ARTICLE

FIRM SIZE DISTRIBUTION IN THE CENTRAL EUROPEAN CONTEXT

Štamfestová, P., Sobíšek, L., Hnilica, J.

Petra Štamfestová (corresponding author) / Prague University of Economics and Business, Faculty of Business Administration, Department of Strategy, W. Churchill Sq. 1938/4, Prague, Czech Republic. Email: petra.stamfestova@vse.cz

Lukáš Sobíšek / Prague University of Economics and Business, Faculty of Business Administration, Department of Strategy, W. Churchill Sq. 1938/4, Prague, Czech Republic. Email: lukas@sobisek.cz

Jiří Hnilica / Prague University of Economics and Business, Faculty of Business Administration, Department of Strategy, W. Churchill Sq. 1938/4, Prague, Czech Republic. Email: hnilica@vse.cz

Abstract

This article analyses the distribution of firms by size in six selected countries in 2012 and 2017. Estimates are always made for the whole economy and two subgroups of firms. We compare the Visegrad Four countries (Czech Republic, Hungary, Poland and Slovakia) with similar economic activities and two neighbouring economically more developed countries (Germany and Austria). As the main objective of the article, we describe the distribution of firms by verifying the validity of Zipf's law in selected economies and their sectors. The results confirm the positively skewed distribution of company sizes measured by sales revenues, but Zipf's law does not apply to the distribution of all companies by the magnitude of sales in the whole economy (or in an economic subgroup), but only to sections in the right tail of the distribution (companies with higher turnover within the whole economy), which is in line with numerous research studies.

Implications for Central European audience: This article attempts to contribute to the state of the art of firm size distribution in two ways. Firstly, it reviews a wide range of research studies looking into the existence of power laws in firm size distribution. Secondly, it analyses the distribution of firms by size in six selected Central European countries in 2012 and 2017, where four sample countries were burdened with interrupted business activities during a communist regime (Czech Republic, Hungary, Poland and Slovakia), and two countries are their neighbours not exposed to changes in the political regime (Germany and Austria).

Keywords: firm size distribution; power law; Zipf´s law

JEL Classification: L11, L25

Introduction

The firm size distribution in economies is remarkably positively skewed, which means that the average size of all companies is affected by a smaller number of very big firms despite the fact that small companies significantly outnumber the larger ones. This goes against concepts of standard macroeconomic theory, which assumes that activities of a large number of firms tend to cancel each other out so that the impact of one individual company matters very little. This assumption breaks down if the distribution of firm sizes is fat-tailed, as documented empirically in many studies (see below).

In light of this evidence, a concept of "granularity" has recently entered the scene. This concept reflects the fact that economies are populated by a few large companies (the big "grains") that coexist with many smaller companies (Gabaix, 2011). Such a distribution of firm sizes is modelled by power laws, a functional relation between two quantities, where a relative change in one quantity yields a proportional relative change in the other, regardless of the initial size of such quantities, meaning that these quantities vary as a power of one another. Economic fluctuations in such environments are not due to small shocks reverberating across all firms; instead, they originate from the incompressible "grains" of economic activity – the large firms – whose idiosyncratic shocks generate material aggregate shocks hitting all the firms (and, of course, GDP).

Phenomena whose distribution is governed by the power law appear widely in both natural and social contexts: population size in cities, frequency of words used in texts, book sales, numbers of telephone calls and emails received and sent, magnitude of earthquakes, diameter of moon craters, intensity of solar flares, size of wars, mass extinctions of species, stock markets and many other (Axtell, 2001).

In many studies on firm size distribution, the power law exponent of firm size distribution oscillates around one, which is known as Zipf's law. This law applies regardless of how the size of the company is measured, whether by means of sales (which seems to be the most frequent form), numbers of employees or the company's capital. To add complexity, Zipf's law does not usually occur across the entire width of the distribution and is only observed from a certain minimum value up to the end of the distribution, e.g., regarding the number of citations of scientific articles, the power law can only be observed in articles with a hundred and more citations (Newman, 2005). Power law distributions cannot be easily characterized by means, variances or higher-order moments because they may become infinite (Mantegna & Stanley, 2000). For an instance of Zipf's law, the aggregate volatility decays only at the rate $logN$, whereas under finite variance, it is $\sqrt{N}$, where N stands for the number of firms in a particular economy (Gabaix, 2011).

First, the article reviews a wide range of existing research studies looking into the existence of power laws in firm size distribution. The body of this article succinctly summarizes their main findings and explanations behind the findings. Data, methodology and results of these studies are presented in detail in Table 1 in the Annex.

Besides, the article makes an original estimation of power law exponents and investigates the applicability of Zipf's laws in the firm size distribution in Central Europe, a region with post-communist countries. The empirical verification of Zipf's law is conducted for Czech, Slovak, Polish and Hungarian (post-communist countries) and for Austrian and German companies (with no interruption of business activity) for the period 2012-2017 at the level of all companies and at the level of selected economic sectors. Our results are further confronted with findings of other studies. The power law exponents are empirically estimated by fitting the OLS linear regression models and applying a smoothing approach.

1 Literature review

The first journal articles focused on the analysis of company size distribution date back to the middle of the twentieth century; one of the first articles on this topic was by Simon and Bonini (1958). They confirmed that the observed distribution of firms in terms of their size is sturdily positively skewed to the right, no matter if measured by sales, assets, numbers of employees or even value-added or profits. Such distributions of companies apply at the level of both individual industries and the entire economy and are robust over time. Researchers at that time attempted to model the dataset by means of lognormal distribution or the Yule distribution (Axtell, 2001).

More recent research departs from the above-mentioned distributions. Stanley et al. (1995) conducted their research across a spectrum of 4,071 US manufacturing companies from the US Compustat database. Company sales were first fitted to the lognormal distribution; however, after the construction of Zipf's plot (double logarithmic plot of sales vs. rank), it turned out that the group of the largest companies showed a lower amount of sales than what would correspond to the lognormal distribution. The deviations from the lognormal distribution were statistically significant. Other papers (e.g., Okuyama et al., 1999; Axtell, 2001), inspired by Harvard linguistics professor George Kingsley Zipf, support these findings known as Zipf’s law, meaning that the power law exponent is close to 1. Okuyama et al. (1999) approximated company size by profit before tax and concluded that Zipf's law is suitable for describing company size distribution. They used Japanese, Italian and US companies in their datasets. Estimated exponents at the aggregate level were around one for the Japanese and Italian companies, while they were around 1.4 for the US companies. Company size distribution at the industry level was analysed only for Japanese companies, where the conclusion was that the values of power law exponents are in the range of 0.7–1.2. As a specific example, they present the results of an analysis of construction companies and companies providing electrical products. In the first group mentioned, the power law exponent was estimated at the level of 1.13, while in the second group, it was at the level of 0.72. Also, a large part of recent studies focuses on examining the distribution of companies only at the industry level. For example, the power law distribution has described the tourism industry in Germany and Italy (Provenzano, 2014) or oil and natural gas production in Texas (Balthrop, 2016). One of the latest articles analysing company size distribution at the industry level was written by Balthrop (2021). The dataset of US trucking firms can be neatly described by the power law (Balthrop, 2021).

A breakthrough and widely cited study (Axtell, 2001) looked for a theoretical power law exponent in US companies for ten years from 1988 to 1997. The conclusion was that the value of the exponent is always around one, and the validity of Zipf's law cannot be denied. More than five and a half million US companies that were in Standard & Poor’s COMPUSTAT commercial database entered the analysis, and the size of the companies was measured by the number of employees.

Gaffeo et al. (2003) examined the average distributed sizes of companies in the group of the seven most economically developed countries in the period from 1987 to 2000. The size of companies was measured by sales, capital and debt. The analysis supports the existence of the power law; however, only when measuring firm size by sales. The other two cases are in line with the power law, but this distribution is less uniform than what would correspond to Zipf's law. Another interesting finding of this study is that firms are distributed more equally during recessions than during expansions (Gaffeo et al., 2003).

