would therefore be possible to suggest that studies written in a similar way, citing similar references and publishing in the same kind of journals exhibit more similar results (controlling for the object of their study, in our case: the similarity of cities, countries and time periods studied) than studies which originate from very different fields, point to very different reference lists and use distinct scientific languages. There is evidence from the MetaZipf corpus that a significant diversity of languages, reference frameworks and disciplines exist. For example, Gabaix & Ibragimov's (2011) article is built like a mathematical demonstration (using terms like "theorem" and "lemma" several times) whereas other articles read more like monographs. Some articles systematically reference back to Zipf (1949) and Auerbach (1913), whereas others start the debate where Gabaix (1999) left it. Some articles cite a very large number of external references (Parr, 1985 or Berry & Okulicz-Kozaryn, 2012) when others do not (such as articles published early or in physics journals). Finally, the range of journals cited and chosen for publication is broad, it ranges from mainstream economics to specialised geography, statistical physics and beyond. The objective of the present work is to assess whether such diversity is reflected systematically in the variation of results reported, in order to better understand urban hierarchies around the world (rather than the scholars who study them).

Networks and meta-analysis

The secondary analysis of data through bibliometrics and systematic reviews has gained momentum since the 1990, and in particular through the use of tools and techniques from network science. A direct example of this trend is the development of network meta-analyses (NMA), i.e. the analysis of networks of evidence to compare indirect treatment effects, usually in randomized clinical trials. This methodology is used to "*detect inconsistency between randomized trials of different treatments, to estimate treatment differences and to assess the uncertainty in these estimates. Estimating treatment differences is straightforward; the main contribution […] is in methods for detecting and estimating the inconsistency among trials*" (Lumley, 2002, p. 2314). A bibliometric analysis of this method shows that it has developed into a growing field, counting over "*2846 studies [...,] published in 771 journals in six languages*" (Shi et al., 2021, p. 1). Unlike in medical studies where the pool of potential patients is very large and trials cannot be replicated endlessly on the same population, the analysis of cities can—and should—use a similar set of observations to estimate Zipf's law, because they rely on published population data (from censuses or surveys every few years) for their observation. Indeed, where statistics are available, they are the same for all observers, they do not need to give consent, and they do not change over time (or only in highly unusual circumstances of data manipulation). The same dataset for cities at a given year can—and should—therefore be the same for all researchers, and all researchers using the same criteria for delineating cities could—and should—end up with the same estimations of Zipf's law. If this is not the case, it means that a diversity of delineations and estimation procedures coexist in the literature. If these procedures covary with disciplinary traditions, with citation behaviours and/or writing styles, we are then faced with a case of publication bias, whereby the reported result of an analysis depends on the context of production of the study as well as on the relations present in the primary data. What could happen is that different authors delineate cities and report Zipf estimations differently depending on the field they come from, which translates into the words they write with and the citations they mobilise. In this contribution, I propose to look at the pairwise similarity in citation, wording and discipline and to include them as variables in the meta-analysis of reported Zipf's exponents. Although some consider co-citation networks as a type of meta-analysis (Fetscherin & Heinrich, 2015), or a preliminary step to a systematic literature review (Chang et al., 2021; Losse & Geissdoerfer, 2021; Radler, 2018), in this contribution I propose to use citation networks as an input in a meta-analysis, defined as the quantitative exploitation of a systematic literature review.

Citation bias, reporting bias and systematic reviews

"*Given the large body of scientific literature, it is often unfeasible to cite all published articles on a specific topic, and so, some selection needs to take place. If this selection is influenced by the actual results of the article, then citation bias occurs. [….] Similarly, in a recent survey among researchers, selective citation was ranked as the most frequently occurring research misbehavior. To assess the potential consequences of citation bias, a proper understanding of its ubiquity is required.*" (Duyx et al., 2017, p.92–93). Whilst acknowledging this caveat, reversing the question is also fruitful: is there a relationship between the set of publications cited in an article the value of results reported in it? This relationship could derive from the fact that reporting biases differ by discipline. In our case, let's assume that positive results related to Zipf's law are different between geography,

economics and physics (the three main disciplines interested in estimating Zipf's law for cities). In economics and in physics, a positive result corresponds to the confirmation of Zipf's law with an exponent equal to 1. By contrast, in geography, finding a deviation from the law consistent with national characteristics (newly urbanised countries having a distribution of city size more uneven than the unity power law for instance) constitutes the positive result rather than the opposite. Since "*positive articles are cited about twice as often as negative ones*" (Duyx et al., 2017, p. 93), articles from the three disciplines would preferentially cite articles with opposite conclusions and preferentially report results marked by their own disciplinary bias. Given the existence of these two types of bias (citation and reporting) and their opposite consequence on different disciplines, we could therefore expect a relationship between the articles cited in empirical papers and the value of the exponents they report.

"*Citation bias distorts this balanced representation and may lead to false beliefs. The good news is that there is a self-correcting mechanism in the form of systematic reviews, which ideally take all published evidence into account regardless of whether it has been cited before or not. Still, although systematic reviews and meta-analyses are often regarded as providing the best form of evidence, they can be flawed and even misleading) […] it has been shown that the conclusions of reviews (both narrative and systematic) can be predicted from the choice of which literature was cited in those reviews. In other words, if this cited literature is biased, wrong conclusions can be drawn.*" (Duyx et al., 2017, p. 98). Consequently, and similarly to empirical studies, two meta-analyses of the same question could lead to different conclusions, depending on the disciplinary orientation of their authors and their openness to a multidisciplinary literature. This suggests the opportunity for a meta–meta-analysis, i.e. a study of the literature which includes the context of publication as an explanatory factor in the meta-analysis.

Meta–meta-analyses

The rapid expansion of knowledge and the faster pace of publication in recent years have stimulated the development of meta–meta-studies, i.e. the research on research. Such studies comprise the systematic literature review (SLR) of systematic literature reviews (cf. Kitchenham et al., 2009 in the field of engineering, Xiao and Watson (2019) and Mohamed Shaffril et al., 2021 on the methodology of SLR themselves), or the meta-analysis of meta-analyses, as in Geyskens et al. (2009). Most of these endeavours are developed to test the robustness of existing research and to promote good practice in literature reviews. Meta–meta-research conducted on citation networks, as in Duyx et al. (2017), can also inform about the link between some article results and their further citation by peers, thus about the magnitude and direction of citation biases. However, this type of analysis does not look at who is using positive or negative results preferentially. The present study, by contrast, investigates whether a similar use of literature between two empirical papers is associated with a similar bias in result reporting, controlling for differences in datasets and estimation specifications.

A second avenue of work around the quantitative research on research is the identification of metaknowledges. "*Metaknowledge, generally speaking, is the knowledge about knowledge. It helps scholars track knowledge through topics […] some articles integrate conflict findings through conventional meta-analysis (Stam* et al*. 2014). However, such articles cannot display the network, or characteristics, of multiple metaknowledges. To fill up those research gaps and identify metaknowledge trends and features, this paper adopts*

two analytical methods: co-citation analysis and network meta-analysis." (Zhang & Guan, 2017, pp. 1177–1178). With this strategy, the authors are able to identify the research fronts, research bursts and metaknowledges from citation analysis of the multidisciplinary field of entrepreneurial ecosystems. Starting from the most significant articles identified through the citation analysis, the network meta-analysis enables them to estimate the correlations between these metaknowledges, thus integrating conflicting findings. They illustrate the opportunity of combining citation network analysis and meta-analysis to conclude on the possibility of publication bias. The present article proposes to start from there and to re arrange the two analyses in a novel way: using citation analysis as a proxy for publication bias into the meta-analysis. "*The paucity of meta-analytical studies testing for publication bias offers cause for concern because confidence in the validity of the findings of a meta-analysis depends on ruling out the presence of publication bias. In conducting a meta-analysis, researchers should always make efforts to assess to what extent publication bias may affect their findings*" (Geyskens et al., 2009, p. 404). Considering this advice, I include one aspect of publication bias, namely the level of similarity between two articles' reference list, as an explanatory variable of their similarity in estimated Zipf coefficients.

