

Article

Robust Characterization of Multidimensional Scaling Relations between Size Measures for Business Firms

Yuh Kobayashi ¹ , Hideki Takayasu ^{2,3} , Shlomo Havlin ^{2,4} and Misako Takayasu ^{1,2,*} 

¹ Department of Mathematical and Computing Science, School of Computing, Tokyo Institute of Technology, Yokohama 226-8502, Japan; kobayashi.y.bz@m.titech.ac.jp

² Institute of Innovative Research, Tokyo Institute of Technology, Yokohama 226-8502, Japan; takayasu@csl.sony.co.jp (H.T.); havlin@ophir.ph.biu.ac.il (S.H.)

³ Sony Computer Science Laboratories, Tokyo 141-0022, Japan

⁴ Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel

* Correspondence: takayasu.m.aa@m.titech.ac.jp

Abstract: Although the sizes of business firms have been a subject of intensive research, the definition of a “size” of a firm remains unclear. In this study, we empirically characterize in detail the scaling relations between size measures of business firms, analyzing them based on allometric scaling. Using a large dataset of Japanese firms that tracked approximately one million firms annually for two decades (1994–2015), we examined up to the trivariate relations between corporate size measures: annual sales, capital stock, total assets, and numbers of employees and trading partners. The data were examined using a multivariate generalization of a previously proposed method for analyzing bivariate scalings. We found that relations between measures other than the capital stock are marked by allometric scaling relations. Power-law exponents for scalings and distributions of multiple firm size measures were mostly robust throughout the years but had fluctuations that appeared to correlate with national economic conditions. We established theoretical relations between the exponents. We expect these results to allow direct estimation of the effects of using alternative size measures of business firms in regression analyses, to facilitate the modeling of firms, and to enhance the current theoretical understanding of complex systems.

Keywords: business firms; companies; econophysics; multivariate scaling; nontrivial power-law exponents; size measures



Citation: Kobayashi, Y.; Takayasu, H.; Havlin, S.; Takayasu, M. Robust Characterization of Multidimensional Scaling Relations between Size Measures for Business Firms. *Entropy* **2021**, *23*, 168. <https://doi.org/10.3390/e23020168>

Academic Editor: Pentti Nieminen

Received: 6 January 2021

Accepted: 25 January 2021

Published: 29 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

An index of the size of a business firm has numerous implications, ranging from its method of corporate finance [1,2] and the quality of its CEO [3] to employee job satisfaction [4] and gender gaps in wages among employees [5]. The exact determinants of the size of a specific firm or a group of firms in an industry have been sought both empirically [6,7] and theoretically [8,9]. Researchers have also studied regularities in the distribution of firm size empirically [10–15], mainly motivated by the theory that the evolution of firm size can be modelled as a stochastic process [10,16–18].

However, the question of what constitutes the “size” of a firm has rarely been addressed, and it is possible that statistical results regarding firm size depend on the measures used. Indeed, many of the effects previously found in empirical studies of corporate finance were not robust when different size measures were used as a control variable [19]. Researchers, who questioned the often implicit assumption that size measurements of firms are adequate indicators of their actual size [19–21], evidently agree that the relations of various size measures commonly used have not been well-explored in the literature. Moreover, some [20,21] propose the notion of multidimensional size for business firms, using various size measures to indicate manifold aspects. Although it might clarify the conceptual status of firm size, differences and relations between firm size measures should be explored empirically for the notion of multidimensional size to be fruitful.

Regarding the relations between corporate size measures, power-law scaling relations have been reported empirically [12,19,22–24]. More interestingly, some size measures have been reported to demonstrate allometric scaling [23], which means that the distribution of relative deviations from the average relation between two size measures was invariant regardless of firm size. In other words, when the logarithms of a size measure were regressed against the logarithms of another measure, the residuals were found to be surprisingly homoscedastic and independent of the “explanatory” variable; note that the use of formulae for regression analysis is solely intended for the characterization of multivariate joint distributions and does not imply causal relations between the size measures. This suggests that we may be able to employ the residual as a variable indicating the adjusted “proportion” between size measures of a firm, allowing us to isolate the effects of using alternative size measures. Nevertheless, the size measures considered in previous studies are not intended to represent all possibilities. It is unknown as to what extent allometric scaling accurately describes empirical relations regarding other size measures. Theories regarding relations between multiple allometric scalings have been lacking as well. Therefore, empirical characterization and theoretical understanding of various size measures are still far from complete.

Analysis of the empirical relations between size measures of firms also has merit for advancing complex systems science, and a few studies [23,24] have indeed been pursued from such a perspective. Power laws of the form $x \propto y^\gamma$ are a typical functional form of scaling found in general complex systems, such as animal bodies [25–28], ecological communities [29,30], and cities [31,32]. For instance, the value of exponent γ in the scaling of the metabolic rate (x) on the body mass (y) is close to $3/4$ in mammals [25,27,28] and is theoretically related to the minimization of the energy consumption in blood pumping [33]. Theoretical considerations have successfully been used to predict numerous other scaling exponents in the natural systems of animal bodies. Similarly, the unique value of $\gamma = 2$ for humans (compared with $\gamma = 3$ for other animals) in the scaling of the body mass (x) against the body length (y) was theoretically accounted for by human bipedalism [34]. In this manner, studies on scaling relations provide a basis for a deeper understanding of a system’s governing principles.

The present study is focused on multivariate relations among five quantities of firms that have often been used as size measures in research and practice to characterize scaling relations exhibited by firms across scales. We based our study on exhaustive annual data for more than 2 million unique Japanese firms over a 22-year period. After examining the properties of individual size measures, we empirically verified statistical models of up to three-dimensional allometric scaling for four of the five size measures, while establishing theoretical relations between multiple scalings. Building upon these results, we further developed a method for estimating scaling exponent γ as well as the range where the data fit well to the scaling relations, thereby obtaining insights into the temporal robustness of our results.

The remainder of this paper is organized as follows. First, we describe our selection of firm size measures and the employed data in Section 2. After the dataset and statistical properties of each size measure are characterized in Section 3.1, we inspect the bivariate scalings between the size measures and specify the pairs of variables that are in an allometric scaling in Section 3.2. We then address the problems of seeming inconsistency between the power-law exponents for the tails of marginal and conditional distributions of size measures with regard to bivariate scalings (Section 3.3). Next, we examine a trivariate scaling relation in Section 3.4. Trivariate allometric scalings were found for some combinations of size measures and shown to be related to the transitivity of allometric scalings both theoretically and empirically. This theoretical development allows robust estimation of bivariate and trivariate scaling exponents (Section 3.5). We found that the number of employees has a higher relevance to sales than does the number of trading partners, and there was a mild variation in the trivariate scaling exponents in the period 1994–2015, which appeared to be correlated to the gross domestic product (GDP) variation (Section 3.6).

Finally, we discuss our results and conclude the paper in Section 4. Our rationales for using the standard regression analysis are given in Appendix A, and problems related to the theoretical relations between several power-law exponents that appear in scalings, conditional distributions, and marginal distributions are discussed in detail in Appendix B.

2. Materials and Methods

2.1. Selection of Firm Size Measures

Despite the recurrent claim that the concept of firm size has not been sufficiently clarified [19–21], the variety of size measures employed in studies has remained mostly unchanged during the past half century. For example, in an economic study in 1975, the authors compared different firm size measures using the data of sales, total or net assets, employment, invested capital, and market value [22]. Almost contemporarily, a review of organizational studies in 1976 listed the number of employees as well as the number of clients, sales, and net assets as commonly used size measures and proposed physical capacities, such as the square footage (area) available for an organization's activities, as a possible measure [20]. Recent research on corporate finance mentions total assets, sales, market value, and the number of employees [19]. Meanwhile, the econophysics school of research has used sales, number of employees, total assets, net sales, net income, and number of trading partners (i.e., the reported number of firms that a firm directly trades with) [12,23,35]. Finally, as small and medium enterprises are legally defined according to their capital stock in Japan [36], capital stock data are widely available for Japanese firms, which is the subject of the database used in our study.

We considered only the firm size measures that did not take negative values and were readily measured for almost all firms. The former criterion concerns the very notion of size in broader contexts. For instance, when the size of an animal body is considered, it is non-negative whether it is measured by weight, length, or how many calories per day it takes for the organism to live. We excluded net assets, net sales, and net income based on this criterion, as these measures can take negative values. Other implications of the concept of a size measure include positive correlations to other size measures—e.g., the weight of a human body is usually positively correlated to the height—but this was not incorporated into our criteria due to the empirical nature of the question. We set the latter criterion pragmatically, allowing us to preserve the comprehensiveness of our original data and to minimize the sampling bias. Market value was rejected with this criterion because data were unavailable for firms that were not publicly traded, which comprised the vast majority by number. Thus, the measures of corporate size that were referred to in the literature and that met our criteria were sales, total assets, invested capital, capital stock, and the numbers of employees and trading partners. We omitted invested capital from our analysis because of its partial conceptual overlap with capital stock, relative scarcity of the data, and its rare appearance in the literature as a size measure.

2.2. Data and Preprocessing

The source of the data used in this study was a collection of brief descriptions of Japanese business firms, including the financial status, type of business, and physical location, and this was provided and maintained by a major Japanese private credit reporter, Teikoku Databank, Ltd., Japan (hereinafter referred to as TDB). The data were mostly based on annual questionnaires presented to Japanese business firms.