A year later, a paper was presented by Fujiwara et al. (2004), who studied the distribution of firms in European countries in 1992–2001. Approximately 260,000 companies from 45 European countries entered the analysis. Nevertheless, the authors present in more detail only results for France and Italy. Again, based on estimates of power law coefficients that ranged around the desired value of one, the authors declared that the analysed data correspond to Zipf's law. In the article, the authors also confirmed the validity of Gibrat's law, which states that growth of companies is independent of their size (Fujiwara et al., 2004).

Three years later, Hernández-Pérez et al. (2006) published research analysing company size distribution in developing countries. This is basically one of the first studies with a focus beyond the developed countries. With annual sales, ten developing countries were subjected to the analysis (Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Mexico, Mozambique, Peru and Venezuela). The authors checked whether the data matched finite size-scaling (FSS), whose function has two components: a power law and an exponential. After this analysis, the authors stated that the size distribution of larger companies is governed by the power law and, according to them, the results are also in accordance with the above-mentioned study by Gaffeo et al. (2003), which showed that the average company size distribution in the G7 countries corresponds to the power law (Hernández-Pérez et al., 2006).

Gabaix and Landier (2008) examined company size distribution in the context of CEO salaries with an assumption that the remuneration of CEOs depends on the size of the company CEOs work at (Gabaix & Landier, 2008). The dataset covered all US traded companies in the period 1978–2004. Again, Zipf's law was confirmed.

Zhang et al. (2009) published a study verifying the validity of Zipf's law on the 500 largest Chinese companies from 2002 to 2007. Podobnik et al. (2010) examined the size distribution of 2,737 medium-sized and large US companies that went bankrupt in the period 1990–2009 and measured their size by the book values of their assets and by market capitalization. The study did not reject the validity of Zipf's law, and the authors also added that companies with a rank value higher than 500 are beginning to deviate from Zipf's law due to the effect of finite size.

A year later, di Giovanni et al. (2011) intended to show that the size distribution of firms is affected by the international trade of scrutinized companies. The authors approximated the size of the companies by sales and employees. Their dataset first considered all French companies from the year 2006. When authors compared the size distribution of exporting and non-exporting companies, non-exporting (exporting) firms had a larger (smaller) exponent than the one found in the whole study dataset. The results were driven by the fact that exporters are, on average, larger than non-exporters. Export firms are bigger on average since all big firms come to all markets over time. Another reason is that fixed export costs have a certain level, so each company above a certain level of productivity considers exports beneficial (di Giovanni et al., 2011). Some articles have studied the size distribution of exports at the product level. One of the latest studies was published by del Rosal (2016). He studied EU country exports by product and found that larger and richer countries have more diversified exports (higher Pareto exponent estimates). This is due to the fact that such countries have a higher R&D intensity (del Rosal, 2016).

In the same year, Kang et al. (2011) looked for a description of the inequality in the size of non-financial tradable Korean firms in 1987, 1997 and 2007 on the stock markets KOSPI and KOSDAQ. The size of 703 firms was measured by the amount of sales, the number of employees, the amount of assets and the amount of capital. Given that the authors examined company size distribution over several decades, they were also able to see whether the crisis in Korea in the period 1997–1999 had an impact on this size distribution. The analysis results confirmed the power law distribution and showed that the values of the power law exponent had increased before the crisis and decreased after the crisis, indicating that the degree of inequality in the size of Korean companies had been higher before the crisis than after the crisis. The coefficient of determination was high and statistically significant, with a significance level of 1% for all the estimated power law exponents. Therefore, the firm size in South Korea follows Zipf's law (Kang et al., 2011).

Spanish manufacturing firms in 2001 and 2006 were analysed by Segarra and Teruel (2012). The authors confirmed the largely right-skewed distribution; however, with a varying power law exponent, implying that there is no constant "power-law" between firm size and firm rank, regardless of the variable, and that it changes with the sample size. Small samples of large firms yielded higher coefficients than large samples that also included smaller firms. The relation between the power law exponent and the sample size was superlinear for the small sample size cases. So these firms were found to exceed the size assumed by Zipf's law, and to be older as well. On the contrary, young firms were found to be smaller than Zipf's law suggests. The superlinear relation in the small samples was explained by the fact that – the biggest firms are more homogenous, which is the conclusion of similar savings from size, diversity level and productivity. Therefore, the sizes of these firms do not vary much (Segarra & Teruel, 2012).

Giovanni and Levchenko (2013) again analysed firm size distribution, this time in 44 countries. The estimated power law exponent varied across countries and ranged between 0.69 and 1.18. Firm size was measured in a standard manner by sales, and Zipf's law approximated the size distribution of firms in most of the countries studied (di Giovanni & Levchenko, 2013). The median of the coefficient of determination (R2) for individual estimates of coefficients was 0,99, and the highest power law coefficient did not differ significantly from minus one.

Gao et al. (2015) studied the temporal evolution of the size distribution of China's listed companies over the period 2001-2013. Just like Zhang et al. (2009), they focused on the largest firms in the economy and estimated the power law coefficient only for companies located in the upper tail of the distribution. Surprisingly, they observed a decrease of the coefficients from 2001 to 2008, and then they fluctuated at the lowest level after 2008, implying that the firm size inequality of China's listed companies continuously increased during these years. They explained this phenomenon by the fact that large firms most likely grow faster than small ones.

A year later, Pascoal et al. (2016) published an article also contributing to this issue. The essence of the presented analysis was to capture the time development of the size distribution of Portuguese firms as measured by annual sales and total assets from 2006 to 2012. Again, Zipf's law was confirmed.

Recent studies include a master thesis from Kazakhstan assessing the validity of power law in local firms and cities (Matayev, 2018). The author analysed approximately 67,000 companies from all industries and regions combined, measuring their size in terms of the number of employees. By testing the existence of power law, there was no evidence of it. He explained this by saying that in Kazakhstan, small companies grow faster than large ones, and there is therefore a discrepancy with Gibrat's law. This explanation was further supported in a study by Duparcq and Konings (2016), who found that the distribution of firms in Kazakhstan is less positively skewed than in most market economies; therefore, the number of smaller companies exceeds the number of large ones less dramatically.

Da Silva et al. (2018) drew upon the above-mentioned granularity hypothesis and tested it on data sent from the top 1000 Brazilian companies for the year 2015. The power law exponent was approximated to one (1.070 ± 0.015), i.e., roughly Zipf’s law.

Wagner and Weche (2020), inspired by the research of Gabaix (2011), tested the granularity hypothesis as well. Gabaix concluded that 100 of the biggest firms in the USA explain about one-third of variations in output growth (Gabaix, 2011). In contradiction, Wagner and Weche (2020) demonstrated that idiosyncratic shocks in the largest firms seem not to be important for an understanding of the aggregate volatility of the German economy. This evidence is in contrast with findings for other countries, and it differs from earlier results for parts of the German economy.

Most of the published literature largely aims to estimate the description of the variability of firm size, i.e., to approximate it with distribution functions on the basis of some empirical evidence. Another stream of research goes further and attempts to explain the power law phenomenon, i.e., derive and describe the elementary processes (principles) that could generate the observed forms of variability, though mostly mathematically rather than economically. Mathematical considerations are behind most of the existing theories about the origin of power distribution, which may be the reason. Probably one of the best-known principles leading to power law distributions is Gibrat's law, which is based on the assumption of random multiplicative growth of firms in determining the minimum size of firms (Gibrat, 1931). According to this theory, the growth of companies does not depend on their size and results from a series of random shocks. The growth of firms can be considered a success, and success is understood here as a random phenomenon, so the growth of forms, according to Gibrat's law, does not depend on the current size of firms (Gibrat, 1931). At the same time, it must be perceived that the number of successful firms is limited, and the final distribution of firms in an economy is established in a form that can be described using the power function.