Methods and materials

This section details the collection of the meta–meta-data and the strategy used to convert it into pairwise similarity matrices along a number of dimensions. It also presents the model used to regress differences in the reported distribution of Zipf estimates by meta-properties, controlling for technical and territorial specifications. The material of the present study consists in a corpus of studies which have all published estimations of Zipf's power law exponent on empirical city size distributions. It makes use of the openly available database MetaZipf,[1] which contains 1962 such empirical estimations from 86 studies, along with their specifications. The 86 studies have been selected to fulfil three criteria: "*they contain at least one estimate of the rank-size exponent based on population; the regression is made on empirical urban data; the regression model is bivariate (i.e. relating populations and ranks or ranks—1/2, but not to any other instrumental variable).*" (Cottineau, 2017, p. 4). In the present work, only 66 of them fulfilled additional criteria detailed below. This subset of 66 studies are subsequently referred to as "the corpus".

Collecting full-texts

For an article from the MetaZipf database to be included in "the corpus", it has to available in open-access or accessible with an extensive institutional subscription, in a machine-readable format. Additionally, only published journal articles written in English were selected, in order to run a coherent textual analysis. This excludes texts in other languages and formats, such as books and dissertations. This choice is detrimental to the recognition that science is plural in forms, languages and origins. However, it did not affect the original sample too much, since most references in MetaZipf were already predominantly in English and in an article format. The corpus is thus composed of 66 full-texts of English-written articles. Out of the original document, only the body of the text was retained. This
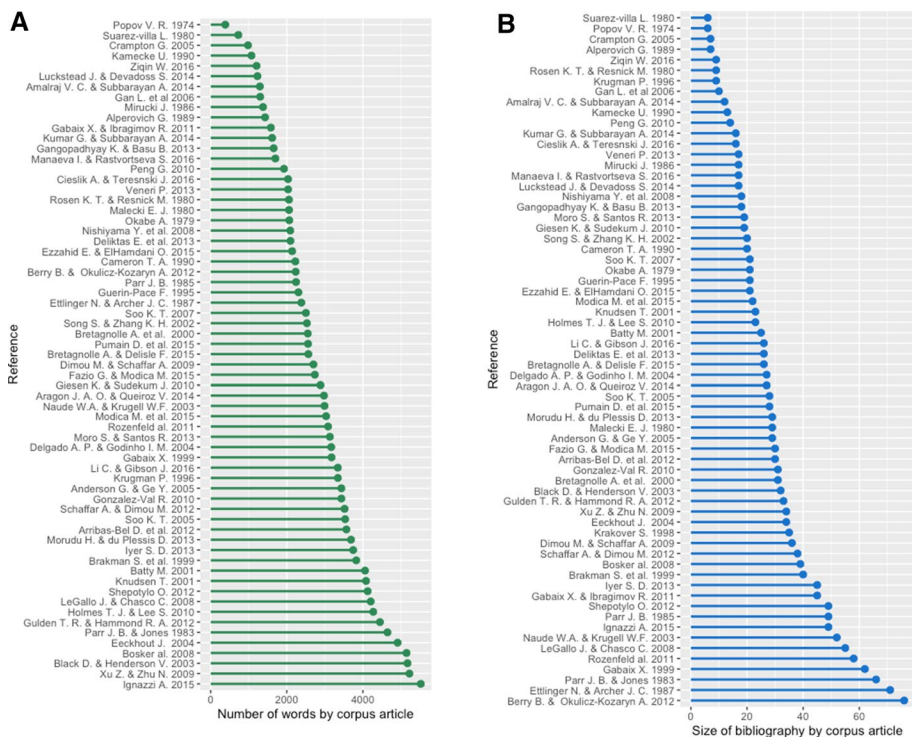

Fig. 1 a (left). Distribution of text size in the corpus (number of words excluding stop-words). **b** (right). Distribution of bibliography size in the corpus (number of external references)

means that titles, affiliations, abstracts, keywords, section titles, figures, tables, equations, references, footnotes and line breaks were removed. The remaining text was used for text mining analysis, after a traditional automated treatment (with the R 'tm' package, cf. Feinerer et al., 2019) in order to remove punctuation, numbers and stop-words and transform the remaining word to lower case. Term frequencies were attached to each reference to allow for a study of wording similarity between them.

After treatment, corpus articles exhibit a continuous array of sizes (Fig. 1a), from 384 words for Popov (1974) to 5522 words for Ignazzi (2015). Apart from significantly shorter sizes in physics articles (around 1600 words on average per corpus article, compared to 3000 on average in economics and 2500 in geography), I do not find any trend by year of publication or else.

Collecting citations

To explore the citation network of corpus articles, each reference from the 66 English-written articles was recorded and formatted in a way that allows to query the authors' names, the year and journal of publication. The 66 corpus articles generated 304 internal citations (i.e. to other articles included in the corpus) and citations to 1155 distinct external references (including references to articles, reports, books or dissertations in various languages) from over 700 different journals or publishing institutions. Corpus articles exhibit once
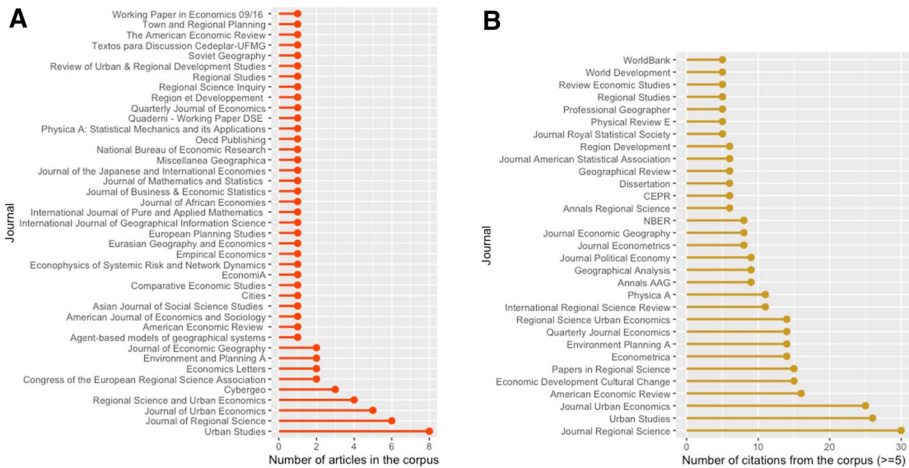
Fig. 2 **a** (left). Distribution of corpus articles by journals (and series) publishing them. **b** (right). Distribution of articles cited by corpus articles by journals (and series) publishing at least 5 of them

again a disparity of (external) bibliography sizes (Fig. 1b), from 6 items in Suarez-Villa (1980) and Popov (1974) to 76 in Berry and Okulicz-Kozaryn (2012). Apart from significantly shorter sizes in physics articles (around 15 items on average, compared to 22 on average in economics and 24 in geography and regional science), I do not find any systematic variation by year of publication or else.

The journals most frequently chosen to publish corpus articles (Fig. 2a) coincide with the journals where bibliographical references most frequently come from (Fig. 2b), i.e. Urban Studies and the Journal of Regional Science, the Journal of Urban Economics, Regional Science and Urban Economics or the Journal of Economic Geography. The average year of publication in the corpus is 2004, ± 1 year for articles from different disciplines except articles published in physics journals, whose interest in city size distributions and average year of publication is more recent (2013). By contrast, the average year of publication for external references is 1989.