The data employed in this study were primarily from the electronic version of the COSMOS 2 database, which is updated every January by TDB. This database describes over two million business firms in Japan and includes two sub-databases: (1) “summaries” of firm profiles, including the sales, capital stock, number of employees, and business type category (approximately one million per year, from 1980) and (2) lists of several million trading relationships between firms for each year since 1993. The profile of the trading relationships included the direction of money flows, i.e., which firm (buyer) paid which (supplier). Our study was focused on the data after 1993, in which the COSMOS 2 database

started to capture the trading relations between firms. We excluded the 1993 data to avoid boundary effects that may have been present at the beginning of the data collection. The data for sales, number of employees, and trading relations that we employed were last updated in January 2017, whereas the data for the capital stock were compiled in January 2020. This confined our analysis to the period ending in 2015 because the 2016 sales were not available for a substantial portion of the firms. Data for the total assets were from the COSMOS 1 database by TDB, in which financial statements of firms were included. This database was last updated in May 2020, and data since 2000 were available to us. Each firm has a unique anonymous ID that allows the firm to be identified in different databases.

To ensure time coherence between different types of data and to make the included firms sufficiently homogeneous for the purpose of our analyses, we performed several data-compilation steps, as follows. First, we excluded the firms that were categorized as governmental (e.g., local governments) or financial (e.g., banks and insurers). This is because the definition of sales for these “firms” differed significantly from that for construction, manufacturing, and wholesale firms, which constituted the majority of the firms in the data. In this step, we filtered out possible outliers in the database. This procedure was applied to both databases. Second, we did not use sales and capital stock data from financial statements published more than 8 years before data entry or without an adequate timestamp. Third, we set the sales to be unknown when the end of the fiscal year had changed because some such data were apparently not the annual sales (the value was sometimes considerably lower than those in previous or subsequent periods), whereas we accepted the data of capital stock and total assets in such cases if more than one instance of inconsistent data did not appear for the same year. Fourth, we determined the year to which a datum of sales, capital stock, or total assets was assigned according to the year in which the fiscal term ended. For example, when a firm had a fiscal year that started in April 2000 and ended in March 2001 (as for the majority of Japanese firms), the sales value of the fiscal term was considered to belong to the year 2001, regardless of whether the data appeared in the database in 2002 or later. In contrast, this principle was not applied to the employee or trading data; we always assigned them to the year prior to the data entry, because the COSMOS 2 database was updated in January of each year.

We directly used the raw data without normalization or adjustment for inflation. The number of trading partners of a firm was determined by counting the trading relations in which the firm participated throughout the year, regardless of whether the firm was a supplier or a buyer. This amounted to computing the sum of the in-degree and out-degree for each node in the directed network of trading.

The data for firm exits were additionally compiled to estimate the number of firm entries and exits. First, if any data of a firm existed in the COSMOS 2 database for a year, the firm was considered to exist in that year. Additionally, when the data of a firm were unavailable for up to two consecutive years, the firm was considered to exist in these years, assuming that its absence from the data was accidental. Before the compilation of the exit data, the first and fourth steps of data exclusion described earlier for the compilation of quantitative data were applied to ensure consistency between the datasets.

When multiple variables were used in the analyses, we excluded the data of firms that lacked any one of the variables. Results for the data amount are presented in Section 3.1 and tabulated in Table S1.

3. Results

3.1. Characterization of Databases and Size Measures

The amounts of data before and after the compilation are plotted in Figure 1 with respect to the years in the 1994–2015 period. The amount of data generally increased. The number of firms with complete data for multiple size measures in the COSMOS 2 database (i.e., number of trading partners k , number of employees ℓ , and annual sales s) was more than 0.8 million per year in 1999 and later, and the data for total assets a were more than 0.2 million per year in 2004 and later, as shown in Figure 1a, after the filtering

procedure described in Section 2.2. The number of firms considered to be existing based on the COSMOS 2 database was usually larger than the number of firms in the sub-database of sales and employees, as existence was assumed for some firms that only appeared in the trading data or those with temporarily missing data. Two sudden increases in the number of trading data (2007–2008 and 2010–2011) are observed in Figure 1b. The second jump was due to trading data involving financial or governmental organizations, as we observed no jump in the same period after the filtering. In contrast, the first jump was the consequence of revised data-collection methods that were specific to trading-relations data; trading relationships that had been mentioned only in on-demand reports started to be included in the database in 2008. Indeed, there was no comparable jump in the number of “summary” profile data. Additionally, the number of exiting or disappearing firms was stable compared with the number of entering firms, as shown in Figure 1c, implying that the former was less affected by the fluctuation of the data-collection efforts by TDB.

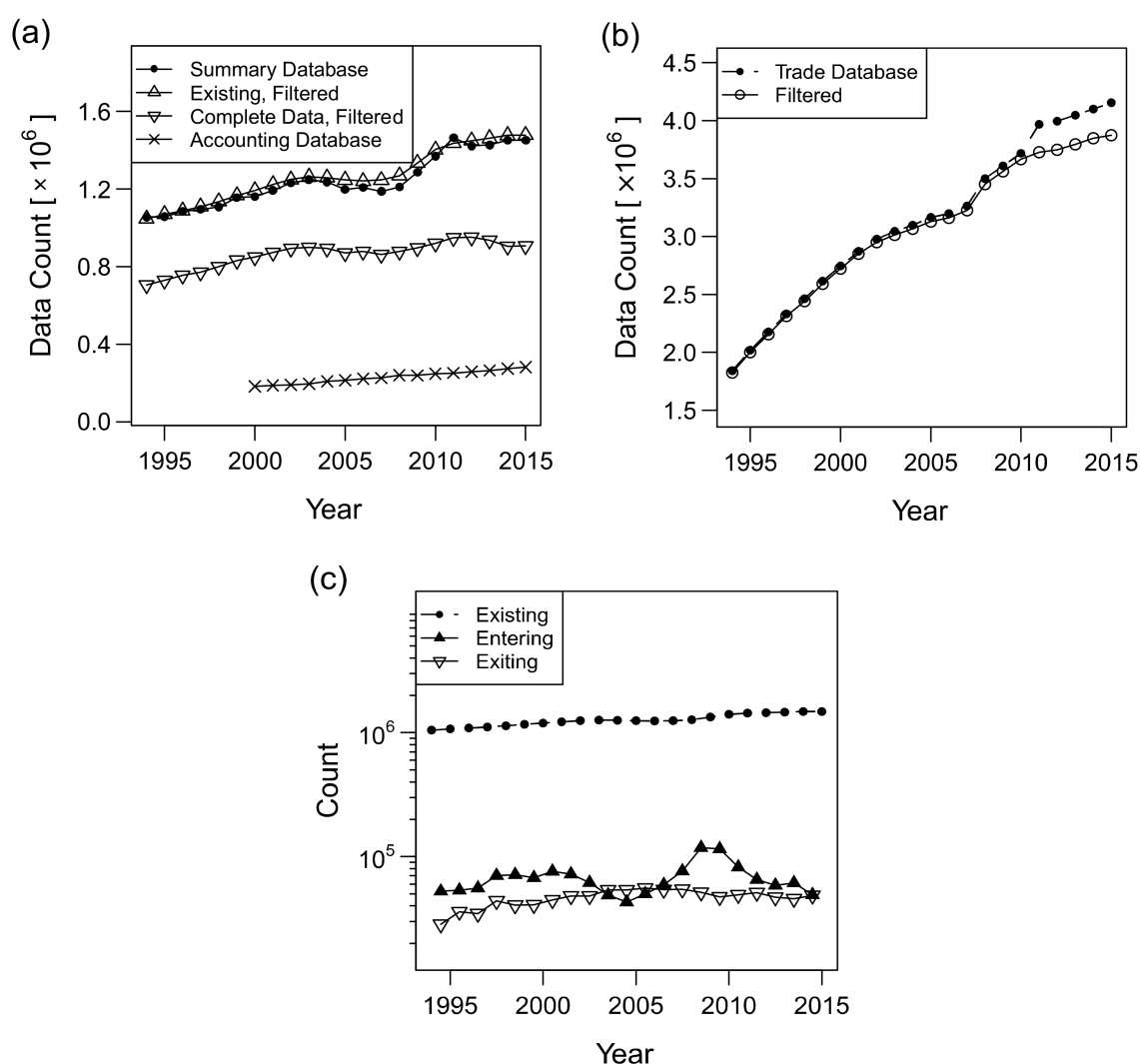


Figure 1. Changes in the data amount for the period 1994–2015. (a) Data amount for the original COSMOS 2 database (black dots), number of existing firms after the filtering and integration with network data (upward triangles), number of firms with data of all the three variables of sales, number of employees, and number of trading partners equal to or larger than one (downward triangles), and number of firms with total asset information in the COSMOS 1 database after the filtering (cross marks). (b) Data amount (number of links) for the original trade database (dots) and after the filtering (circles). (c) Numbers of entering (upward triangles) and exiting (downward triangles) firms. The number of existing firms (dots) is plotted for comparison. The vertical axis is in the logarithmic scale.

All the size variables were distributed in fat-tailed manners on the right side, as shown in Figure 2. Their distribution functions for every year are plotted in the figure, with the color gradient from blue through black to red, indicating the direction from older data to newer ones. The distributions were fairly stable despite the increase in data and, in particular, the exponents of the power-law tails were evidently invariant when the tails could be approximated by a power-law distribution. They were -2.2 for k (the number of trading partners), -2.2 for ℓ (employee number), -2.0 for s (annual sales in million yen), and -1.85 for a (total assets in thousand yen). The exponent of the sales distribution was consistent with that in previous studies [13–15]. The distribution of capital stock (hereafter denoted by p) was also fairly fat-tailed, although it did not seem to be characterized by a power law because the curve of the estimated density function was convex, as seen in Figure 2e. Consequently, the arithmetic mean of s and a did not indicate the representative value, and the standard deviation of all the variables was not a robust index of the width of the distribution (Table S1). We alternatively indicated the representative value and width of distribution by the median and the interquartile range, respectively, as shown in the table. However, these values were only representative of firms of middle-range sizes, not the ones at the tails, which were characterized by power-law distributions and scaling relations.

The capital stock was unique because the right tail did not exhibit a power law and there were multiple modes near the median value. A manual inspection of the data revealed that these multiple modes were the result of preference for rounded values such as 10,000. This phenomenon was also notable in other published data [36].