In 2015, Acemoglu and Cao (2015) published an article where they extended the standard endogenous growth model by assuming the implementation of an innovation process in firms, which is then perceived as the basic source of the final Pareto distribution of firms with an exponent value around one, i.e., in the form of Zipf's law. The authors assumed that incumbents undertake innovations to improve their products while entrants engage in more radical innovations to replace and outperform incumbents. According to this theory, the growth of companies is accelerated by the entry (establishment) of new firms, which creates pressure to increase the productivity of existing firms, and ultimately the final distribution of firms is established in a form that can be described in terms of the power law, especially in sectors that are characterized by high costs of imitation. Of course, the model also assumes the possible departure of the incumbents in the event of failure. However, the difficulty of imitating and catching up with incumbents is also indicated by the small size of a large number of participants (Acemoglu & Cao, 2015).

Luttmer (2007) saw the reason for the size distribution of firms according to the power law in the model of balanced growth in the economy. The essence of growth here is mainly the increase in firms' productivity, selection of successful companies and the assumption that new entrants are trying to imitate incumbents. If new entrants can imitate incumbents without major difficulties and if barriers to entry are low, growth in the economy is rapid. Reducing the costs of entry speeds up the rate of economy-wide experimentation and raises the growth rate of the economy (Lutmer, 2007). Otherwise, when entry of new entrants is difficult, for example, due to high costs of initial capital or specific knowledge, or imitation of existing companies is also difficult, the distribution of firms is established according to Zipf's law. The large number of small firms proves that it is difficult to imitate successful and established companies (Luttmer, 2007). A small number of successful companies are then formed into the tail of the distribution, according to Zipf's law, if the entry of new firms into the economy is difficult. The Pareto tail of the size distribution of enterprises is therefore based on the random growth of productivity of successful companies, or it can also be based on the random accumulation of capital (Luttmer, 2011). In 2007, Rossi-Hansberg and Wright introduced another random growth model or, more precisely, mechanisms that lead to it (Rossi-Hansberg & Wright, 2007). The theory emphasizes the role of the accumulation of specific human capital in shaping the final size distribution of firms.

Other theories and some older theories explaining the origin of distributions in the form of power functions include the theory of superstars. This phenomenon was explained by Rosen (1981). In economies, companies are constantly competing to obtain the best possible resources (for example, talent in the form of human capital) to enable the highest possible increase in the firms' profits. This can be associated with the so-called preferential binding process, where in general new entities tend to connect more with existing and established entities, i.e., for example, larger and more successful companies attract rarer talents and resources, the most prestigious companies are more likely to gain new customers, just like larger cities attract more people or more cited scientific articles attract more researchers, etc. This phenomenon is also referred to as Matthew effect. This effect brings the distribution closer to manifesting the power law.

2 Data description

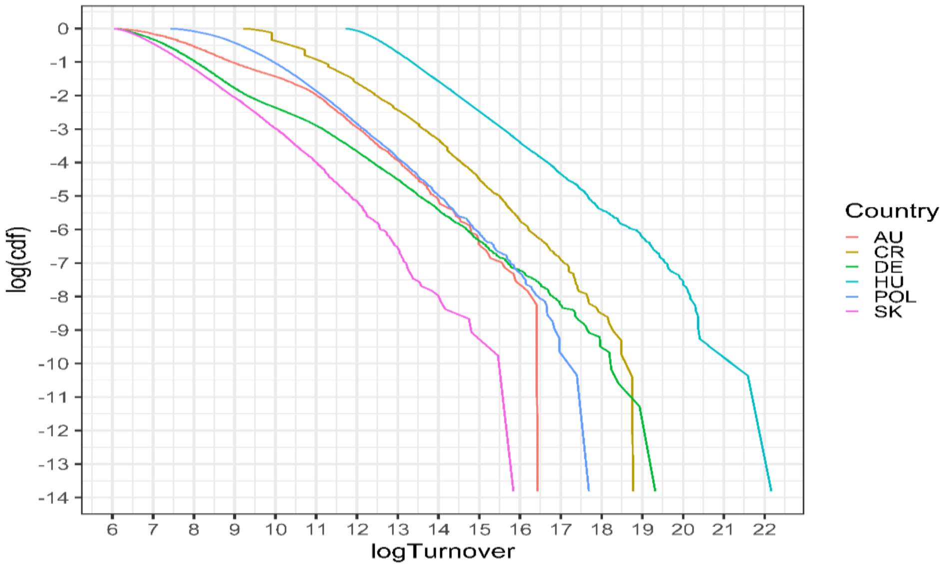
Bureau van Dijk's Amadeus database (Bisnode Czech Republic, 2019) was used to identify enterprises in selected economies. From this database, we selected available data from 2017 and data five years older than 2012 (for sensitivity analysis). As the first search criterion, all active enterprises in the given economy were identified, and in order to eliminate extremely small companies, the minimum size of companies was measured according to the size of sales. It was set at CZK 10 million in the Czech Republic and in other countries using the same cash equivalent in the given currency as of the search date (4 July 2019). This condition of the minimum amount of sales applied to the entire available period in the database used from 2012 to 2017.

Estimates are always made for the whole economy and two subgroups of companies. The Visegrad Four countries (Czech Republic, Hungary, Poland and Slovakia) with similar economic activity and two neighbouring economically more developed countries (Germany and Austria) are compared. This statement can be substantiated by looking at the development of production and the unemployment rate in individual countries. These two indicators are commonly used to measure cyclical changes. The development of changes in real GDP and the unemployment rate in V4 has been monitored and documented, for example by Mura et al. (2020) since 2008. When comparing the absolute values of the real GDP in the V4 countries over time, Poland's superiority over the other three countries can be traced; however, when compared with the values of the real GDP of Germany, all the four countries lag behind significantly. In fact, Germany's real GDP in millions of EUR exceeds the real GDP of all V4 countries combined more than three times (Eurostat, 2023).

We have chosen the subgroups so that each subgroup represents a different group of economic entities and, at the same time, is represented by a sufficient number of companies in order to estimate the PLEE (power law exponent estimate) by the smoothing approach (see Methodology). One subgroup consists of all reported, medium-sized companies from sector C – MANUFACTURING with 50–249 employees. The second group consists of all reported small companies from sector G – WHOLESALE AND RETAIL TRADE; REPAIR AND MAINTENANCE OF MOTOR VEHICLES with 10–49 employees.

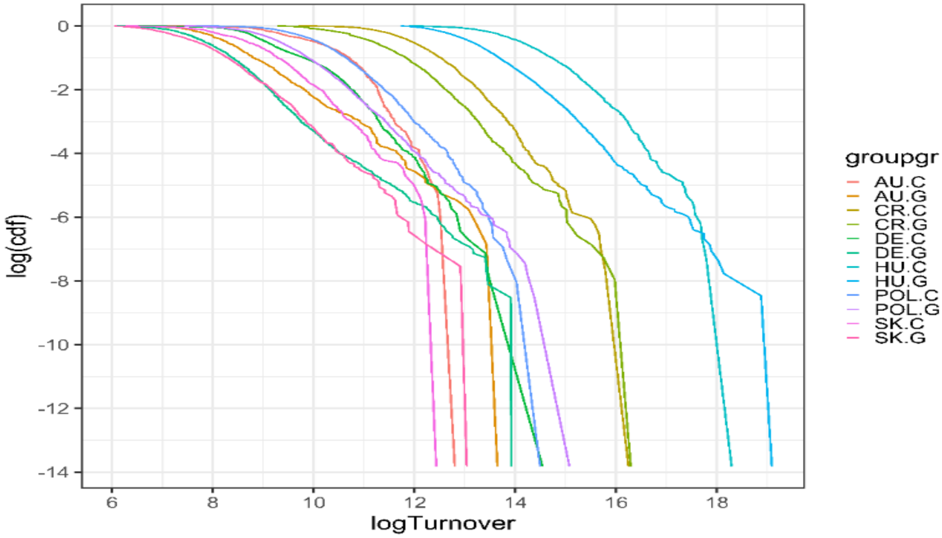
Figure 1 shows the CDF (cumulative distribution function) distribution of sales of companies in 2017 on log scales, and Figure 2 shows the distribution in two selected industries in the same year.