The most cited external reference is to Zipf himself. 37 out of the 66 corpus articles cite it for its 1949 book on the "principle of the least effort" and 5 cite it for its 1941 work "National unity and disunity; the nation as a bio-social organism". Corpus papers not citing any of the two Zipf references are frequent among those published at earlier dates (Fig. S1 in Supplement), and proportionally more in geography and economics journals where it might be considered obvious. By contrast, 4 out of 4 articles published in physics journals cite Zipf. It is interesting to note, however, that Zipf's work is not the most cited reference in the corpus: two internal references appear even more frequently (Fig. 3A): Gabaix's theoretical 1999 paper (41 out of the 50 other articles published in or after 1999) and Rosen & Resnick's comparative 1980 paper (39 out of the 62 other articles published in or after 1980). Externally (Fig. 3b), the reference to Auerbach's work from 1913 is in the top 3 with 23 external references, but it is less frequently cited than Gabaix and Ioannides's (2004) chapter in the Handbook of Urban and Regional economics, with external citations from 29 corpus articles.

The graph on Fig. 3B shows that many top references cited externally are early classics of urban theory (Christaller, 1933 [6 cites], Lösch, 1940 [8]) and statistics (Gibrat, 1931
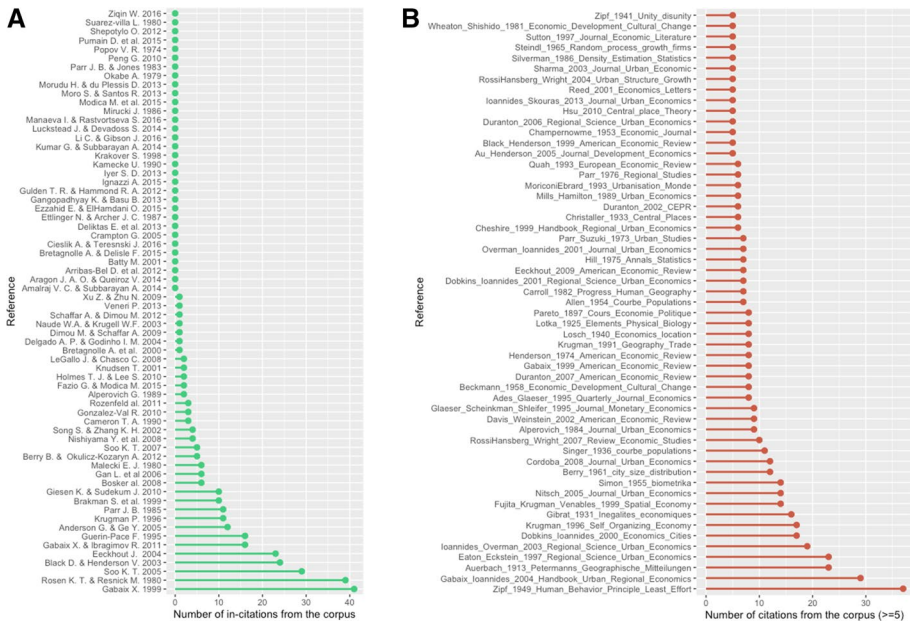
Fig. 3 **a** (left). Distribution of citations to corpus articles by corpus articles. **b** (right) Distribution of citations (over 5) to non-corpus articles by corpus articles

[16 cites], Simon, 1955 [14], Pareto, 1897 [8], Hill, 1975 [7]). Some highly cited references such as Singer, 1936 [11] or Eaton & Eckstein, 1997 [23] actually include empirical estimations of Zipf's exponent, which suggests that they could have been included in the corpus. However, the former was not accessible and the latter contains instruments in the regression. It could be interesting to reconsidered this criterion to include their findings in the future, given their influence on the corpus' reference frameworks.

The most striking feature of this list however is the prominence of post-1995 contributions from three economists in the top cited references (Gabaix, Krugman and Ioannides) compared to earlier works by geographers (like B. Berry in 1961, J. Parr since the 1970s, or F. Moriconi-Ebrard in the early 1990s). As pointed by C. Webster (2006, p. 489–490) in the context of planning journals, "*there is both a publishing and a cognitive limitation on the number of citations included in a paper and this means that the rate of citation growth will be higher, the higher the citation count of a paper. Well-cited papers will become more well cited. If the total number of citations per paper grew to accommodate the increasing number of papers as a field grows, then this inequality might not be inevitable. But reference lists do not get ever longer and, as a result, the frequency of paper citation counts tends to follow a rank-size pattern (sic)*". In the case of top cited papers in this study, they indeed belong to highly visible academics of large, established and dominant disciplinary fields, whose articles in general and the Zipf ones in particular, generate hundreds

to thousands of citations (2133 for Gabaix's, 1999 "Zipf's law for cities, an explanation"). Finally, Nitsch (2005)'s meta-analysis is frequently cited (21% of all corpus articles and 34% of those published in or after 2005). Many externally cited references do not appear on this graph for they receive less than 5 mentions from the 66 corpus bibliographies.[2]

Translating journals into disciplines/disciplinary fields

In order to study the disciplinary dynamics of corpus articles on Zipf's law for cities, I assigned a "discipline" to each of the 707 journals and publishing institutions from which internal and external references of this meta–meta-analysis were taken. I identified 5 fields: Economics (ECO), Geography (GEO), Regional Science and Planning (REG), Statistics (STAT) and Physics (PHY). Although identification of journals in the last two fields was rather straightforward, the lines between Economics, Geography and Regional Science were quite blurry. However, it seemed interesting to distinguish them for two reasons. Firstly, economics and geography are recognised disciplines whose practitioners do not frequently publish in each other's journals, whereas Regional Science sits precisely at the intersection between economics, geography and planning. In regional sciences/studies conferences and journals, it is not unusual to find references to both disciplines. I thus wanted to see if regional science occupied this middle ground position in city size distribution studies as well. Secondly, the separation acknowledges the fact that publication strategies vary greatly between the journals of these fields, in terms of exposure, sphere of impact, formal and theoretical requirements, even when articles deal with the same object.

The keys used to allocate journals between the three fields were the following:

- "ECO" for general economics journals (such as the Quarterly Journal of Economics) as well as journals with "economics" in their names (Journal of Urban Economics for instance), if they do not have "regional science" in it as well.
- "REG" when the subject is "urban affairs", "urban studies", or has "urban and regional" in the name.
- "GEO" for general geography journals (for example, the Annals of the Association American Geographers) as well as journals with names referring to the processes of urban development and urbanisation.

This approach contains some ad-hoc character. I have tried to alleviate it first by providing access to the lookup table. I am aware of existing journal classifications but find that they do not reflect entirely the stakes of this sub-field (nor do they provide guidance for the classification of books, reports and dissertation). Second, an assessment of the most frequent journals with the Scimago classification shows for instance that the ad-hoc fields I attributed to journals matches at least one of the Scopus subject areas proposed by Scimago,[3] considering that "Environmental Science (miscellaneous)" corresponds to Regional Science and planning (table S1 in supplement). The advantage of my system is that it provides a single category for each source, whereas Scimago has a varying number of entries for different journals, and no entry for French journals like "Région et

---

[2] Some of them are indeed quite specific, for instance those from the aerosol literature cited in to Eeckhou (2004): Haaf, Amin and Jaenicke, Rainer. "Results of Improved Size Distribution Measurement in the Aitken Range of Atmospheric Aerosols." *Journal of Aerosol Science*, 1980, *11*(3), pp. 321–330. & Hinds, William C. *Aerosol technology*. New York: Wiley, 1982.

[3] https://www.scimagojr.com/ (accessed on 20/08/2020).

Table 1 Distribution of references by discipline of the journal they were published in

Disciplinary field	ECO	GEO	STAT	REG	PHY	OTHER
Corpus	23	13	0	26	4	0
External references	341	233	126	125	49	281

Développement" which is externally cited 6 times in our corpus, or for dissertations and World bank databases.

After applying this ad-hoc translation to all external references, I can see a stark difference between the distribution of corpus articles by disciplinary field and that of their reference list (Table 1). Indeed, while most corpus articles are published in regional science and economics journals, their framework of reference comes primarily from economics and geography, or at least from articles published in economics and geography journals. Secondarily, corpus articles draw from the statistics (for estimation methods and tools) and regional science literatures. Thirdly, they cite articles published in physical science journals. The large number of "OTHER" references indicates the diversity of Zipf-related work bibliographies, which frequently reference other disciplines (political science, architecture, etc.) and formats (reports, dissertation, etc.).