3.2. Bivariate Scalings between Size Measures

The dependencies between the size measures of business firms were characterized in terms of bivariate “allometric” scaling relations in a previous study [23], and we visualized them using the method described in the study. The data in the previous study were gathered by a credit reporting company independent of TDB. However, we expected the results to be similar between the two datasets, as the data were collected for the same system (Japanese firms).

Bivariate allometric scaling is denoted by $x \propto y^\gamma$, where γ represents the exponent, and x and y are two different size measures. We first consider the three size measures studied previously: the number of trading partners (k), number of employees (ℓ), and annual sales in millions of yen (s). As described in the previous study [23], the concept of allometric scaling relations incorporates scale-invariant relative fluctuations and is defined in terms of conditional distributions:

$$P(\ell|k) = \tilde{P}_{\ell|k}(\ell/k^{\gamma_1})/k^{\gamma_1}, \quad (1)$$

$$P(s|k) = \tilde{P}_{s|k}(s/k^{\gamma_2})/k^{\gamma_2}, \quad (2)$$

$$P(s|\ell) = \tilde{P}_{s|\ell}(s/\ell^{\gamma_3})/\ell^{\gamma_3}, \quad (3)$$

where $P(x|y)$ represents the probability density function of variable x conditional on a specific y -value; γ_1 , γ_2 , and γ_3 represent the scaling exponents; $\tilde{P}_{x|y}$ represents a scaling function for variable x conditional on the y -value, which indicates the distribution of the fluctuation ratio around the scaling line, $x \propto y^\gamma$.

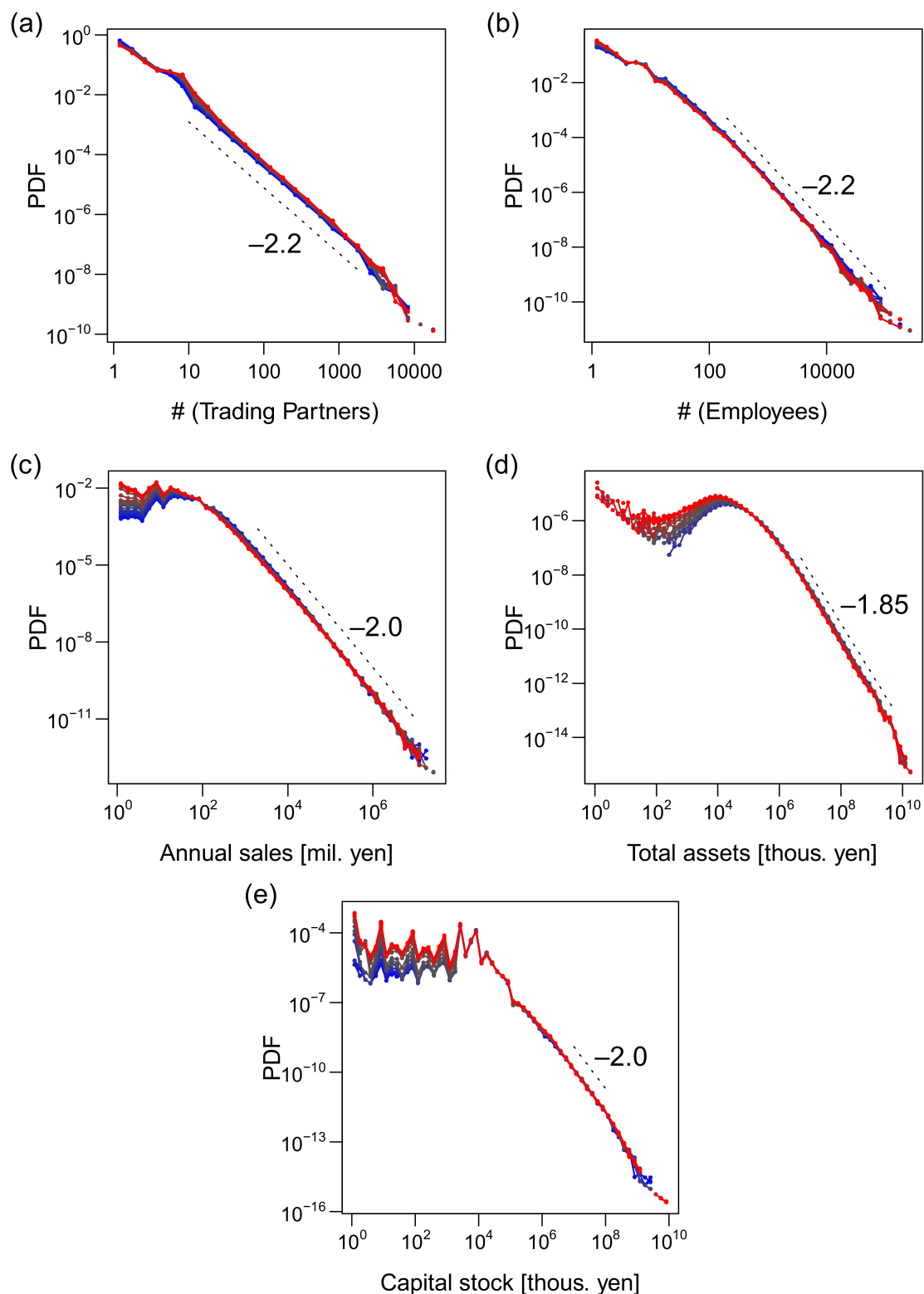


Figure 2. Probability distribution function (PDF) of the size variables for the period 1994–2015. The figures are plotted on log–log scales. Numerals near the dashed lines indicate the slope of the lines. The blue-to-red gradient of the color indicates the direction from old to new data. (a) Number of trading partners. (b) Number of employees. (c) Annual sales in millions of yen. (d) Total assets in thousands of yen. (e) Capital stock in thousands of yen. Data of total assets are plotted for the year 2000 or after, where the data were available.

All the probability distribution functions of the three size variables (k , ℓ , and s) were fat-tailed for large values in each year (see Figure 1a–c and Table S1). To avoid extreme values, which are usually observed in variables distributed in such a manner, we log-transformed the raw size data so that the variables were exponentially distributed. After the log-transformation, Equations (1)–(3) were more intuitive linear regression-type formulae, as discussed in detail in Appendix A:

$$\log l = \gamma_1 \log k + \varepsilon_{l|k}, \quad (4)$$

$$\log s = \gamma_2 \log k + \varepsilon_{s|k}, \quad (5)$$

$$\log s = \gamma_3 \log l + \varepsilon_{s|l}. \quad (6)$$

Here, the error terms, $\varepsilon_{l|k}$, $\varepsilon_{s|k}$, and $\varepsilon_{s|l}$, can be regarded as stochastic variables that correspond to $\tilde{P}_{l|k}$, $\tilde{P}_{s|k}$, and $\tilde{P}_{s|l}$, respectively. The mean of these “error terms” is generally not equal to zero, and the intercept term (as it is called in regression analysis) is included among them. These bivariate scaling laws, presented in a previous work [23], can be naturally interpreted as projections of a single three-dimensional scaling line onto two-dimensional planes, as illustrated in Figure 3a, around which the firms are densely distributed [37]. A similar concept of multivariate scaling was used in a study on the morphology of organisms [38].

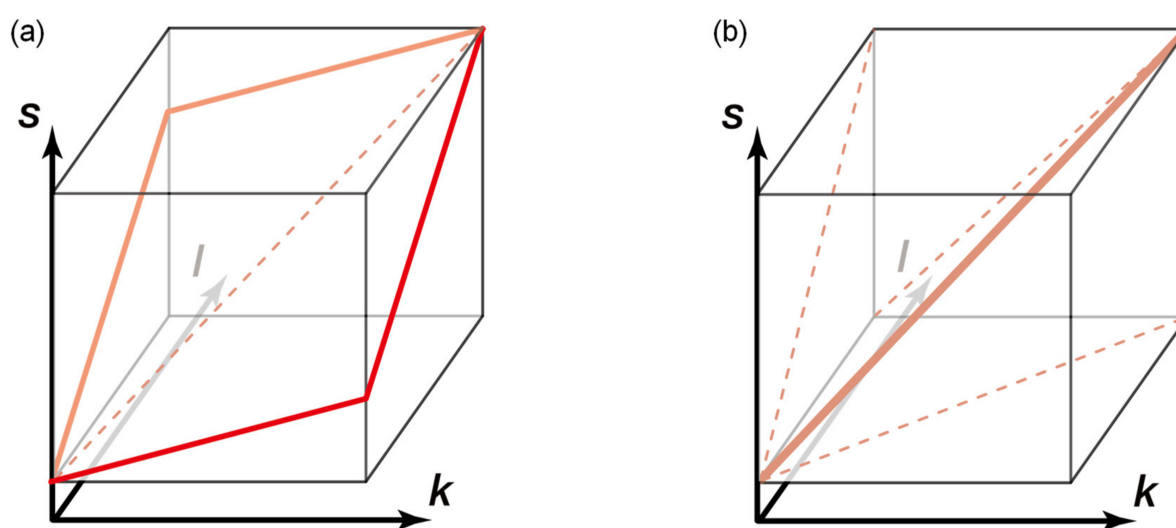


Figure 3. Schematics of the scaling relations. A firm is represented as a point in the three-dimensional phase space. Variables k , ℓ , and s represent the number of trading partners, number of employees, and annual sales, respectively. (a) Multivariate scaling $s \propto k^\alpha \ell^\beta$ is illustrated as a plane (red solid line). (b) Three bivariate scaling relations (indicated by the dashed lines) can be understood as projections of a single (red bold) “scaling line.” The plane in panel (a) must include the scaling line (red dashed), but this line clearly cannot determine a unique plane.