Figure 1 | Distribution of ln(turnover) in 2017 in individual countries for the whole economy



Source: authors

Figure 2 | Distribution of ln(turnover) in 2017 in individual countries for sectors C and G



Note: AU = Austria, CR = Czech Republic, DE = Germany, HU = Hungary, POL = Poland, SK = Slovakia. C or G (after the country abbreviation and dot) indicates sector C or sector G. The x-axis shows the natural logarithm of turnover, and the y-axis is the natural logarithm of CDF.

Source: authors

3 Methodology

The power law is mathematically defined as follows:

$$Pr(X > x) = Cx^{-\xi}, \tag{1}$$

where $C > 0$ and $\xi > 0$ are constants and $Pr(X > x)$ is the probability that the random variable X will attain a value greater than x, which can be empirically interpreted as

$$\Pr(X > x) = \frac{number\ of\ firms\ meeting\ X > x}{total\ number\ of\ firms}. \tag{2}$$

This law should govern the distribution of various characteristics of companies X, where values $x > 0$, such as company size (number of employees), revenues, turnover and others in the selected group (e.g., country).

The aim of the article is to find out whether Zipf's law applies to turnover (X), i.e., estimate ξ and determine whether the estimated (absolute) value of this coefficient (power law exponent estimate, PLEE) is close to one, because in that case, we refer to the exponent law as the Zipf's law.

Di Giovanni et al. (2011) summarized three empirical procedures of regression estimation of the slope of the power law ξ using the OLS (ordinary least square) method: (1) based on the complement CDF (cumulative distribution function) to one, (2) based on the PDF (probability distribution function), and (3) based on rank. In practice, these three OLS estimators, relying on an appropriate modification of the sample or variables used in estimation, delivered remarkably similar results.

In this article, we estimate the distribution of turnover for selected countries and years using the first simplest approach (CDF) based on Axtell (2001), who made direct use of the definition of the power law, i.e., taking the logarithm of Equation (1)

$$ln\big(Pr(X > x)\big) = ln(C) - \xi \ln(x), \tag{3}$$

and the OLS method estimates PLEE (ξ). In our OLS linear regression model

$$Y = b_0 + b_1 X, \tag{4}$$

the conditional mean value *Y* of the outcome is estimated, which is *ln*(proportion of firms with turnover greater than *x*) for the given value *x*, in this case *ln*(turnover), the regression coefficient b_0 is *ln*(*C*), and b_1 is -ξ.

Linear model (4) should not be estimated by means of the OLS method for the whole distribution (all companies in the given country/economic sector) as the relationship between *X* and *Y* is not linear; see Figure 1. For this reason (to ensure linearity), we estimated the PLEE multiple times in a moving manner gradually for the whole group by individual sections (windows, subgroups) with a fixed quantity of 100 (neighbouring) firms. The whole smoothing approach procedure can be summarized in four steps:

1. We sort firms in ascending order (from smallest to largest) according to the values of *ln*(turnover).
2. We take the first 100 firms where a linear relationship between *X* and *Y* exists and estimate the PLEE for them using the OLS method.
3. We move the window by 1 (2^{nd} – 101^{st} firm) and repeat the procedure under 2.
4. We gradually move the window by 1 until we reach the last firm (with the maximum ln(turnover), and repeat the PLEE estimate according to 2.

The sequence of the estimated PLEE (for the given (sub)group) is smoothened in the graphic visualization using the LOESS method. We consider the validity of Zipf's law and present it in tables for estimated (absolute) values of the PLEE in the interval from 0.9 to 1.1 (around 1). All the calculations and graphical visualization of the results were done using the R statistical system, version 3.5.2 (R Core Team, 2018).

4 Results

Table 2 presents the numbers of firms and summarizes the estimated sections in a given economy using the percentile interval and their turnover, where Zipf's law applies (PLEE estimate in the interval 0.9–1.1).

In the whole economy, the PLEE estimates are stable over time (years 2012 and 2017) and in different parts of the distribution, Zipf's law applies in three of the six countries examined, namely Austria (approximately 81–100%), the Czech Republic (94–100%) and Slovakia (89–99%). The estimates are unstable in the remaining three countries. In Germany, the percentile is estimated to be 94–100% in 2012, and based on data from 2017, the validity was not confirmed in any section. In Hungary, the estimates are 94–100% for 2012 vs. 65–88% for 2017, and in Poland, the figures are 91–100% vs. 80–100%.

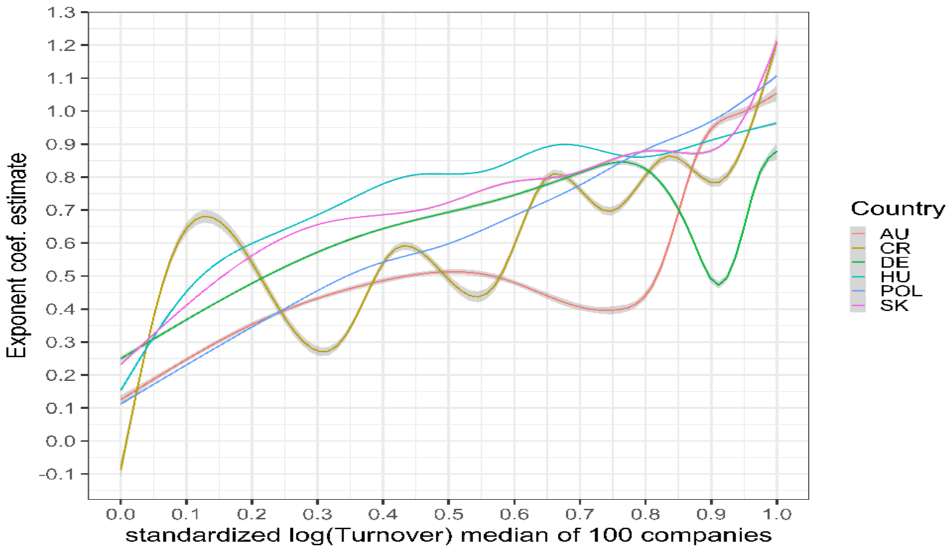
Table 2 | Power law exponent (coefficient) estimate from 0.9 to 1.1 (around 1 = Zipf's law)

Country	Year	Sample size	Percentile from	Percentile to	Turnover from (CZK, in millions)	Turnover to (CZK, in millions)
All sectors and company sizes						
Austria	2012	4,511 (100%)	83.9%	100.0%	2,873	10,540
	2017	7,541 (100%)	81.4%	100.0%	962	362,099
Czech Republic	2012	36,098 (100%)	94.1%	100.0%	568	1,252
	2017	33,004 (100%)	93.5%	99.5%	643	6,617
Germany	2012	22,175 (100%)	94.2%	100.0%	9,072	21,225
	2017	79,370 (100%)	N/A	N/A	N/A	N/A
Hungary	2012	33,878 (100%)	94.2%	100.0%	278	12,446
	2017	31,448 (100%)	64.8%	87.7%	56	184
Poland	2012	9,371 (100%)	91.4%	100.0%	784	12,263
	2017	31,098 (100%)	79.8%	100.0%	277	298,183
Slovakia	2012	16,179 (100%)	89.3%	99.0%	213	1576
	2017	17,387 (100%)	90.7%	96.3%	306	800

Source: authors

Figures 3 and 4 show the PLEE estimates for each country on a standardized (due to the comparison of differently strong economies) scale of turnover in 2017 and 2012. The figures show that the PLEE estimates for Germany and Hungary differ significantly compared to the three other Visegrad Group countries and Austria, where the estimates are comparable.