From individual studies to reference networks

In order to assess whether the common meta-properties of articles dedicated to the empirical estimation of Zipf's law play a role in the variation of results they report, I constructed nine non-oriented networks of similarity between corpus articles. The networks all have 66 vertices corresponding to corpus articles. They differ in the distribution of edge weights connecting vertices. The first three networks ("wording", "citation" and "discipline") were built to test our three hypotheses: that similarity in text, citation and discipline contexts signal similarity of goals and research design resulting in similarities in values reported for Zipf's exponent alpha. The next three networks ("country", "decades" and "city definition") were built to control for the similarity in the objects actually studied by corpus articles. The networks "mean alpha" and "sd alpha" are the one to eventually "explain": they correspond to the networks of corpus articles drawn by the similarity of the distribution of Zipf estimates they report (with the mean and standard deviation—sd). A last network ("n alpha") reflects the similarity of corpus articles based on the number of estimates they report, which has a direct influence on standard deviation calculation and is thus used as a control variable in the regression model. Similarity for all networks was measured pairwise, using the cosine similarity (unless stated otherwise). A subset of each network is visualised using the 'igraph' R package (Csardi & Nepusz, 2006). For better visibility, I apply a cutoff to the weight of edges represented, exclude non-connected vertices, and colour vertices using Louvain community clusters (except in Fig. 6). These representations provide clues for interpretation. However, the modelling analysis is run on the complete networks. The entire analysis is available and reproducible online: https://clementinecttn. github.io/MetaZipf/metametazipf_notebook.nb.html.
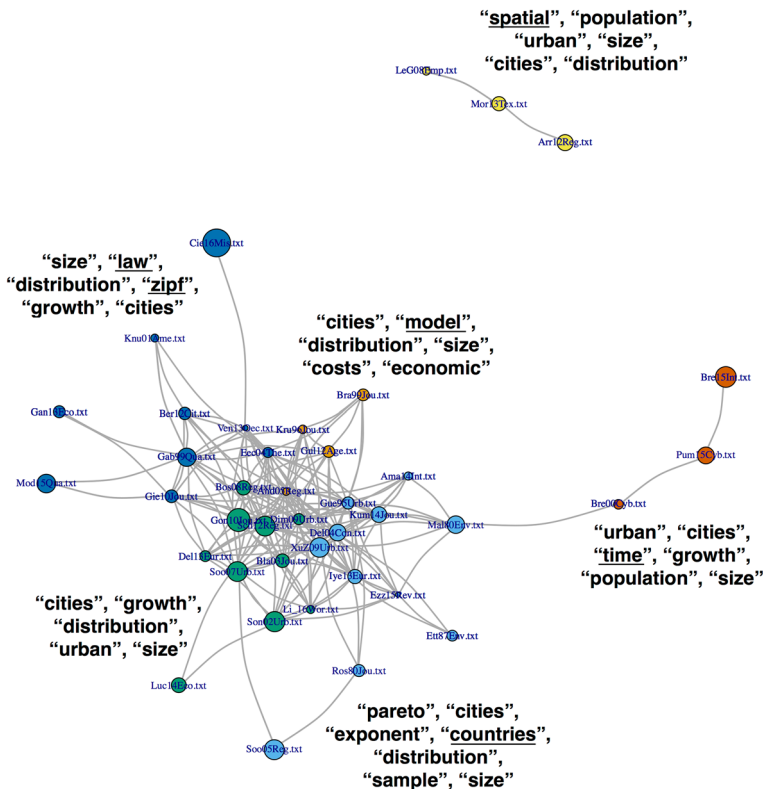
Fig. 4 Similarity network of corpus articles by frequency of words used (cut-off at 0.65). cf. Table S0 in supplement for a lookup table between article identifiers and bibliographic references

The "wording" network

The "wording" network represents the similarity between corpus articles based on the frequency of words they have used to write their paper and to present empirical estimations of Zipf's exponents. Using the 66 full-texts collected, I computed the frequency distribution of all of the corpus non-stop words (10,791 in total) in each article. The vectors used as inputs for the "wording" cosine similarity are therefore composed of 10,791 values comprised between 0 and 1. A visualisation of a subset of the network created is visible on Fig. 4, along with some of the most frequent terms used by corpus articles of the different communities. The variation in vertices sizes represents their total number of terms.

The figure shows a network with strong connectivity. Indeed, most articles use, at the very least, the words "cities", "size" and "distribution" very frequently. However, a disconnected community of three articles (Le Gallo & Chasco, 2008; Moro & Santos, 2013 and Arribas-Bel et al., 2012, in yellow) exhibits a more important use of the word "spatial". These works even have "spatial" in their title. Their aim is not to verify Zipf's law but to present and analyse a national urban system, respectively Spain, Brasil and Australia. Another cluster (in red) shows a specific use of the term "time". Originating from a single research group in France, Bretagnolle et al., (2000, 2015) and Pumain et al. (2015) indeed present long-term evolutions of systems of cities, reporting on their growth and
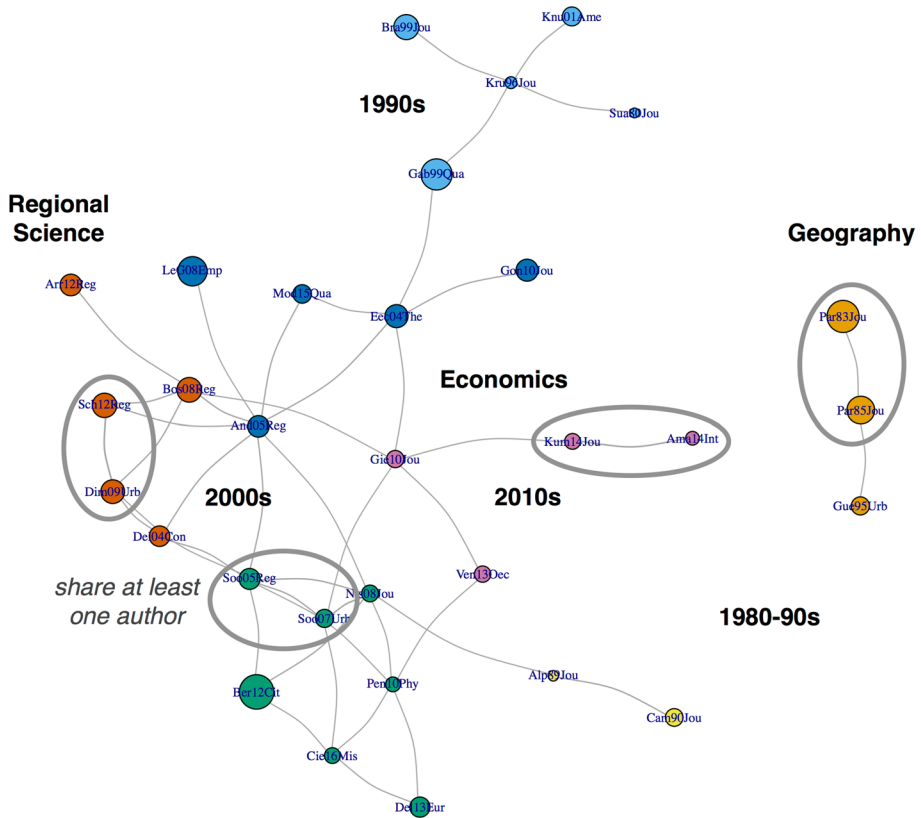
Fig. 5 Similarity network of corpus articles by external articles cited (cut-off at 0.25)

structure of several decades. The light blue cluster gathers comparative studies who therefore make a more thorough use of the term "countries". Two other clusters represent articles less devoted to empirical analysis and more to the testing of "Zipf"'s "law" (dark blue) or "model"-ling its generation (orange). This network thus represent the way Zipf's law is approached by authors and signal somehow the finality of the argument and how estimation results will be used.