To test the validity of these relations (Equations (1)–(6)) in our data, we examined both the conditional quantiles and the distributions of the error terms, where the conditional quantile was defined as follows: for any given number, q , in the range of $(0, 1)$, the q -quantile of x conditional on y , which is denoted as $\langle x|y \rangle_q$, satisfies the following equation:

$$\int_0^{\langle x|y \rangle_q} P(x|y) dx = q. \quad (7)$$

If Equations (1)–(6) hold, all the quantiles of the same q -value conform to the power-law relation, $\langle x|y \rangle_q \propto y^\gamma$, and the curves of $\langle x|y \rangle_q$ plotted against y for different q -values should collapse into a single curve when vertically shifted. Furthermore, the distributions of $x/\langle x|y \rangle_{0.5}$ conditional on y should be identical for all y -values and amount to the distribution specified by $\tilde{P}_{x|y}$ with normalization by its median.

As shown in Figure 4a–c, the plots of $\log \langle x|y \rangle_q$ vs $\log y$ indicated linear relations for a wide range of y , indicating that the power-law relation, $\langle x|y \rangle_q \propto y^\gamma$ (abbreviated as $x \propto y^\gamma$), holds for all combinations of k , ℓ , and s . In Figure 4d–f, the distribution of relative fluctuation from the conditional median ($x/\langle x|y \rangle_{0.5}$) is plotted for eight intervals of equal length on the logarithmic scale of y . All the curves collapse into a single “scaling function” for all three cases, validating the scaling relations of Equations (1)–(3).

We next examined the total assets, a , with regard to k , ℓ , and s . The relation between k and the quantiles of a ($\langle a|k \rangle_q$) was not clearly linear, being slightly curved in a logarithmic plot, as shown in Figure 5a; however, the distributions of relative fluctuations from the conditional median ($a/\langle a|k \rangle_{0.5}$) were almost independent of k (Figure 5d). More obvious allometric scaling can be seen for ℓ – a (Figure 5b,e) and s – a (Figure 5c,f) relations. The total assets seemed to scale superlinearly with ℓ ($\gamma > 1$) and linearly with s ($\gamma = 1$). The power-law scaling exponents of a with regard to ℓ and s are hereafter denoted by γ_4 and γ_5 , respectively. Note that the distributions of relative fluctuations of a (i.e., $a/\langle a|s \rangle_{0.5}$) were wider for small values of s (blue curves) than for large values (red curves), as seen in Figure 5f. We also noted an increase in a slightly larger than a linear function of s for sales values more than 10^5 . Therefore, allometric scaling most precisely approximated the s – a relation for the range of s between 10^2 and 10^5 . We conclude that total assets scaled allometrically with ℓ and s , and the relation between a and k did not conform to the allometric scaling.

The capital stock, p , was not found to be in an allometric scaling with any of k , ℓ , s , or a , as shown in Figure 6. The vertically shifted curves of conditional quantiles did not agree for k – p , ℓ – p , and s – p relations (see Figure 6a–c), indicating that the forms of conditional distributions of p differed substantially as the values of k , ℓ , or s varied. Whereas the a – p relation appeared to be close to allometric in the plot of conditional quantiles, as shown in Figure 6d, the conditional relative fluctuations ($p/\langle p|a \rangle_{0.5}$) were not constant over the range of a (Figure 6e).

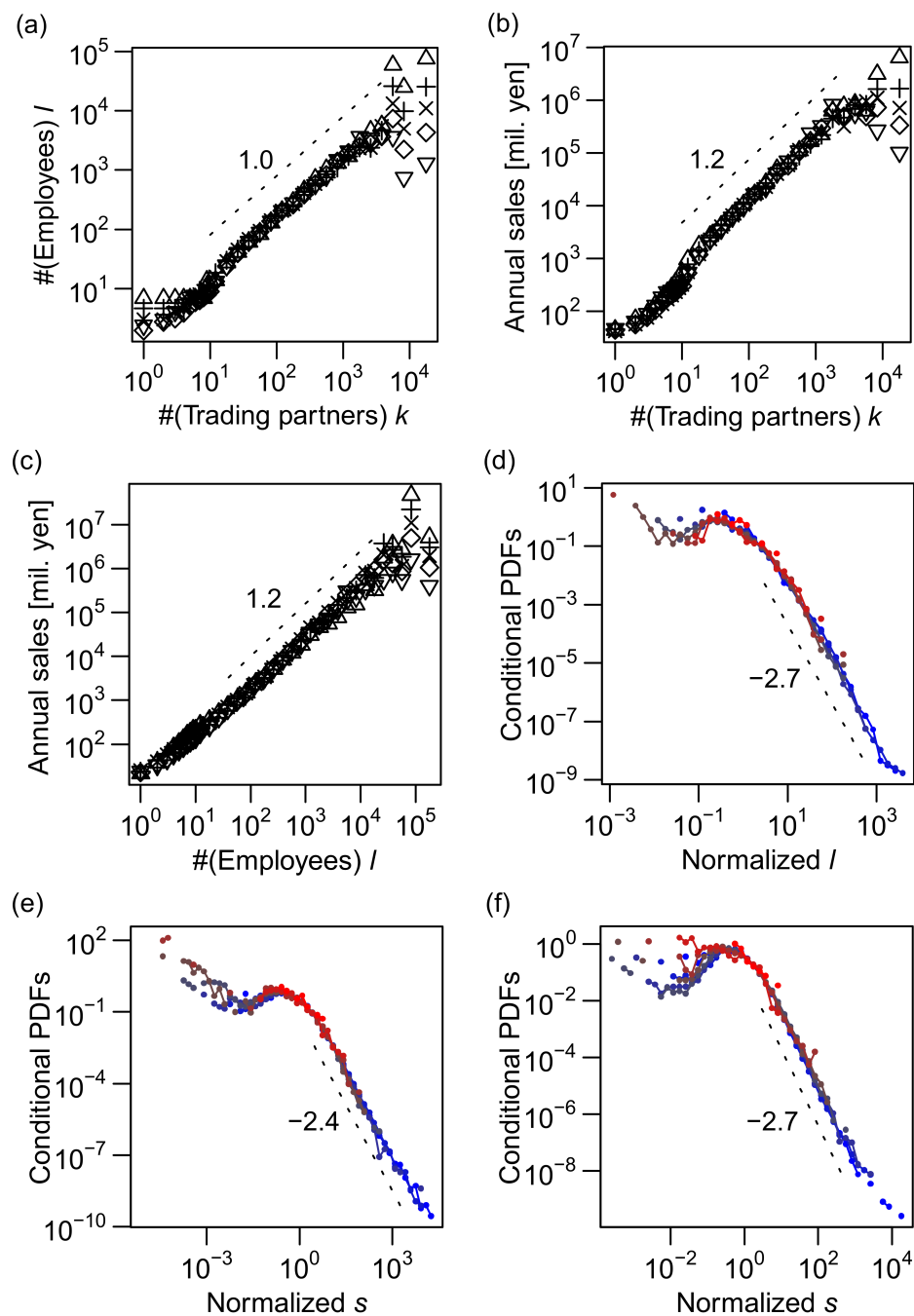


Figure 4. Scaling relations between pairs of measures: number of trading partners k , number of employees l , and annual sales in millions of yen s . All the data used in the figures are for 2014. (a–c) 0.95(∇), 0.75(+), 0.5(\times), 0.25(\diamond), and 0.05(\triangle) quantiles of the conditional distribution of l or s (vertical axis) are plotted against another variable k or l (horizontal axis) on a log–log scale. The horizontal axis is divided into intervals of identical lengths on the log scale (6 segments per a 10-fold interval such as 10–100) for k and l values of >10 and intervals of length unity on the linear scale for k and l values of ≤ 10 . Quantiles other than those of $q = 0.5$ (i.e., medians) are plotted with horizontal shift, so that all the curves pass through a point whose x -axis is slightly above 100. (d–f) Probability distributions (PDFs) of l or s conditional on k or l (i.e., $P(l|k)$, $P(s|k)$, and $P(s|l)$) for 8 different “bins” (i.e., mutually exclusive intervals) of the “explanatory” variable (k or l), normalized by their conditional medians, plotted on a log–log scale. The bins are obtained by evenly dividing the entire range (from the minimal value of unity to the maximum value of the variable shown in Table S1) into eight segments on the logarithmic scale. The color gradient of blue to red indicates low to high values of the explanatory variable.