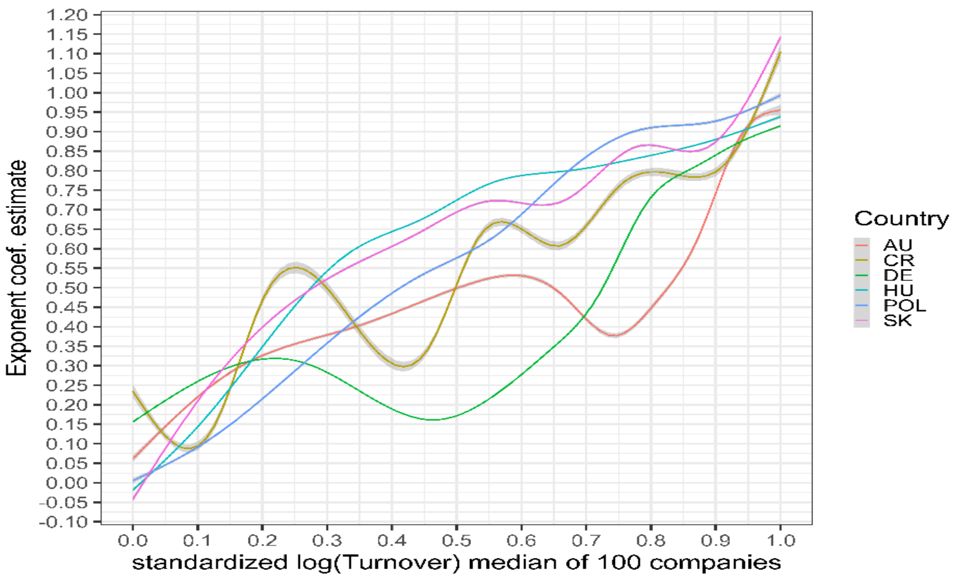
Figure 3 | PLEE estimate in 2017 for each country



Note: The standardized logarithm of turnover is on the x-axis and the PLEE estimate for the given distribution section (x) is on the y-axis.

Source: authors

Figure 4 | PLEE estimate in 2012 for each country



Note: The standardized logarithm of turnover is on the x-axis and the PLEE estimate for the given distribution section (x) is on the y-axis.

Source: authors

Table 3 presents the numbers of firms and summarizes the estimated sections in the manufacturing sector in each economy using the percentile interval and their sales, where Zipf's law applies (PLEE estimate in the interval 0.9–1.1).

In the subgroup of medium-sized companies from sector C, Zipf's law applies in the 40–85% percentile for Austria and Hungary, around 49–75% percentile for the Czech Republic, around 42–70% for Poland, and 49–80% for Slovakia. The percentiles for these countries were similar in 2012 and 2017. In Germany, the estimates are fundamentally different (no intersection) in 2012 (48–64%) and in 2017 (72–80%).

Table 3 | Power law exponent (coefficient) estimate from 0.9 to 1.1 (around 1 = Zipf's law)

Country	Year	Sample size	Percentile from	Percentile to	Turnover from (CZK, in millions)	Turnover to (CZK, in millions)
Sector C – MANUFACTURING firms with 50-249 employees						
Austria	2012	192 (4%)	40.0%	80.4%	910	1,111
	2017	389 (5%)	44.6%	80.4%	708	1,767
Czech Republic	2012	2,163 (6%)	49.3%	75.3%	173	297
	2017	2,100 (6%)	47.6%	68.9%	188	317
Germany	2012	1,358 (6%)	48.1%	63.7%	883	1,057
	2017	2,486 (3%)	72.0%	80.0%	696	952
Hungary	2012	1,266 (4%)	42.4%	85.4%	119	327
	2017	1,477 (5%)	51.6%	75.9%	156	318
Poland	2012	1,156 (12%)	41.9%	69.7%	160	229
	2017	3,072 (10%)	45.4%	63.6%	172	262
Slovakia	2012	787 (5%)	57.6%	80.2%	236	306
	2017	876 (5%)	49.3%	73.2%	205	383

Source: authors

Table 4 presents the numbers of firms and summarizes the estimated sections in the wholesale and retail sectors in each economy using the percentile interval and their turnover, where Zipf's law applies (PLEE estimate in the interval 0.9–1.1).

In the subgroup of small firms from sector G, the estimates are stable (comparable between 2012 and 2017) for the Czech Republic (around 57–72%) and Hungary (49–79%), comparable for Poland (47–76%) and Slovakia (48–85%). Based on our data, the estimate appears to be unstable or similar for Austria (30–80% in 2012 vs. 47–100% in 2017) and Germany (95–100% vs. 47–64%).

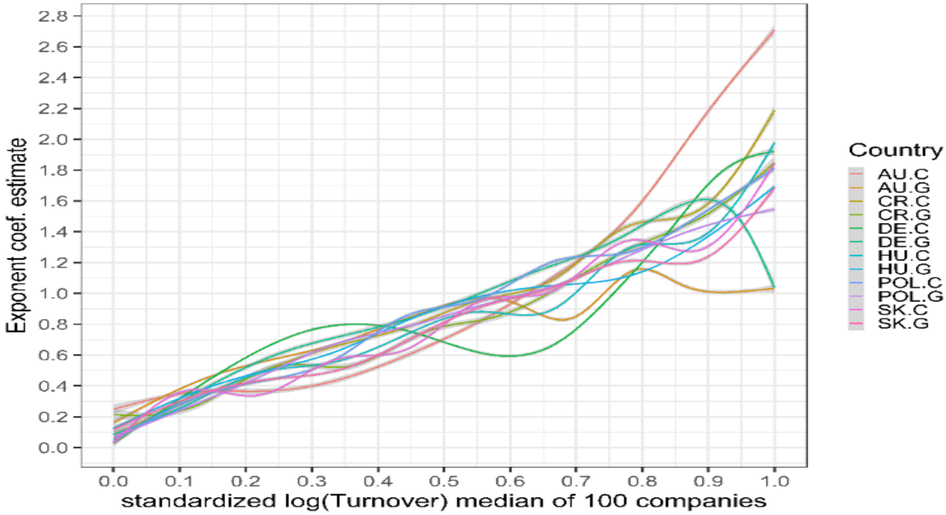
Table 4 | Power law exponent (coefficient) estimate from 0.9 to 1.1 (around 1 = Zipf's law)

Country	Year	Sample size	Percentile from	Percentile to	Turnover from (CZK, in millions)	Turnover to (CZK, in millions)
Sector G – WHOLESALE AND RETAIL TRADE; REPAIR AND MAINTENANCE OF MOTOR VEHICLES SECTION firms with 10-49 employees.						
Austria	2012	492 (11%)	30.4%	80.4%	101	231
	2017	957 (13%)	47.3%	100.0%	124	22,164
Czech Republic	2012	2,947 (8%)	56.5%	71.4%	94	129
	2017	2,869 (9%)	58.3%	71.7%	118	176
Germany	2012	1,698 (8%)	94.8%	100.0%	2,024	3,020
	2017	10,039 (13%)	46.5%	63.5%	79	118
Hungary	2012	3,620 (11%)	54.9%	74.9%	51	84
	2017	4,744 (15%)	48.9%	78.6%	53	124
Poland	2012	1,402 (15%)	46.6%	76.0%	96	158
	2017	5,143 (17%)	51.5%	71.5%	92	156
Slovakia	2012	1,350 (8%)	48.4%	85.0%	78	175
	2017	1,887 (11%)	50.4%	73.6%	76	143

Source: authors

Figures 5 and 6 show the PLEE estimates in 2017 and 2012 for individual subgroups and countries. The figures indicate that the estimates are similar up to about 70% of the percentile, and then differ significantly between sectors and countries.