The "citation" network

The "citation" network represents the similarity between corpus articles based on the external references they cite in bibliography. It could be argued that two papers citing the exact same corpus of references would more frequently share the same aim, such as "proving" or "disproving Zipf's law", and therefore report more similar estimate values. The similarity was measured using the 66 vectors of 1155 external references, coded 1 if the reference was cited and 0 otherwise. A subset of the network is visible in Fig. 5, with the size of vertices showing the total number of citations of each corpus article.

Fig. 6 Similarity network of corpus articles by external disciplines cited (cut-off at 0.9). Node colours show the discipline of the article rather than its community cluster. Yellow: physics. Green: geography. Light blue: economics. Dark blue: regional science and planning

This network is less connected than others, suggesting that the reference framework of authors depends on a diversity of factors besides a common object of study. Indeed, the subnetwork shown in Fig. 5 is organised along time, co-authorship and disciplinary lines. The similarity of citations is, quite trivially, lower for articles of different periods because of the unavailability of later references for earlier articles. Therefore, we see clusters of corpus articles with similar publication dates (orange and yellow clusters in the 1980s and early 1990s, light blue cluster in the late 1990s, red and green clusters in the 2000s and 2010s, pink cluster in the 2010s). The strong similarity of co-authored articles (Soo, 2005, 2007 or Dimou & Schaffar, 2009 and Schaffar & Dimou, 2012 for instance) reveal the inertia of reference frameworks of individual researchers over time and their relative subjectivity. Finally, articles published in economics journals seem to share more bibliography among them than they do with articles published in geography.

The "discipline" network

The "discipline" network represents the similarity between corpus articles based on the discipline of the journal their external references were published in. The cosine similarity was measured on vectors of 6 items (the number of external references from each discipline). For this representation, I did not use community clusters for colouring nodes but instead the discipline of the journal where corpus articles were published (Fig. 6). The figure shows some entanglement of disciplinary references, with some regional science corpus articles citing a similar pool of disciplines as geography corpus articles. However, this might be an artifact of publication strategies, since the articles in question (such as Parr & Jones, 1983, Guérin-Pace, 1995 or Batty, 2001) are authored by people also recognised as geographers. Corpus articles in economics and in regional science also share similar disciplinary references. The divergence of disciplinary frameworks appears mainly between geography and economics articles, although some articles (Krugman, 1996; Berry and Okulicz-Kozalyn, 2012 or Dimou & Schaffar, 2009) work as bridges, citing from a more varied pool of disciplinary references.

The "country" network

The "country" network represents the similarity between corpus articles based on the countries on which they perform empirical estimations of Zipf's exponents (Fig. S2). High similarity in the country network characterise articles dedicated to the same area (USA, China, South Africa) and articles dedicated to comparative studies (like the most extensive of that kind: Rosen & Resnick, 1980; Soo, 2005). It is expected that these pairs of articles should report similar Zipf estimations, since they are performed on the same set of observations. The two densest clusters are composed of studies reporting Zipf's estimates exclusively for American (in orange) and Chinese (in blue) cities.

The "decades" network

The "decades" network represents the similarity between corpus articles based on the decades on which they perform empirical estimations of Zipf's exponents (Fig. S3). High similarity characterise articles dedicated to the same period. The densest clusters are composed by corpus articles reporting Zipf's estimates exclusively for a single decade (such as Cameron, 1990 or Krugman, 1996).

The "city definition" network

The "city definition" network represents the similarity between corpus articles based on the city definition used to collect city populations (municipalities, agglomerations or metropolitan areas mostly) on which empirical estimations of Zipf's exponents are performed (Fig. S4). This network is polarised by the use of one or more city definitions in the corpus article. The larger cluster (in orange) unfortunately reflects the fact that most city size distributions are analysed within improper urban delineations (the 'city proper' or municipal boundaries), as their shape and defining principles vary greatly across countries while they tend to stay fixed over time whereas cities expand spatially and functionally.
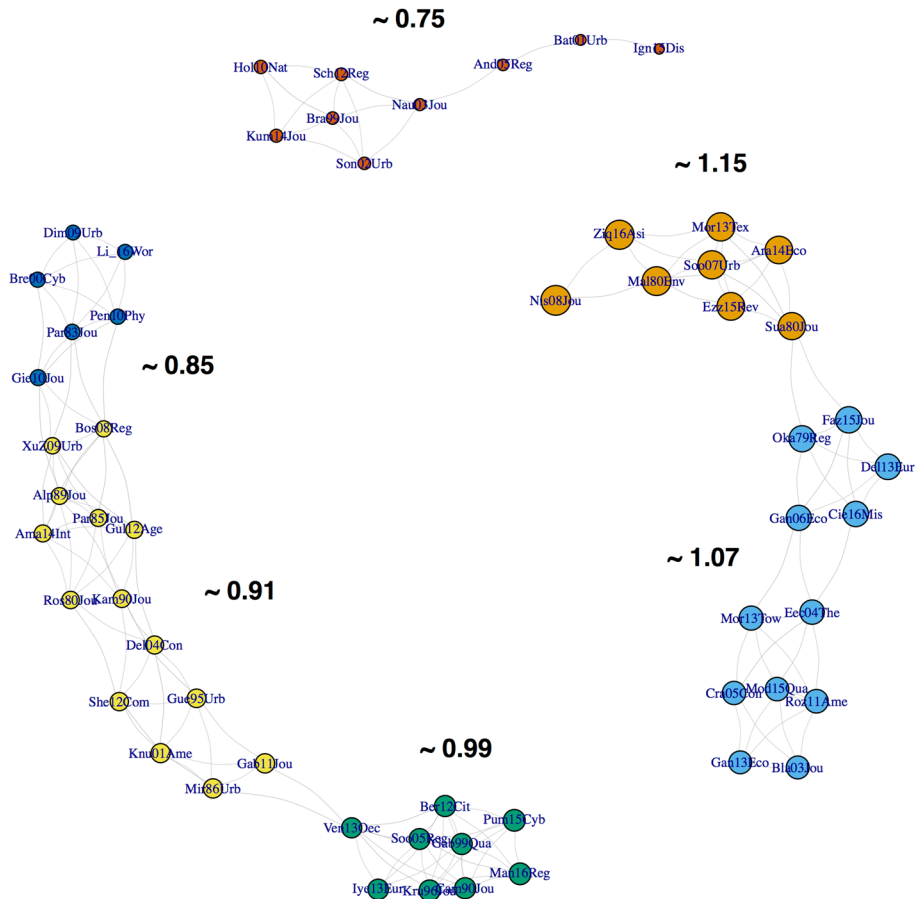
Fig. 7 Similarity network of corpus articles by average value of estimates reported (cut-off at −0.025). The size of nodes reflects the average value of estimates reported in the article and the numbers in black correspond to the average value reported for the community

The "alpha" network

The "alpha" network represents the similarity between corpus articles based on the distribution of empirical estimations of Zipf's exponents (alpha expressed under the Lotka form, or 1/alpha expressed in the Pareto form) they report. I choose to model two aspects of this distribution: the average value of alpha reported on the one hand, and its standard deviation on the other hand. Additionally, I use the number of estimates reported as an extra control. To construct the "mean alpha" network, I compute the average value of alpha estimates a_i per study i and the average value of alpha estimates a for the entire sample (1962 estimates in total). I then compute a distance da_{ij} between studies as follows:

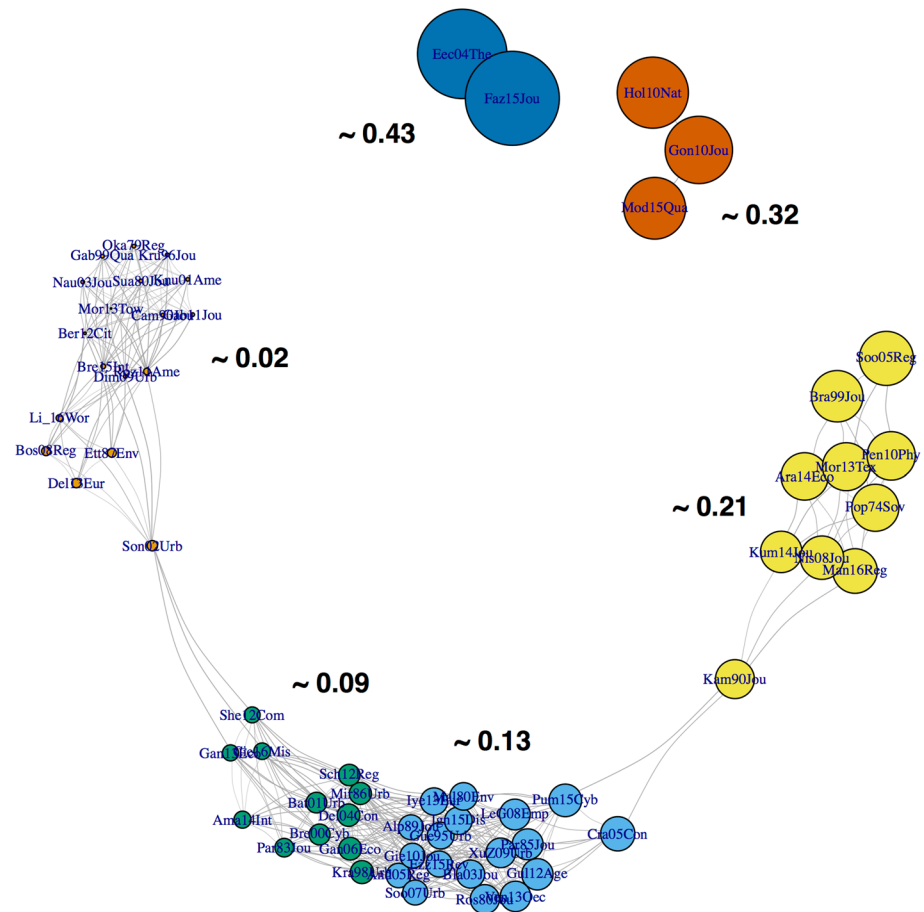$$da_{ij} = \left| a_i - a_j \right| / a, \quad \text{with } i \neq j$$

Fig. 8 Similarity network of corpus articles by standard deviation of estimates reported (cut-off at −0.1). The size of nodes reflects the standard deviation of estimates reported in the article and the numbers in black correspond to the average standard deviation for the community

The smaller this distance the more studies i and j report Zipf estimates close in value. To transform this distance into a similarity, I simply multiply da_{ij} by −1. The network emerging from this similarity is therefore organised around groups of studies based on the average values of alpha estimates they report (Fig. 7). At the low end of the spectrum, studies like Holmes and Lee (2010) or Kumar and Subbarayan (2014) report very low values of estimates (0.75 on average for the group in red), which indicates city sizes more evenly distributed than predicted by Zipf. At the other end of the spectrum, studies like Ziqin (2016) or Nishiyama et al. (2008) report high values of estimates (1.15 on average for the group in orange), which reflects highly uneven city size distributions.

To construct the "standard deviation" network, I compute the standard deviation σ^2_i of alpha estimates per study i and the standard deviation of alpha estimates σ^2_a for the entire sample. I then compute a distance $D\sigma^2_{ij}$ between studies as follows:

$$Do^2_{ij} = \left| \sigma^2_i - \sigma^2_j \right| / \sigma^2_a, \quad \text{with } i \neq j$$

The smaller this distance the more studies i and j report a similar dispersion of Zipf estimates. To transform this distance into a similarity, I simply multiply Do^2_{ij} by -1. The network emerging from this similarity is therefore organised around groups of studies based on the average dispersion of alpha estimates they report (Fig. 8). At the low end of the spectrum, studies like Okabe (1979) or Gabaix (1999) report estimates very close to one another (0.02 standard deviation on average for the group in orange), frequently because such studies only report only 1 or 2 estimates. At the other end of the spectrum, studies like Eeckout (2004) or Fazio and Modica (2015) report very dispersed distributions of estimates (0.43 standard deviation on average for the group in dark blue). In these two examples, such dispersion is produced by estimations all made for the USA in 2000 and 2010, but with large variations of truncation points (i.e. the minimum population size to include cities in the sample), from 135 residents to 29,000, which changes the number of places included in the regression from about 156,000 to only 35. As noted in Cottineau (2017), the truncation point for city populations is one of the most important technical specifications with respect to the variation of Zipf's estimates in the literature.

Finally, I constructed a "n alpha" network to control for the number of estimates reported (especially when modelling their dispersion). I computed the number alpha estimates n_i per study and the average value of alpha estimates n in the entire sample. I then computed a distance dn_{ij} between studies as follows:

$$Dn_{ij} = \left| n_i - n_j \right| / n, \quad \text{with } i \neq j$$

The smaller this distance the more studies i and j report a similar number of Zipf estimates. To transform this distance into a similarity, I simply multiplied dn_{ij} by -1. The network emerging from this similarity is shown in supplementary Fig. S3.

Modelling dyad similarities

I run two series of models, one aimed at "explaining" the similarity in mean alpha values reported between corpus articles, and one aimed at explaining their similarity in alpha dispersion. "Explaining" variables for each series of models are similarity measures of the "wording", "citation", "disciplinary", "country", "decades", "city definition" and "n alpha" networks. All variables were centred and scaled prior to modelling. I estimate the coefficients b_k and the residuals e_{ij} by running step-wise OLS regressions.

$$\text{MeanAlpha}_{ij} = b_1\text{Wording}_{ij} + b_2\text{Citation}_{ij} + b_3\text{Discipline}_{ij} + b_4 n\text{Alpha}_{ij}$$
$$+ b_4\text{Country}_{ij} + b_5\text{Decade}_{ij} + b_6\text{CityDef}_{ij} + e_{ij}$$

$$sd\text{Alpha}_{ij} = b_1\text{Wording}_{ij} + b_2\text{Citation}_{ij} + b_3\text{Discipline}_{ij} + b_4 n\text{Alpha}_{ij}$$
$$+ b_4\text{Country}_{ij} + b_5\text{Decade}_{ij} + b_6\text{CityDef}_{ij} + e_{ij}$$

with $i \neq j$, i and j being articles from the corpus.

I also look at interactions between control variables to identify studies which have studied similar national systems in the same decade and under the same definition of cities. These studies should report the most similar values of rank-size estimations.

Table 2 OLS regression of the similarity in average value of alpha reported in the corpus

Similarity in …	Dependent variable				
	similarity in meanAlpha				
	(1)	(2)	(3)	(4)	(5)
wording	0.048**				0.050**
	(0.022)				(0.023)
citation		0.062***			0.069***
		(0.022)			(0.024)
discipline			0.043*		0.015
			(0.022)		(0.024)
nAlpha				−0.053**	−0.046**
				(0.022)	(0.022)
country				0.063***	0.063***
				(0.023)	(0.023)
decade				−0.100***	−0.123***
				(0.022)	(0.023)
cityDef				−0.001	−0.015
				(0.022)	(0.023)
country:decade				−0.014	−0.013
				(0.021)	(0.021)
country:cityDef				0.049**	0.053**
				(0.022)	(0.022)
decade:cityDef				0.037*	0.038*
				(0.022)	(0.022)
country:decade:cityDef				−0.017	−0.014
				(0.021)	(0.021)
Constant	0.015	0.015	0.015	0.013	0.012
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Observations	2016	2016	2016	2016	2016
R^2	0.002	0.004	0.002	0.021	0.030