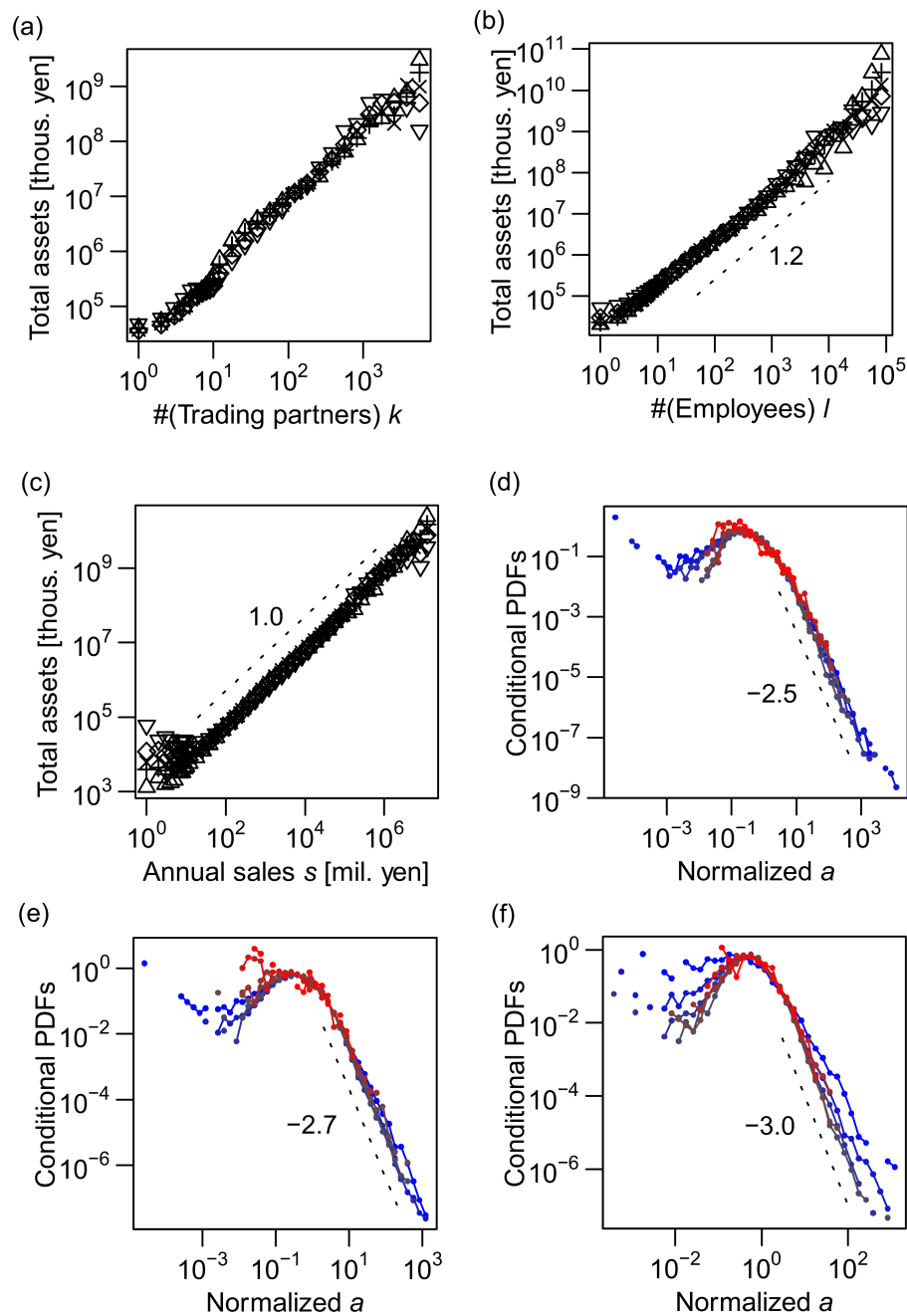


Figure 5. Scaling relations between total assets in thousands of yen a and number of trading partners k , number of employees ℓ , and annual sales in millions of yen s . All the data used in the figures are for 2014. (a–c) Vertically shifted quantiles of the conditional distribution of a (vertical axis) are plotted against another variable k , ℓ , or s (horizontal axis) on a log–log scale. (d–f) Probability distributions (PDFs) of a conditional on k , ℓ , or s (i.e., $P(a|k)$, $P(a|\ell)$, and $P(a|s)$, respectively) for 8 different “bins” of the “explanatory” variable (k , ℓ , or s), normalized by their conditional medians, plotted on a log–log scale. See the caption of Figure 4 for details on plotting.

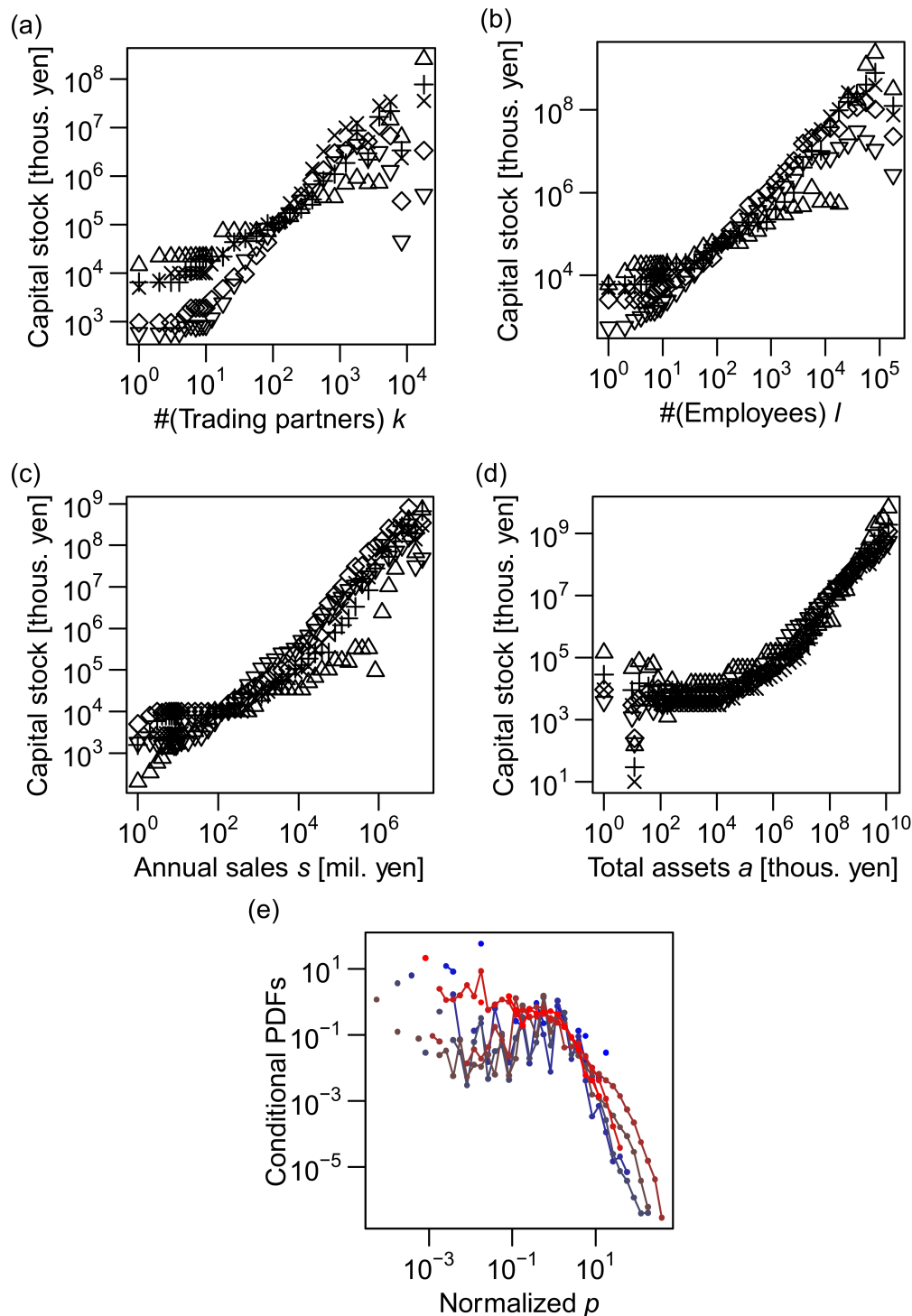


Figure 6. Scaling relations between capital stock and number of trading partners k , number of employees ℓ , annual sales in millions of yen s , and total assets in thousands of yen a . All the data used in the figures are for 2014. (a–d) Vertically shifted quantiles of the conditional distribution of p (vertical axis) are plotted against another variable k , ℓ , s , or a (horizontal axis) on a log–log scale. (e) Probability distributions (PDFs) of p conditional on a (i.e., $P(p|a)$) for 8 different “bins” of the “explanatory” variable (a), normalized by their conditional medians, plotted on a log–log scale. See the caption of Figure 4 for details on plotting.

3.3. Explaining Puzzling Scaling Exponents

Next, we studied the open question of the different power-law exponents in the conditional and marginal sales distributions, which has not been addressed previously. This is related to the asymmetric levels of relevance of the numbers of employees (ℓ) and trading partners (k) to the firm sales (s), as described in the following section.

It is well established that the distribution of annual sales s of firms roughly follows Zipf's law [13–15]; i.e., the probability density tail follows a power law, $P(s) \propto s^{-2}$. This was also observed in our data (see Figure 2c). However, for the conditional sales distributions, the power-law exponents increased to approximately 2.4 (for the number of trading partners (k); see Figure 4e) and approximately 2.7 (for the number of employees (ℓ); see Figure 4f), substantially varying from 2.0 (see Figure 7).

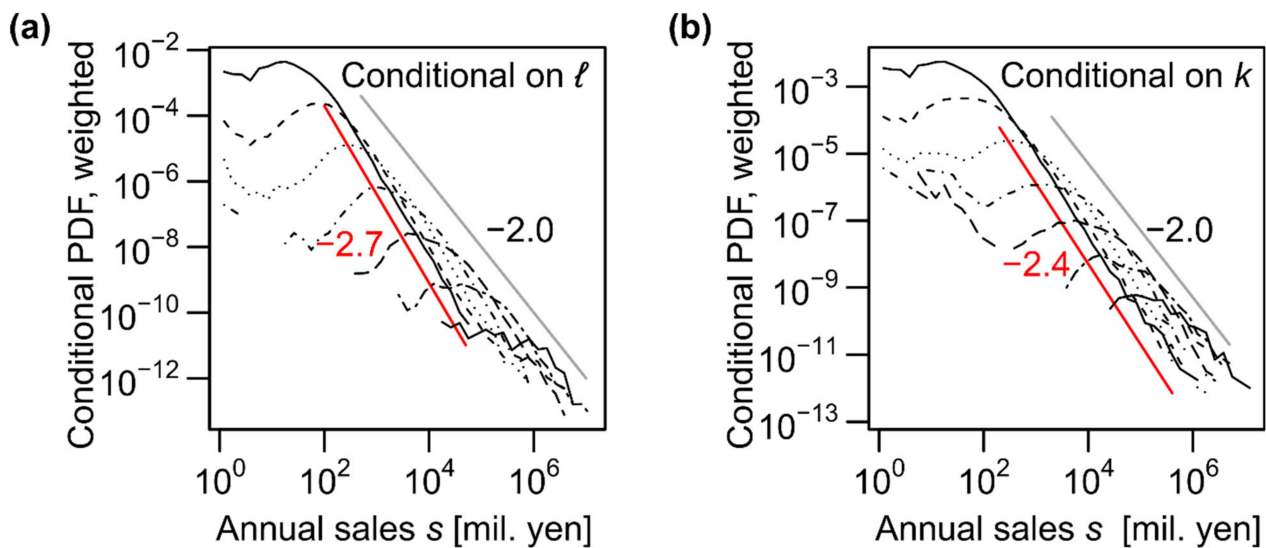


Figure 7. Apparent inconsistency of the power-law tail exponents for marginal and conditional PDFs, with an intuitive understanding based on Bayes' theorem. Conditional probability distributions (a) $P(s|\ell)$ and (b) $P(s|k)$, weighted with $P(\ell)$ and $P(k)$ (see Equation (12)), respectively, plotted on log–log scales. The entire range of k and ℓ is divided into eight levels corresponding to eight curves, so that each interval has an identical range on the logarithmic scale. The weight of an interval is defined as the average probability density in the interval. All the data used in the figures are for 2014.