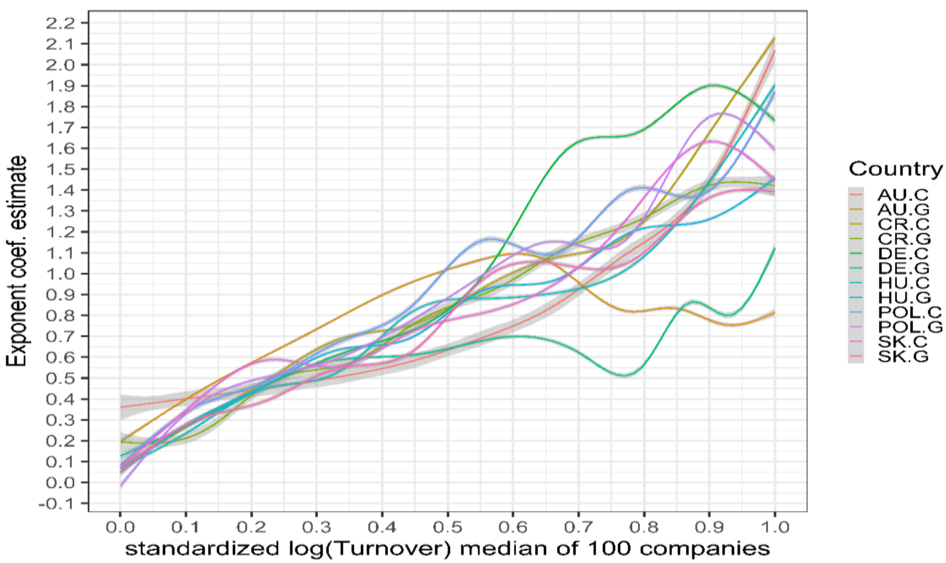
Figure 5 | PLEE estimate in 2017 for each subgroup and country



Note: The standardized logarithm of turnover is on the x-axis and the PLEE estimate for the given distribution section (x) is on the y-axis.

Source: authors

Figure 6 | PLEE estimate in 2012 for each subgroup and country



Note: The standardized logarithm of turnover is on the x-axis and the PLEE estimate for the given distribution section (x) is on the y-axis.

Source: authors

5 Discussion

Based on our data, we found that Zipf's law does not apply to the overall distribution of sales of all firms in an economy (or in an economic subgroup), but only in the sections in the right part of the distribution (companies with above-average sales revenues within the whole economy), which is in line with many previous studies, which have focused only on the largest firms in the economy and estimated the power law coefficient only for companies located in the upper tail of the distribution (for example, Zhang et al., 2009; Gao et al., 2015; Jenkins, 2017; Blanchet et al., 2017; Maia et al., 2020).

In this empirical study, subgroups (medium-sized companies from sector C and small companies from sector G) include a greater percentage (interval) of firms where Zipf's law applies. Zipf's law applies to different parts of the distribution (percentiles of sales) in individual countries. The sensitivity analysis showed that estimates in 2012 are comparable for most countries (in whole economies and selected subgroups) with the (primarily monitored) year 2017. The Czech Republic, Poland and Hungary show similar results for individual sectors.

The results of the analysis naturally depend on the quality of the data used. Franchina and Sergiani (2019) classified Bureau van Dijk's data as class A based on the performed analysis. These data are reliable for analysis and forecasts, which have been performed (Franchina & Sergiani, 2019). Financial data from the selected database are used widely for studying European companies (Liu, 2020) because it is the largest database of this type. This fact is confirmed by Bajgar (2020) as well. Moreover, he adds that data representativeness is increased significantly by excluding extremely small firms. This was performed in this analysis as well. There is a high probability that other commercial datasets would be worse than Bureau van Dijk's Amadeus database (Bajgar, 2020).

In order to reduce the selection bias, we chose, in addition to the PLEE estimates for the whole economy, two subgroups of firms and also performed a sensitivity analysis for the year 2012. The estimates of the percentile of firms in the whole economy, where Zipf's law applies to three out of the six countries, are similar in 2012 and 2017. Specifically, they are Austria (approximately 81–100%), the Czech Republic (94–100%) and Slovakia (89–99%).

As another limitation of the present research and its results, it is necessary to mention the fact that the firm size distribution changes according to the phase of the business cycle in which it is studied. The reaction of firms to the economic cycle is influenced by their size and also by their age. The distribution of companies is more uniform during recessions. Small firms are more sensitive to changes in the business cycle than large ones (Fort et al., 2013). Therefore, the distribution of firms changes according to the phases of the business cycle. The increasing value of the power law exponent indicates the general growth of economic activity, which leads to an increase in the concentration of firms and, conversely, in times of recession, economic activity decreases, which generally leads to a reduction in the concentration of firms, and thus the value of the estimated exponent should be lower.

Conclusion

This article contributes to the understanding of the distribution of firms in economies or in their industries. Knowledge of the proportion of representation of large and small firms is important for estimating the macroeconomic impacts of the change, if any, in the distribution of firms. The more market power is concentrated in a small fraction of firms, the better the chances that even minor changes in the distribution of firms can have relatively large macroeconomic impacts, in terms of either employment or the volume of production offered. The idea that in the distribution of firms according to the power law, shocks (changes) at the firm level can cause changes at the aggregate level has been supported, for example, by Gabaix (2011) or Acemoglu et al. (2010).

Firm size distribution can be obtained from the firm growth dynamics. It is considered the result of the stochastic growth models coming from Gibrat's law (Gibrat, 1931). Even though Gibrat originally explained the lognormal distribution of the firm size by the fact that firm growth is a stochastic process that is independent of their size, a number of researchers later demonstrated that the power law can be a consequence of stochastic firm growth as well (Crosato & Ganugi, 2007; Saichev et al., 2008). This attribute turns out to be very important in the research into the factors affecting aggregate economic outcomes. This has been proven by the relatively new study of Carvalho and Grassi (2019) as well. It followed abovementioned Gabaix's (2011) work. Carvalho and Grassi showed that if the firm size distribution is in compliance with the power law, then the big firms' dynamics are able to control the economic cycle. In other words, idiosyncratic shocks at the firm level cause a significant part of the aggregate volatility (Carvalho and Grassi, 2019). There is another consequence of firm size power law distribution – the impact on the concentration of wealth and income, which can affect the structure of aggregate demand and the growth rate of aggregate productivity (Stachurski, 2019).

The power law basically indicates a higher concentration of firms in a given economy or economic sector compared to another distribution. This means that there are larger companies in a given economy (or sector), and they are larger compared to a different distribution (for example, lognormal distribution). A higher estimated value of the power law exponent over time in a given economy (or sector) indicates that the number of large firms is increasing in proportion to the small ones. The recommendation resulting from our study is that Zipf's law does not apply to the entire distribution of firms according to the size of sales, and we recommend verifying its validity for each economy or, as the case may be, subgroup separately. Due to the above considerations about the quality of data in the Amadeus database, it would be appropriate to validate our results and conclusions on data from other sources. Another benefit of the article is that we described the size distribution of firms in countries where we could not find similar analyses, except one source, namely Di Giovanni et al. (2011).

At the same time, we were able to use our solution procedure to identify the range of values that behave according to the power law within each examined country, which is often mentioned in the literature as one of the problems accompanying such analyses. In 2012, the validity of the power law can be observed in all the countries studied in approximately the last decile of the values (with a slight deviation of the validity interval for firms in Austria, where the law also applies to parts of the ninth decile of the values); therefore, no significant differences can be observed between the Visegrad Four countries and Germany and Austria, which were also examined. We found basically similar results from data which are five years older in the distribution of Czech, Slovak and Austrian firms, but in the case of the two other Visegrad Four countries, the section on the power law validity was extended over time;

specifically for Poland, the beginning of the power law validity interval was moved to the end of the eighth decile and for Hungary to the seventh decile, with the end in the ninth decile of the values. The validity of the power law was not observed at all in the German economy in 2017. Here, one can also see a topic for further research, where one of the ways to explain this fact is to analyse the development and changes, if any, in the growth dynamics of German firms, which increasingly drove the growth of the local economy until the period in question. Of course, these effects have an impact on the distribution of firms in an economy that does not copy the power law if small firms grow faster than large ones, as has also been demonstrated in this text. If we look at the validity of the power law within the sectors, we have found that it is possible to observe that this law applies to longer sections rather than entire economies. Within individual sectors, the Visegrad Four countries show more similar results than the other two examined countries outside the Visegrad Group.