*$p<0.1$; **$p<0.05$; ***$p<0.01$

Results

The results of the regressions are reported in Tables 2 and 3. Regarding the similarity in mean alpha reported in the corpus (Table 2), I find a confirmation of two of my three initial hypotheses. Although the R^2 are low, the similarity in mean alpha varies positively and significantly with both the similarity in wording and the similarity in citations (models 1 and 2). This means that articles written with a similar set of words and references tend to report similar values of Zipf estimates on average. This interesting feature persists (model 5) even when I account for the similarity in countries, decades and city definitions which the pair of corpus articles studies. As the wording network showed, this could result from a different setup from which the estimation originates. In some articles, the goal is to validate a "law" and the adequacy of one case to the "model". It is thus more probable that such studies report estimates centred around 1, as in the strict version of Zipf's law. On the other

Table 3 OLS regression of the similarity in standard deviation of alpha reported

| Similarity in … | Dependent variable | | | | | |
| | Similarity in sdAlpha | | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)
wording	0.112***					0.116***
	(0.022)					(0.023)
citation		−0.013				−0.006
		(0.022)				(0.024)
discipline			−0.004			−0.009
			(0.022)			(0.024)
nAlpha				0.134***		0.133***
				(0.022)		(0.022)
country					−0.096***	−0.078***
					(0.023)	(0.022)
decade					−0.077***	−0.076***
					(0.022)	(0.023)
cityDef					0.088***	0.083***
					(0.022)	(0.022)
country:decade					0.041*	0.042**
					(0.021)	(0.021)
country:cityDef					0.050**	0.051**
					(0.022)	(0.022)
decade:cityDef					0.050**	0.051**
					(0.022)	(0.022)
country:decade:cityDef					−0.023	−0.016
					(0.021)	(0.021)
Constant	−0.009	−0.009	−0.009	−0.008	−0.014	−0.013
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Observations	2016	2016	2016	2016	2016	2016
R^2	0.012	0.0002	0.00002	0.018	0.027	0.055

hand, articles citing the same pool of references can exhibit a similar interest in validating or challenging the law. The evidence for the similarity in disciplinary references is more mixed, since the significant effect of the simple model 3 disappears when other variables and controls are accounted for. In terms of controls, I find that articles reporting a similar number of estimates tend to differ in mean alpha. This can be the effect of sensitivity analysis studies which explore the effect of threshold values or other specification criteria: they generally report a high number of estimates but their dispersion is such that the average value varies a lot. As expected, studies dedicated to the same set of countries tend to report similar values of estimates on average, however the opposite is true for time periods. The effect of similar city definitions chosen to analyse size distributions was not found significant by itself but appeared positive in conjunction with a similarity in the set of countries and with a similarity in the set of decades studied, as expected.

The distribution of positive residuals (Fig. S4) shows pairs with higher similarity than expected by the model. No obvious pattern seems to govern the association between such

pairs, where luck might play a role. However, negative residuals are driven by three studies whose average estimate value differ from that of all others: Luckstead and Devadoss (2014), Le Gallo and Chasco (2008) and Popov (1974). They report an average value of alpha respectively of 1.91, 1.73 and 1.45. Those are very far away from the expected linear exponent of Zipf's law, which might suggest considering them as outliers for a subsequent meta-analysis.

Regarding the similarity in dispersion (Table 3), I find that only one of my main hypotheses is verified: the more articles are written with similar words, the more similar they are in terms of standard deviation of alphas reported (models 1 and 6). Again, some articles are similar in their attempts at verifying the "law": they are written with mathematical language and tend to report few estimates close in value. Other articles have the goal of exploring the national variation of city size distributions or their sensitivity to technical specifications: they use words like "countries", "spatial" and "comparison" and tend to report a very dispersed set of results. I do not find any significant evidence of covariation between the similarity in reference list and disciplines cited on the one hand and the similarity in alpha dispersion on the other hand. However, the number of estimates is shown to positively influence the similarity in dispersion, since more estimates tends to increase the dispersion on average. Studies which use similar city definitions tend to report similar dispersion. Finally, although the similarity in countries and decades studied is negatively associated with a similarity in dispersion per se, they are positively associated when in interaction with one another and with city definition (model 6). The distribution of residuals (figure S5) exhibits the same properties as that of the previous model: elective similarity between more or less isolated pairs of studies and polarised dissimilarity with a couple of articles, including Luckstead and Devadoss (2014).

Discussion and conclusion

In this article, I have looked at the empirical literature on Zipf's law for cities from a network perspective. As a complement to previous meta-analyses, the present approach has shed light on the scientific text and context mobilized to report on city size distributions. As in Raimbault et al. (2019), it has used textual analysis and citation networks to reflect various proximities between articles of a corpus. The analysis of each network had produced insights about the wording, reference and discipline frameworks mobilised by the empirical literature on Zipf's law for cities. Their use as regression variables in a model of the similarity in the distribution of estimates reported has shown that wording is important in both cases, whereas similar citation patterns mostly impact the average value of Zipf's estimate reported. These results point towards the existence of a combined publication and reporting bias in the multidisciplinary literature on Zipf's law for cities. Indeed, despite the fact that city sizes could be observed identically by all researchers using the same specifications, I find that different authors tend to delineate cities and report Zipf estimations differently depending on the field they come from, which translates into the words they use and the citations they mobilise.

The contribution of this paper to meta-analyses is two-fold. Firstly, using the citation networks of studies included in a meta-analysis, it allows to identify gaps and outliers in the corpus and potentially overlooked articles. These have appeared respectively when looking at the most cited external references and the model residuals. For instance, the article of Eaton and Eckstein (1997) is one of the most externally cited reference to report

empirical estimations of Zipf's law. It was initially rejected from the corpus (Cottineau, 2017) because the estimation included instruments. The present analysis suggests that relaxing this criterion could allow its inclusion as a major reference in the field. Symmetrically, the analysis of model residuals has shown that some very atypical studies drive a large share of the difference in mean values and dispersion used in the meta-analysis, suggesting that removing them as outliers could provide clearer results. Secondly, the data and code of the present study has been made open on Github, including an R notebook with all visualisations, in order to be reused by the community.

Although this article does not close the debate on city size distribution, it has tried to reveal a newer aspect of a literature in rapid development: the fact that it mixes studies from different disciplines, with very different aims and methods, potentially characterised by reporting biases. What seems quite obvious from the corpus is also the fact that Zipf's law estimation is a research field where many authors contribute at one point of their scientific career in urban studies, economics or physics, but mostly is not a dominant object of individual research per se. A further point of inquiry in the reflexive meta-analysis could thus be to trace various authors' contribution to the empirical Zipf literature as part of their own scientific topic trajectory (Zeng et al., 2019). However, it is not obvious at this point to which extend it would help provide guidelines for rigorous analysis of city size distribution.

Declarations

Conflict of interest The author declares no conflict of interest.

References

Arribas-Bel, D., Gracia, F. S., & Ximénez-de-Embún, D. (2012). Kangaroos, cities and space: A first approach to the australian urban system. *Region Et Development, 36*, 165–187.

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

Batty, M. (2001). Polynucleated urban landscapes. *Urban Studies, 38*(4), 635–655. https://doi.org/10.1080/00420980120035268

Berry, B. J. L., & Okulicz-Kozaryn, A. (2012). The city size distribution debate: Resolution for US urban regions and megalopolitan areas. *Cities, 29*, S17–S23.

Bretagnolle, A., Mathian, H., Pumain, D., & Rozenblat, C. (2000). Long-term dynamics of European towns and cities: towards a spatial model of urban growth. *Cybergeo: European Journal of Geography*. https://doi.org/10.4000/cybergeo.566

Bretagnolle, A., Delisle, F., Mathian, H., & Vatin, G. (2015). Urbanization of the United States over two centuries: an approach based on a long-term database (1790–2010). *International Journal of Geographical Information Science*, pp 1–18.

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

Chang, C. Y., Lee, D. C., Tang, K. Y., & Hwang, G. J. (2021). Effect sizes and research directions of peer assessments: From an integrated perspective of meta-analysis and co-citation network. *Computers & Education*, 164, 104123.