These seemingly contradictory results are explained by Bayes' theorem. Considering number of employees ℓ and annual sales s , we can approximate the integral of $P(s|\ell)P(\ell)$ with respect to ℓ using the maximum values of $P(s|\ell)P(\ell)$:

$$P(s) = \int_{\ell} P(s|\ell)P(\ell)d\ell \sim P(s|\ell_{\text{lead}})P(\ell_{\text{lead}})\Delta\ell_{\text{lead}}, \quad (8)$$

where $\ell_{\text{lead}} = \arg\max_{\ell} P(s|\ell)P(\ell)$, such that $P(s|\ell_{\text{lead}})P(\ell_{\text{lead}})$ is the “leading order” contribution, and $\Delta\ell_{\text{lead}}$ is the width of ℓ at ℓ_{lead} , which is assumed to be a constant. In Figure 5a, the functional form of $P(s|\ell)P(\ell)$ is plotted for several typical values of ℓ based on real data. As shown, the envelope function of $P(s|\ell)P(\ell)$ followed a power law with an exponent close to -2.0 at its tail. Similar results were obtained for the number of trading partners (k), as shown in Figure 5b. A more rigorous derivation is presented in Appendix B.

Because the scaling relations between the size variables, i.e., $\ell \propto k^{1.0}$, $s \propto k^{1.2}$, and $s \propto \ell^{1.2}$ (see the following sections for the determination of these exponents), suggest symmetric relevance of k and ℓ to annual sales s , it is surprising that different exponents of the power-law tails for different conditional distributions (shown in Figure 5a,b) suggest strong asymmetry between the levels of relevance of k and ℓ to s . We discuss this novel feature in more detail in the next section in connection with a trivariate scaling relation.

3.4. Multivariate Scaling with Scale-Invariant Fluctuation

So far, we have studied bivariate relations between size measures of business firms. In spite of the intuitive picture of triple bivariate scalings as the two-dimensional projections of a single three-dimensional scaling line (Figure 3b), the transitivity of allometric scalings is not theoretically guaranteed; for example, $\ell \propto k^{\gamma_1}$ and $s \propto \ell^{\gamma_3}$ does not necessarily mean $k \propto s^{\gamma_1\gamma_3}$. Although the transitivity might be justified if the distribution of s is constant for an ℓ -value regardless of the k -value, such an assumption directly contradicts the previous results [23]. Therefore, trivariate relations should be considered to understand the theoretical basis of the transitivity of allometric scaling relations in the empirical datasets.

We first inspected the scaling of annual sales s with regard to the numbers of trading partners and employees, k and ℓ . Our analysis assesses the relative relevance of the number of trading partners (k) and the number of employees (ℓ) to the prediction of annual sales s . We generalize the allometric scaling relations to a multivariate relation, as follows:

$$\log s = \alpha \log k + \beta \log \ell + \varepsilon_{s|k,\ell} \quad (9)$$

or

$$P(s|k, \ell) = \tilde{P}_{s|k,\ell} \left(s / k^\alpha \ell^\beta \right) / k^\alpha \ell^\beta, \quad (10)$$

where α and β are the scaling exponents indicating the relative relevance of k and ℓ , respectively; $\varepsilon_{s|k,\ell}$ is a stochastic fluctuation term of $\log s$ conditional on both k and ℓ ; $P(s|k, \ell)$ represents the conditional probability density, which is dependent on both k and ℓ ; $\tilde{P}_{s|k,\ell}$ represents the scaling function. This multivariate scaling relation was proposed in [23] but not confirmed directly by real data. Assuming that Equation (7) is satisfied, the median value can be straightforwardly derived:

$$\log \langle s|k, \ell \rangle_{0.5} = \alpha \log k + \beta \log \ell + \langle \varepsilon_{s|k,\ell} \rangle_{0.5}, \quad (11)$$

where $\langle s|k, \ell \rangle_{0.5}$ represents the median value of s conditional on a specific set of k and ℓ values. If Equation (9) holds, the contour plots of conditional median sales $\langle s|k, \ell \rangle_{0.5}$ on the k - ℓ logarithmic coordinate plane should exhibit nearly regular and parallel contours. Because $\langle \varepsilon_{s|k,\ell} \rangle_{0.5}$ in Equation (9) is a constant, the multivariate scaling relation can be represented by a plane, as shown in Figure 3a. Importantly, this differs from the “scaling line” [37], which is implied by the three scaling relations for the pairs of variables (Figure 3b).

Although the relation is not a perfect plane but a surface because it is curved in the high- k and low- ℓ regions, as shown in Figure 8a, the data support the scaling of conditional median sales against the k and ℓ values (Equation (9)), particularly for medium or large values. Moreover, the statistical fluctuations around the median value are clearly independent of k and ℓ (Figure 8b). These facts indicate that the assumption of scaling represented by Equation (8) is valid for most k and ℓ values. When the distribution of scaled sales $s / k^\alpha \ell^\beta$ conditional on k and ℓ was plotted using the values of α and β estimated according to the data (Figure 8c), a remarkable fraction of the curves scaled with each other. Additionally, function $\tilde{P}_{s|k,\ell}$ was surprisingly stable across the years (Figure 8d). Thus, the scaling assumptions of Equations (7) and (8) are well supported by a large amount of available data.

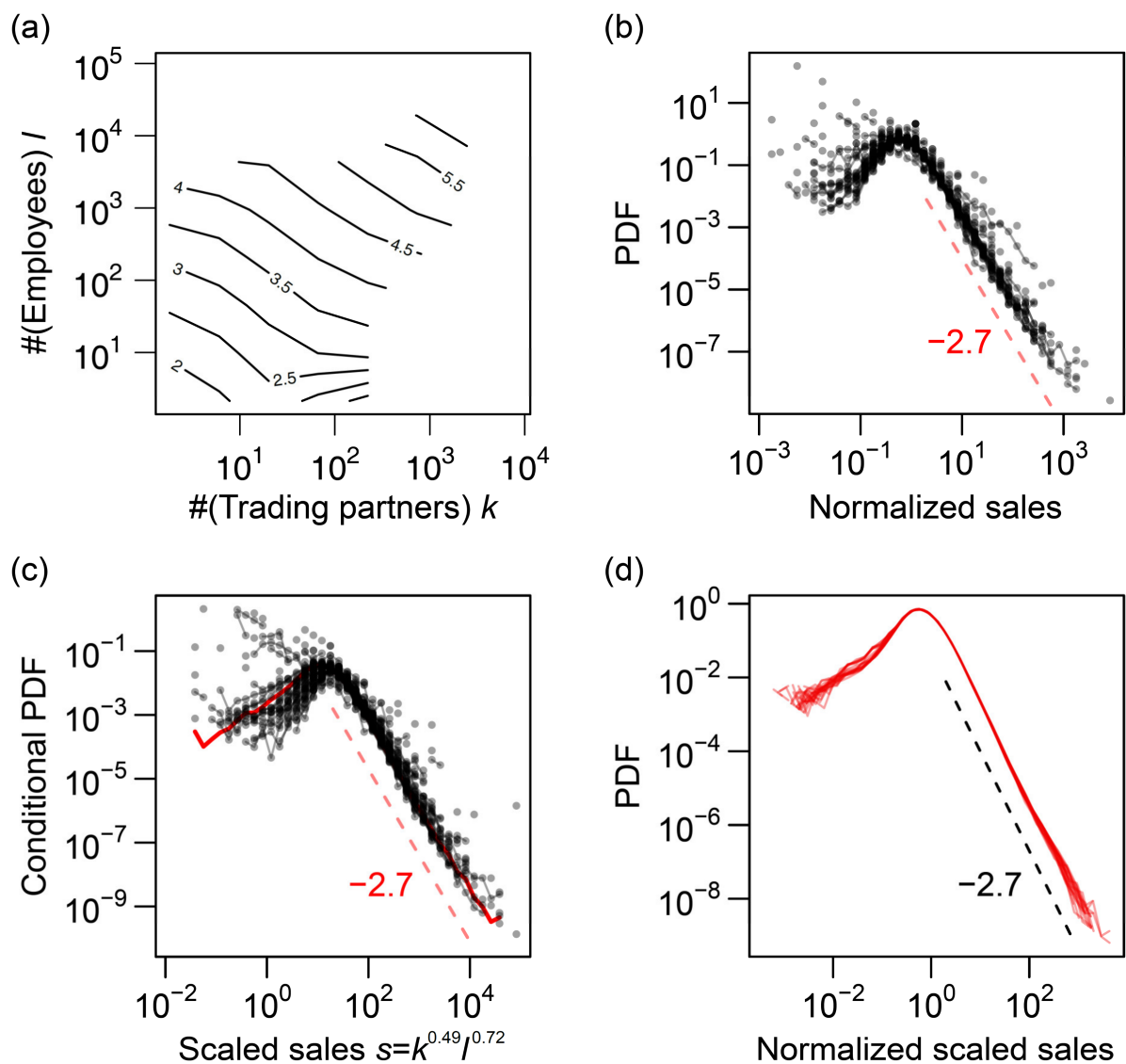


Figure 8. Multivariate scaling relations of annual sales s against number of trading partners k and number of employees ℓ . The plots in all the panels are on a log–log scale and present results for the 2014 data, unless otherwise mentioned. (a) Contour plot of the median values of annual sales s in millions of yen, conditional on both the number of trading partners, k (horizontal), and the number of employees, ℓ (vertical). The entire ranges of k and ℓ are divided into eight levels, so that each interval has an identical range on the logarithmic scale. The contours are obtained by linearly interpolating the log-transformed median sales values. Data of the conditional median from a grid with less than 10 samples were omitted in the plot. (b) Probability distributions (PDFs) of s conditional on both k and ℓ , normalized by their medians. The conditional distributions are obtained for grids of the conditioning variables where both dimensions are divided into intervals of identical length on a log scale (2 segments per a 10-fold interval). (c) PDFs of scaled sales $s/k^{\alpha}\ell^{\beta}$ conditional on both k and ℓ , where α and β represent the estimated exponents. The binning method of k and ℓ is identical to that for panel (b). Red lines indicate the marginal probability distribution of $s/k^{\alpha}\ell^{\beta}$. (d) Probability distributions of $s/k^{\alpha}\ell^{\beta}$ for each of the 22 years (1994–2015), normalized with respect to the medians. The exponents used are the estimated values for each year, as shown in Figure 10c.