Acknowledgement

Funding: There was no special funding towards this study.

Conflict of interest: The authors hereby declare that this article was not submitted nor published elsewhere. Authors do not have any conflict of interest.

References

Acemoglu, D., & Cao, D. (2015). Innovation by entrants and incumbents. *Journal of Economic Theory*, 157, 255–294. https://doi.org/10.1016/j.jet.2015.01.001.

Acemoglu, D., Ozdaglar, A., Tahbaz-Salehi, A. (2010). Cascades in Networks and Aggregate Volatility. NBER Working Paper No. 16516.National Bureau of Economic Research.

Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes. *Science*, 293, 1818–1820. https://doi.org/10.1126/science.1062081.

Bajgar, M., Berlingieri, G., Calligaris, S., Criscuolo, Ch., Timmis, J. (2020). Coverage and representativeness of Orbis data. OECD Science. Technology and Industry Working Papers No. 2020/06, OECD Publishing, Paris.

Balthrop, A. T. (2016). Power laws in oil and natural gas production. *Empirical Economics*, 51, 1521–1539. https://doi.org/10.1007/s00181-015-1054-4.

Balthrop, A. T. (2021). Gibrat's law in the trucking industry. *Empirical Economics*, 61, 339–354. https://doi.org/10.1007/s00181-020-01851-1.

Bisnode Czech Republic (2019). http://binsode.cz. Date of access: 4 July 2019.

Blanchet, T., Fournier, J., Piketty, T. (2017). Generalized Pareto Curves: Theory and Applications. WID.world Working Paper.

Carvalho, V. M., & Grassi, B. (2019). Large firm dynamics and the business cycle. *American Economic Review*, 109, 1375–1425. https://doi.org/10.1257/aer.20151317.

Crosato, L., & Ganugi, P. (2007). Statistical regularity of firm size distribution: the Pareto IV and truncated Yule for Italian SCI manufacturing. *Statistical Methods and Applications*, 16(1), 85–115. https://doi.org/10.1007/s10260-006-0023-7.

Da Silva, S., Matsushita, R., Giglio, R., & Massena, G. (2018). Granularity of the top 1,000 Brazilian companies. *Physica A*, 512, 68–73. https://doi.org/10.1016/j.physa.2018.08.027.

Del Rosal, I. (2016). Power laws in EU country exports. *Empirica*, 45(2), 311–337. https://doi.org/10.1007/s10663-016-9362-2.

Di Giovanni, J., & Levchenko, A. A. (2013). Firm entry, trade, and welfare in Zipf's world. *Journal of International Economics*, 89, 283–296. https://doi.org/10.1016/j.jinteco.2012.08.002.

Di Giovanni, J., Levchenko, A. A., & Ranciére, R. (2011). Power laws in firm size and openness to trade: Measurement and implications. *Journal of International Economics*, 85, 42–52. https://doi.org/10.1016/j.jinteco.2011.05.003.

Duparcq, P., & Konings, J. (2016). *An Enterprise Map of Kazakhstan: A Guide for Industrial Policy*. Retrieved July 15, 2020, from http://shss.nu.edu.kz/shss/academics/departments/economics.

Eurostat. (2023). *GDP and main components (output, expenditure and income)*. Retrieved June 7, 2023, from https://ec.europa.eu/eurostat/databrowser/view/namq_10_gdp/default/table?lang=en.

Fort, T. C., Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2013). How firms respond to business cycles: The role of firm age and firm size. *IMF Economic Review*, 61(3), 520–559. https://doi.org/10.1057/imfer.2013.15.

Fujiwara, Y., Di Guilmi, C., Aoyama, H., Gallegati, M., & Souma, W. (2004). Do Pareto-Zipf and Gibrat laws hold true? An analysis with European Firms. *Physica A*, 335, 197–216. https://doi.org/10.1016/j.physa.2003.12.015.

Franchina, L., & Sergiani, F. (2019). High-quality dataset for machine learning in the business intelligence domain. In Y. Bi, R. Bhatia, & S. Kapoor (Eds.), Intelligent Systems and Applications: Proceedings of the 2019 Intelligent Systems Conference (IntelliSys) (pp 391–401). Springer International Publishing.

Gabaix, X. (2011). The Granular Origins of Aggregate Fluctuations. *Econometrica*, 79(3), 733–772. https://doi.org/10.3982/ECTA8769.

Gabaix, X., & Landier, A. (2008). Why has CEO pay increased so much? *Quarterly Journal of Economics*, 123(1), 49–100. https://doi.org/10.1162/qjec.2008.123.1.49.

Gaffeo, E., Gallegati, M., & Palestrini, A. (2003). On the size distribution of firms: additional evidence from the G7 countries. *Physica A*, 324, 117–123. https://doi.org/10.1016/S0378-4371(02)01890-3.

Gao, B., Chan, W. K. V., & Li, H. (2015). On the increasing inequality in size distribution of China's listed companies. *China Economic Review*, 36, 25–41. https://doi.org/10.1016/j.chieco.2015.08.005.

Gibrat, R. (1931). *Les Inégalités Économiques*. Paris: Librairie du Recueil Sirey.

Hernández-Pérez, R., Angulo-Brown, F., Tun, D. (2006). Company size distribution for developing countries. *Physica A*, 359, 607–618. https://doi.org/10.1016/j.physa.2005.04.027.

Jenkins, S. P. (2017). Pareto Models, Top Incomes and Recent Trends in UK Income Inequality. *Economica*, 84, 261–289. https://doi.org/10.1111/ecca.12217.

Kang, S. H., Jiang, Z., Cheong, Ch., & Yoon, S. M. (2011). Changes of firm size distribution: The case of Korea. *Physica A*, 390(2), 319–327. https://doi.org/10.1016/j.physa.2010.10.007.

Liu, G. (2020). Data quality problems troubling business and financial researchers: A literature review and synthetic analysis. *Journal of Business & Finance Librarianship*, 25, 315–371. http://dx.doi.org/10.1080/08963568.2020.1847555.

Luttmer, E. G. J. (2007). Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics*, 122, 1103–1144. https://doi.org/10.1162/qjec.122.3.1103.

Luttmer, E. G. J. (2011). On the mechanics of firm growth. *Review of Economic Studies*, 78(3), 1042–1068. https://doi.org/10.1093/restud/rdq028.

Maia, A., Matsushita, R., & Da Silva, S. (2020). Earnings distributions of scalable vs. non-scalable occupations. *Physica A*, 560, https://doi.org/10.1016/j.physa.2020.125192.

Mantegna, R. N., & Stanley, H. E. (2000). *An introduction to econophysics: Correlations and complexity in finance*. New York: Cambridge.

Matayev, CH. (2018). Power-Law Distributions in Economics: Empirical Evidence for an Emerging Economy, Kazakhstan. [Master thesis, Nazarbayev University], https://nur.nu.edu.kz/bitstream/handle/123456789/3166/13.04.18.pdf?sequence=1

Mura, L., Zsigmond, T., Kovács, A., & Baloghová, É. (2020). Unemployment and GDP relationship analysis in the Visegrad Four countries. *On-line Journal Modelling the New Europe*, 34, 118–134. https://doi.org/10.24193/OJMNE.2020.34.06.

Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46, 323–351. https://doi.org/10.1080/00107510500052444.

Okuyama, K., Takayasu, M., & Takayasu, H. (1999). Zipf's law in income distribution of firms. *Physica A*, 269, 125–131. https://doi.org/10.1016/S0378-4371(99)00086-2.

Pascoal, R., Augusto, M., & Monteiro, A. M. (2016). Size distribution of Portuguese firms between 2006 and 2012. *Physica A*, 458, 342–355. https://doi.org/10.1016/j.physa.2016.04.010.

Podobnik, B., Horvatic, D., Petersen, A. M., Urosevic, B., & Stanley, H. E. (2010). Bankruptcy risk model and empirical tests. *Proceedings of the National Academy of Sciences*, 107(43), 18325–18330. https://doi.org/10.1073/pnas.1011942107.

Provenzano, D. (2014). Power laws and the market structure of tourism industry. *Empirical Economics*, 47, 1055–1066. https://doi.org/10.1007/s00181-013-0769-3.

R. C. Team. (2018). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [software]. https://www.R-project.org/.

Rosen, S. (1981). The economics of superstars. *America Economic Review*, 71, 845–858.

Rossi-Hansberg, E., & Wright, M. L. J. (2007). Establishment Size Dynamics in the Aggregate Economy. *American Economic Review*, 97(5), 1639–1666. https://doi.org/10.1257/aer.97.5.1639.

Saichev, A., Malevergne, Y., & Sornette, D. (2008). Theory of Zipf’s Law and of General Power Law Distributions with Gibrat’s law of Proportional Growth. arXiv preprint arXiv: 0808.1828, arXiv.org. https://doi.org/10.48550/arXiv.0808.1828.

Segarra, A., & Teruel, M. (2012). An appraisal of firm size distribution: Does sample size matter? *Journal of Economic Behavior & Organization*, 82(1), 314–328. https://doi.org/10.1016/j.jebo.2012.02.012.

Simon, H. A., & Bonini, C. P. (1958). The size distribution of business firms. *American Economic Review*, 48(4), 607–617.

Stachurski, J. (2019). Power Laws without Gibrat’s Law. arXiv preprint arXiv:1910.14023, aXiv.org. https://doi.org/10.48550/arXiv.1910.14023.

Stanley, M. H. R., Buldyrev, S. V., Havlin, S., Mantegna, R. N., Salinger, M. A., & Stanley, H. E. (1995). Zipf plots and the size distribution of firms. *Economics Letters*, 49(4), 453–457. https://doi.org/10.1016/0165-1765(95)00696-D.

Wagner, J., & Weche, J. P. (2020). On the granularity of the German economy – first evidence from the top 100 companies panel database. *Applied Economics Letters*, 27(21), 1768–1771. https://doi.org/10.1080/13504851.2020.1722790.

Zhang, J., Chen, Q., & Wang, Y. (2009). Zipf distribution in top Chinese firms and an economic explanation. *Physica A*, 388(10), 2020–2024. https://doi.org/10.1016/j.physa.2009.01.027.

The research article has been peer-reviewed. | Received: 20 April 2023**; Revised:** 11 June 2023**; Accepted:** 5 July 2023**; Available online:** 4 September 2023**; Published in the regular issue:** 18December 2023.

Annex

Table 1 | Literature overview – power law in firm size distribution

Article	Journal	Power law exponent (point estimate)	Country	Sample size	Size measure	Key findings
Okuyama et al. (1999)	Physica A	α ≅ 1	Japan	85,375	Income before taxes	At the sectoral level only in Japan α ∈ [0.7, 1.2]
		α ≅ 1	Italy	Over 11,000	Income before taxes	
		α ≅ 1.4	USA	Over 10,000	Income before taxes	
Axtell (2001)	Science	α = 1.059	USA	5,541,918	Employees	In 1992 α = 0.994
		α = 0.994			Sales	No analysis with sales in 1992
Gaffeo et al. (2003)	Physica A	α = 0.96	G7	22,064	Sales	Zipf's law not rejected only with sales as size variable
		α = 1.16			Capital	
		α = 1.14			Debt	
Fujiwara et al. (2003)	Physica A	α = 0.886	France	8,313	Assets	Validity of Gibrat's law also confirmed
		α = 0.896	France	15,776	Sales	
		α = 0.995	UK	15,055	Employees	
Hernández-Pérez et al. (2006)	Physica A	α = 0.576	Argentina	501	Sales	Firm size distribution corresponds to power
		α = 0.687	Bolivia	100		

Volume 12 | Issue 5 | 2023
https://doi.org/10.18267/j.cebr.345

		α = 0.696	Brazil	150		law in larger firms in the sample
		α = 0.445	Chile	100		
		α = 0.819	Colombia	300		
		α = 1.432	Ecuador	51		
		α = 0.810	Mexico	500		
		α = 0.975	Mozambique	100		
		α = 0.615	Peru	1,963		
		α = 0.425	Venezuela	219		
Gabaix and Landier (2008)	Quarterly Journal of Economics	Varies by method. Average over time: α = 1.095/ 0.869	USA	Top 500 firms	Market value	Specific value provided only for 2004 according to Hill method: α = 1.01
Zhang et al. (2009)	Physica A	In 2007: α = 0.960 In 2006: α = 0.937 In 2005: α = 0.971 In 2004: α = 1.010 In 2003: α = 1.006 In 2002: α = 1.013	China	Top 500 firms	Sales	Very good data fit to the law subject to verification
Podobnik et al. (2010)	Proceedings of the National Academy of Sciences	α = 1.11	USA	2,737	Assets (before bankruptcy)	When using book and market values, the hypothesis of the
		α = 1.44		462	Assets (in bankruptcy)	

		α = 1.1		2,545 firms in sub-sample NASDAQ	Market capitalization	validity of Zipf's law is not refuted
		α = 1.02			Equity	
Di Giovanni et al. (2011)	Journal of International Economics	Varies by method: α = 1.017/1.019/0.825	France	157,084	Sales	Non-exporting firms have a larger exponent than the one found in the whole dataset and vice versa
		Varies by method: α = 1.078/1.093/0.79		152,429	Employees	
Kang et al. (2011)	Physica A	In 2007: α = 1.268 In 1997: α = 1.259 In 1987: α = 1.101	Korea	703	Sales	Before the crisis, the degree of inequality in the size of Korean companies increased, and it decreased after the crisis
		In 2007: α = 1.266 In 1997: α = 1.316 In 1987: α = 1.153		704	Assets	
		In 2007: α = 0.985 In 1997: α = 1.035 In 1987: α = 0.934		703	Employees	
		In 2007: α = 1.221 In 1997: α = 1.212 In 1987: α = 1.131		704	Capital	
Segarra and Teruel (2012)	Journal of Economic Behavior and Organization	In 2006: α = 0.970 In 2001: α = 0.945	Spain	61,455	Employees	As the size of the dataset decreases, the power law exponent value increases
		In 2006: α = 0.675 In 2001: α = 0.661		61,322	Sales	
di Giovanni and Levchenko (2013)	Journal of International Economics	Varies by method and year: α ∈ [0.69, 1.18]	44 countries	Minimum 1,000 firms	Sales	

Gao et al. (2015)	China Economic Review	Varies by year: α ∈ [0.757, 1.499]	China	Varies by year: From 1,140 to 2,136	Assets	The inequality in the distribution of Chinese tradable firms is increasing over time
		Varies by year: α ∈ [0.934, 1.640]			Equity	
Pascoal et al. (2016)	Physica A	Varies by year: α ∈ [1.425, 1.489]	Portugal	8,480	Assets	To approximate company size, it is more appropriate to use the amount of sales rather than assets
		Varies by year: α ∈ [1.695, 1.821]			Sales	
Matayev (2019)	Master thesis	α = 2.36	Kazakhstan	Approximately 67,000	Employees	No evidence of existence of power law!

Source: authors