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

Cottineau, C. (2017). MetaZipf. A dynamic meta-analysis of city size distributions. *PloS one*, *12*(8), e0183919.

Cottineau, C. (2020). MetaMetaZipf. What do analyses of city size distributions have in common?. arXiv preprint arXiv:2008.11473.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems, 1695*(5), 1–9.

Dimou, M., & Schaffar, A. (2009). Urban hierarchies and city growth in the Balkans. *Urban Studies*. http://usj.sagepub.com/content/early/2009/09/04/0042098009344993.short

Duyx, B., Urlings, M. J., Swaen, G. M., Bouter, L. M., & Zeegers, M. P. (2017). Scientific citations favor positive results: A systematic review and meta-analysis. *Journal of Clinical Epidemiology, 88*, 92–101.

Eaton, J., & Eckstein, Z. (1997). Cities and growth: Theory and evidence from France and Japan. *Regional Science and Urban Economics, 27*(4–5), 443–474.

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

Fazio, G., & Modica, M. (2015). Pareto or log-normal? Best fit and truncation in the distribution of all cities*. *Journal of Regional Science, 55*(5), 736–756. https://doi.org/10.1111/jors.12205

Feinerer I., Hornik K. (2019) Artifex Software Inc., 2019, "tm: Text Mining Package", R package, version 0.7–7, https://CRAN.R-project.org/package=tm

Fetscherin, M., & Heinrich, D. (2015). Consumer brand relationships research: A bibliometric citation meta-analysis. *Journal of Business Research, 68*(2), 380–390.

Gabaix, X. (1999). Zipf's law for cities: anexplanation. *Quarterly Journal of Economics*, pp 739–767

Gabaix, X., & Ibragimov, R. (2011). Rank − 1/2: A simple way to improve the OLS estimation of tail exponents. *Journal of Business & Economic Statistics, 29*(1), 24–39. https://doi.org/10.1198/jbes.2009.06157.

Gabaix, X., & Ioannides, Y. M. (2004). The evolution of city size distributions. *In Handbook of Regional and Urban Economics, 4*, 2341–2378.

Geyskens, I., Krishnan, R., Steenkamp, J. B. E., & Cunha, P. V. (2009). A review and evaluation of meta-analysis practices in management research. *Journal of Management, 35*(2), 393–419.

Gibrat, R. (1931). *Les inégalités économiques*. Libraire du Recueil Sierey.

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

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

Holmes, T. J., & Lee, S. (2010). Cities as Six-by-Six-Mile Squares: Zipf's Law? *NBER*, pp 105–131

Ignazzi, C.A. (2015). Coevolution in the Brazilian system of cities, Dissertation, Université Paris 1, Panthéon-Sorbonne, Paris, France

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and Software Technology, 51*(1), 7–15.

Krugman, P. (1996). Confronting the mystery of urban hierarchy. *Journal of the Japanese and International Economies, 10*(4), 399–418.

Kumar, G., & Subbarayan, A. (2014). The temporal dynamics of regional city size distribution: Andhra Pradesh (1951–2001). *Journal of Mathematics and Statistics, 10*(2), 221.

Latour, B., & Woolgar, S. (1986). *Laboratory life*. Princeton University Press.

Le Gallo, J., & Chasco, C. (2008). Spatial analysis of urban growth in Spain, 1900–2001. *Empirical Economics, 34*(1), 59–80.

Lösch, A. (1940). Die räumliche Ordnung der Wirtschaft: eine Untersuchung über Standort. *Gustav Fisher, Jena*.

Losse, M., & Geissdoerfer, M. (2021). Mapping socially responsible investing: A bibliometric and citation network analysis. *Journal of Cleaner Production*, 126376.

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

Lumley, T. (2002). Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine, 21*(16), 2313–2324.

Mohamed Shaffril, H. A., Samsuddin, S. F., & Abu Samah, A. (2021). The ABC of systematic literature review: The basic methodological guidance for beginners. *Quality & Quantity, 55*(4), 1319–1346.

Moro, S., & Santos, R. (2013). The characteristics and evolution of the Brazilian spatial urban system: empirical evidences for the long-run, 1970–2010(Textos para Discussão Cedeplar-UFMG No. 474). Cedeplar, Universidade Federal de Minas Gerais. Retrieved from http://econpapers.repec.org/paper/cdptexdis/td474.htm

Nishiyama, Y., Osada, S., & Sato, Y. (2008). OLS estimation and the t test revisited in rank-size rule regression. *Journal of Regional Science, 48*(4), 691–716.

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

Okabe, A. (1979). An expected rank-size rule: A theoretical relationship between the rank size rule and city size distributions. *Regional Science and Urban Economics, 9*(1), 21–40.

Pareto, V. (1897). Cours d'économie politique, Vol. 2. *Lausanne*, Rouge.

Parr, J. B. (1985). A note on the size distribution of cities over time. *Journal of Urban Economics, 18*(2), 199–212.

Parr, J. B., & Jones, C. (1983). City size distributions and urban density functions: Some interrelationships. *Journal of Regional Science, 23*(3), 283–307. https://doi.org/10.1111/j.1467-9787.1983.tb00991.x.

Popov, V. R. (1974). Investigation of the system of urban places of Crimea Oblast. *Soviet Geography, 15*(1), 19–23.

Pumain, D., Swerts, E., Cottineau, C., Vacchiani-Marcuzzo, C., Ignazzi, A., Bretagnolle, A., Delisle, F., Cura, R., Baffi, S. (2015). Multilevel comparison of large urban systems. *Cybergeo: European Journal of Geography*. https://doi.org/10.4000/cybergeo.26730

Radler, V. M. (2018). 20 Years of brand personality: A bibliometric review and research agenda. *Journal of Brand Management, 25*(4), 370–383.

Raimbault, J., Chasset, P. O., Cottineau, C., Commenges, H., Pumain, D., Kosmopoulos, C., & Banos, A. (2019). Empowering open science with reflexive and spatialised indicators. *Environment and Planning B: Urban Analytics and City Science*. https://doi.org/10.1177/2399808319870816.

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

Schaffar, A., & Dimou, M. (2012). Rank-size city dynamics in China and India 1981–2004. *Regional Studies, 46*(6), 707–721. https://doi.org/10.1080/00343404.2010.521146.

Shi, J., Gao, Y., Ming, L., Yang, K., Sun, Y., Chen, J., Shi, S., Geng, J., Li, L., Wu, J., & Tian, J. (2021). A bibliometric analysis of global research output on network meta-analysis. *BMC Medical Informatics and Decision Making, 21*(1), 1–12.

Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika, 42*(3/4), 425–440.

Singer, H. W. (1936). The" courbe des populations." A parallel to Pareto's Law. *The Economic Journal*, *46*(182), 254–263.

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

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

Suarez-Villa, L. (1980). Rank size distribution, city size hierarchies and the Beckmann model: Some empirical results. *Journal of Regional Science, 20*(1), 91–95.

Webster C. (2006) Editorial. Ranking planning journals. *Environment and planning B, 33*, 485–490. https://doi.org/10.1068/b3304ed

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *J. Plan. Educ., 39*(1), 93–112.

Zeng, A., Shen, Z., Zhou, J., Fan, Y., Di, Z., Wang, Y., & Havlin, S. (2019). Increasing trend of scientists to switch between topics. *Nature Communications, 10*(1), 1–11.

Zhang, C., & Guan, J. (2017). How to identify metaknowledge trends and features in a certain research field? Evidences from innovation and entrepreneurial ecosystem. *Scientometrics, 113*(2), 1177–1197.

Zipf, G. K. (1941). *National Unity and Disunity*. Principia Press.

Zipf, G. K. (1949). *Human behavior and the principle of least effort* (p. 573). Hafner.

Ziqin, W. (2016). Zipf law analysis of Urban Scale in China. *Asian Journal of Social Science Studies, 1*(1), 53.