The scaling exponent for k (α) was smaller than the scaling exponent for ℓ (β), as indicated by the moderate slopes of the contours compared with -1.0 in Figure 8a. We quantitatively verify this statement in the following section. This finding can be linked to the asymmetry of k and ℓ , as discussed in the previous section. In view of the magnitude of errors or fluctuations around the scaling relations, the distribution of residuals ($\varepsilon_{s|k}$ and $\varepsilon_{s|\ell}$, as defined in Equations (5) and (6)) was more fat-tailed when s was regressed against k ($\tilde{P}_{s|k}(\tilde{s}) \propto \tilde{s}^{-2.4}$; see Figure 4e) than when s was regressed against ℓ ($\tilde{P}_{s|\ell}(\tilde{s}) \propto \tilde{s}^{-2.7}$; see Figure 4f). This difference in the tails of the error distributions implies that the number of trading partners (k) was less useful in predicting the sales value than the number of employees (ℓ), which was indeed the case.

The trivariate allometric scaling, $s \propto k^\alpha \ell^\beta$, along with $\ell \propto k^{\gamma_1}$ and a power-law tail of the distribution of k , is sufficient to derive $s \propto \ell^{\gamma_3}$ and $k \propto s^{\gamma_1 \gamma_3}$ asymptotically under fairly mild assumptions (see Appendix B). In other words, the transitivity of bivariate scaling relations depends on whether the trivariate allometric scaling holds among the variables. We compared the triplets of (k, ℓ, a) and (ℓ, s, a) to exemplify this point. Note that the transitivity held among the latter but not among the former, because allometric scaling holds between k and ℓ and between ℓ and a , but not between k and a . We expected that the tri-variate allometric scaling would hold only for the latter. The results for these combinations of variables that are analogous to Figure 8b are plotted in Figure 9a–d. Although the deviations of a from the conditional median were independent of k and ℓ (panel (b)), the contour curves representing the conditional median of a were notably curved against k and ℓ (panel (a)). In contrast, contours for a conditional on ℓ and s were straight, parallel to each other, and evenly spaced (panel (c)), and most of the deviations might be approximated by a single function (panel (d)). Therefore, the trivariate scaling of a with regard to ℓ and s can be approximated by the following equation:

$$\log a = \alpha' \log \ell + \beta' \log s + \varepsilon_{a|\ell,s}. \quad (12)$$

However, the trivariate scaling among k , ℓ , and a cannot be described in an analogous way as expected.

Finally, we inspected the trivariate scaling of the capital stock, p , against other size measures. The results for the three variables (ℓ, s, p) are shown in Figure 9e,f. Both figures indicate clearly that these three variables were not in an allometric scaling, as we can expect from the non-allometric bivariate scaling property of p discussed in Section 3.2.

3.5. Robust Estimation of Scaling Exponents

Using the trivariate scaling with a scale-invariant relative fluctuation, we can derive several mathematical relations between the scaling exponents (see Appendix B). Here, we show that these relations allow the asymptotic scaling exponents to be estimated robustly against scaling for small firms that does not conform to the power law.

By assuming the trivariate scaling formulated by Equation (7), bivariate scaling of a similar form between k and ℓ , and power-law tails of several conditional and marginal distributions, we can determine the joint probability distribution of k , ℓ , and s . Thus, one can obtain a set of relations between numerous scaling exponents. Central to these are

$$\gamma_2 = \alpha + \beta \gamma_1, \quad (13)$$

and

$$\gamma_3 = (\alpha + \beta \gamma_1) / \gamma_1, \quad (14)$$

where γ_1 , γ_2 , and γ_3 represent the k – ℓ , k – s , and ℓ – s scaling exponents, respectively (see Appendix B for the derivations). These theoretical equations are useful for estimating the exponents, as determining the threshold size over which the behaviors of firms can be approximated as asymptotic is critical to such an estimation. The aforementioned relations are satisfied only when the data with $k < 100$ or $\ell < 100$ are excluded (see below), indicating that firms do not follow exactly the same power-law scaling below and above the threshold.

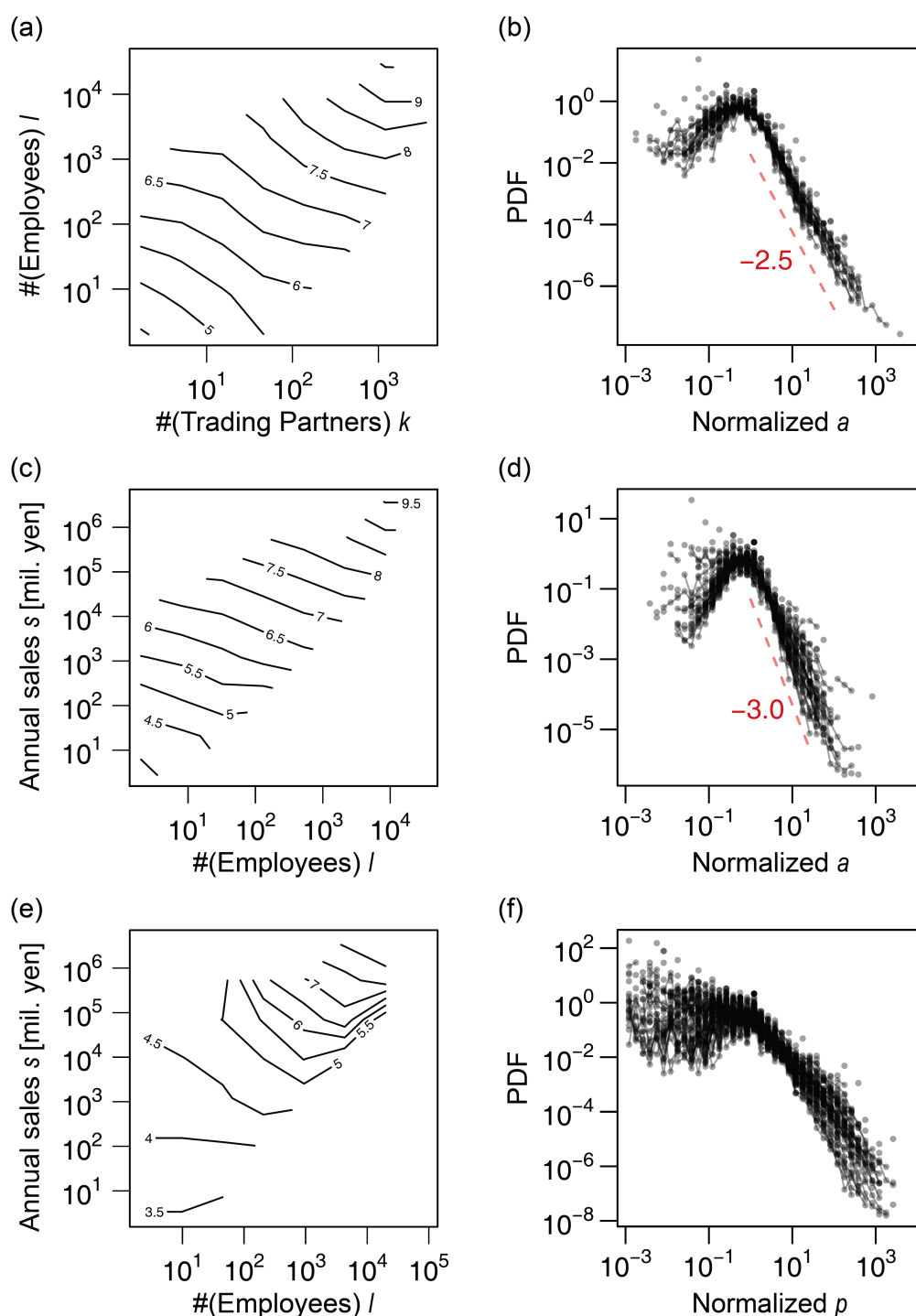


Figure 9. Multivariate scaling relations among number of trading partners k , number of employees ℓ , annual sales s , total assets a , and capital stock p . The plots in all the panels are on a log–log scale and present results for the 2014 data. (a) Contour plot of the median values of total assets a in thousands of yen, conditional on both the number of trading partners, k (horizontal), and the number of employees, ℓ (vertical). (b) Probability distributions (PDFs) of a conditional on both k and ℓ , normalized by their medians. (c,e) Contour plot of the median values of total assets a or capital stock p in thousands of yen, conditional on both the number of employees, ℓ (horizontal), and the annual sales, s , in millions of yen (vertical). (d,f) Probability distributions (PDFs) of a or p conditional on both ℓ and s , normalized by their medians. In contour plots, the entire ranges of the horizontal and vertical axes are divided into eight levels, so that each interval has an identical range on the logarithmic scale. See the caption of Figure 8 for further details on plotting.

We observed deviations from the scaling relations, particularly when the number of trading partners (k) and the number of employees (ℓ) were less than 10 or the annual sales (s) were less than 100 (Figure 4a–c, Figures 5a–c and 8a; also see [37]). Because smaller firms dominated the data (Figure 2), their deviation from the scaling relations substantially affected the estimation of the scaling exponents. When all the available data were included in the regression analyses, there was a clear difference between the expected value of γ_3 (dashed purple line) based on Equation (11) and the value obtained via direct estimation (solid purple line), as shown in Figure 10a.

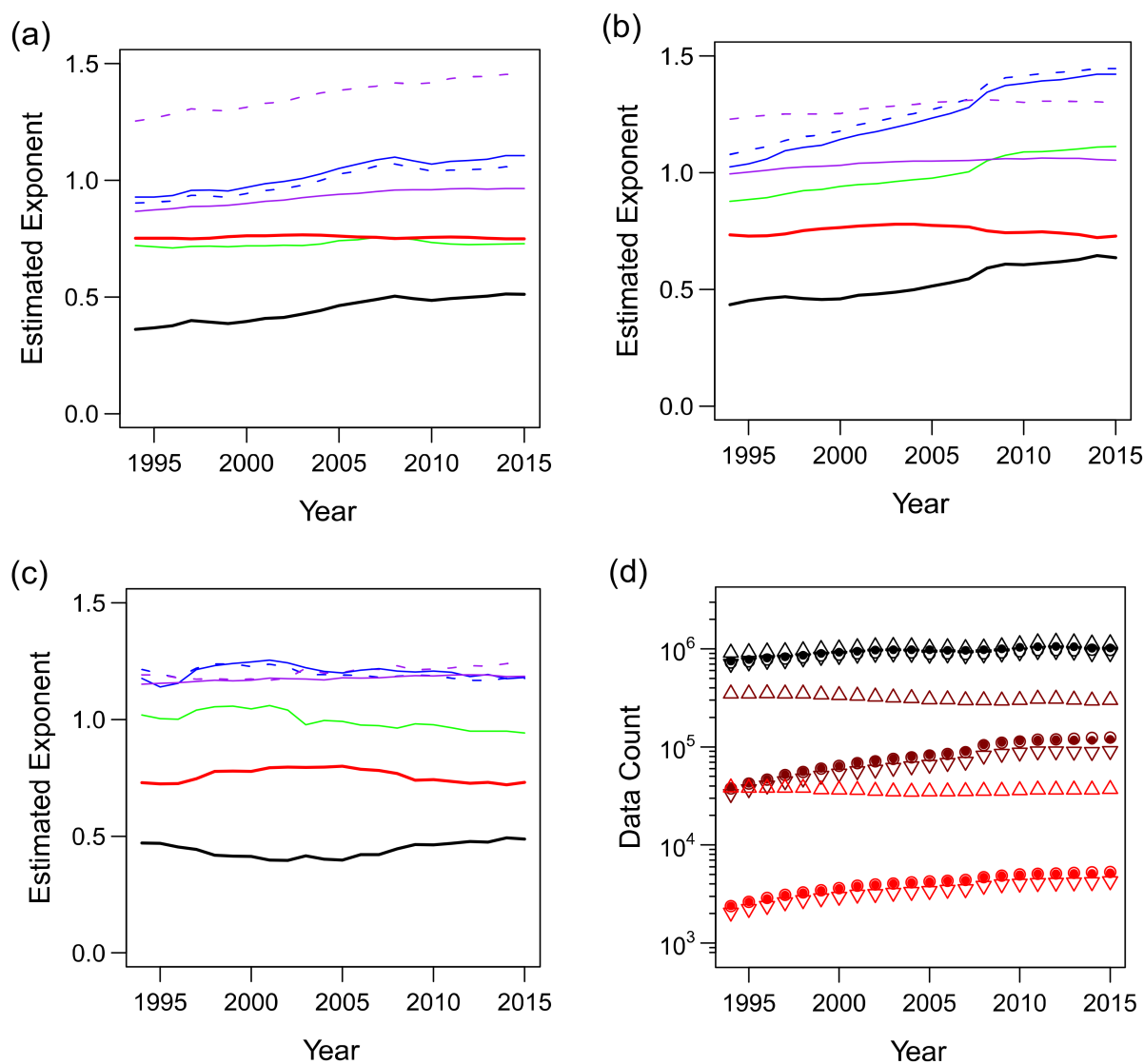


Figure 10. Estimated exponents of the trivariate scaling relations of s against k and ℓ for the period 1994–2015. Estimated exponents γ_1 in $\ell \propto k^{\gamma_1}$ (green), γ_2 in $s \propto k^{\gamma_2}$ (blue), γ_3 in $s \propto \ell^{\gamma_3}$ (purple), and α (black) and β (red) in $s \propto k^\alpha \ell^\beta$ are plotted. Additionally, the expected values of γ_2 and γ_3 derived mathematically from α , β , and γ_1 in the case of perfect scaling (Equations (10) and (11)) are juxtaposed (blue and purple dashed lines, respectively). (a) All the available data are used to estimate the exponents. (b) Only the data for which the “explanatory” variable is >10 are used. (c) Only the data for which the “explanatory” variable is >100 are used. (d) Plot of the size of samples from which the exponents are estimated against the year. The black, brown, and red points indicate the thresholds of 0 (no exclusion), 10, and 100, corresponding to panels (a–c), respectively. The hollow (\circ) and filled (\bullet) circles and upward (\triangle) and downward (∇) triangles represent the scaling of ℓ vs. k , s vs. k , s vs. ℓ , and s vs. k and ℓ , respectively.

We performed a standard regression analysis using R (ver. 3.1.2) to estimate the scaling exponents in the bivariate and multivariate scaling relations of the firm data. See Appendix A for the rationales for applying the regression analysis to the problem of estimating scaling exponents of power-law relations with scale-invariant relative fluctuations. To reject the data of small firms that did not fit the “linear” assumption of the model, we excluded the data with low values of “explanatory” variables in the regression formulae by applying a threshold. We used two values for possible thresholds: 10 and 100. The threshold was determined according to the consistency of the resulting multivariate scaling exponents with the bivariate ones. If the “explanatory” variables on the right-hand sides of Equations (4)–(7) were below the threshold, the datum was neglected. Although a considerable proportion of the data might be excluded from the analysis, this process ensured that the final set of exponents conformed to the model assumptions.

The results of regression of s against k and ℓ using thresholds of 10 and 100 are presented in Figure 10b,c, respectively. As shown, γ_1 nearly reached 1.0, and the direct and indirect estimates of γ_3 agreed only when the threshold was set as 100. This suggests that the threshold should be ≥ 100 for k and ℓ . We preferred to employ a smaller threshold value owing to the large sample size; thus, we adopted the value of 100. The sample sizes, before and after the threshold was applied, are shown in Figure 10d. After the threshold of 100 was applied, as shown in Figure 10c, the estimated γ_1 value was between 0.94 and 1.06, and the estimated γ_2 and γ_3 values were both between 1.14 and 1.25. We estimated the values of these exponents (averaged for the entire period) as $\gamma_1 \approx 1.0$, $\gamma_2 \approx 1.2$, and $\gamma_3 \approx 1.2$. However, the slow and systematic fluctuations of the estimated α (black line) and β (red line) values, i.e., the trivariate scaling exponents defined in Equation (7), were noticeable.

Despite our use of linear regression, estimation of the uncertainty requires a nonparametric method, because the “error terms” are not normally distributed even with their homoscedasticity. Thus, we used the bootstrap method [39] to determine the confidence intervals (CIs). Resampling was performed 10,000 times, and the resampling size was identical to the sample size. The 95% CIs were determined as the 2.5- and 97.5-percentiles of the bootstrap distribution.

Considering the uncertainty of the estimations, our finding of $\beta > \alpha$ was true for all the years, as shown in Figure 12a, where the estimated scaling exponents for different years are plotted with the uncertainties based on the bootstrap method. For example, the results for 2014 were $\alpha = 0.49$ (95% CI (0.455, 0.531)) and $\beta = 0.72$ (95% CI (0.697, 0.744)). Thus, we conclude that the number of employees (ℓ) has a higher relevance to sales s than the number of trading partners (k).

Although the changes were not large and inequality $\alpha < \beta$ was invariably satisfied, exponents α and β became smaller and larger, respectively, in the 2000–2005 period compared with the 2013–2015 period. The difference was “significant” in the sense that the 95% CIs did not overlap. Additionally, α and β were negatively correlated, which we expected from Equation (10) and the relatively robust γ_1 and γ_2 (Figure 10c).

We then applied the aforementioned procedure to the trivariate scaling of total assets a with ℓ and s . The results are shown in Figure 11. In contrast to the scaling of s with k and ℓ , there was no obvious improvement in the estimation of α' and β' by the use of thresholding. This should be related to better fitting of the trivariate data of (ℓ, s, a) to the scaling surface hypothesis (see Figure 3a) compared to the (k, ℓ, s) data, as seen in Figures 8a and 9c. The degree of the time variation of α' and β' was noticeably smaller in the (ℓ, s, a) data than in the (k, ℓ, s) data. The estimated α' (exponent for ℓ) and β' (exponent for s) values were between 0.17 and 0.24 and between 0.82 and 0.92, respectively, in all years and for all threshold values that we applied. In contrast, the bivariate scaling exponents were estimated without data with the s - or a -value less than 100 to exclude the data that did not conform to the power law, as seen in Figure 5b,c. Therefore, the final values of estimated γ_4 (for ℓ - a scaling) and γ_5 (for s - a scaling) were between 1.16 and 1.19 and between 1.01 and 1.04, respectively, as shown in Figure 11c.

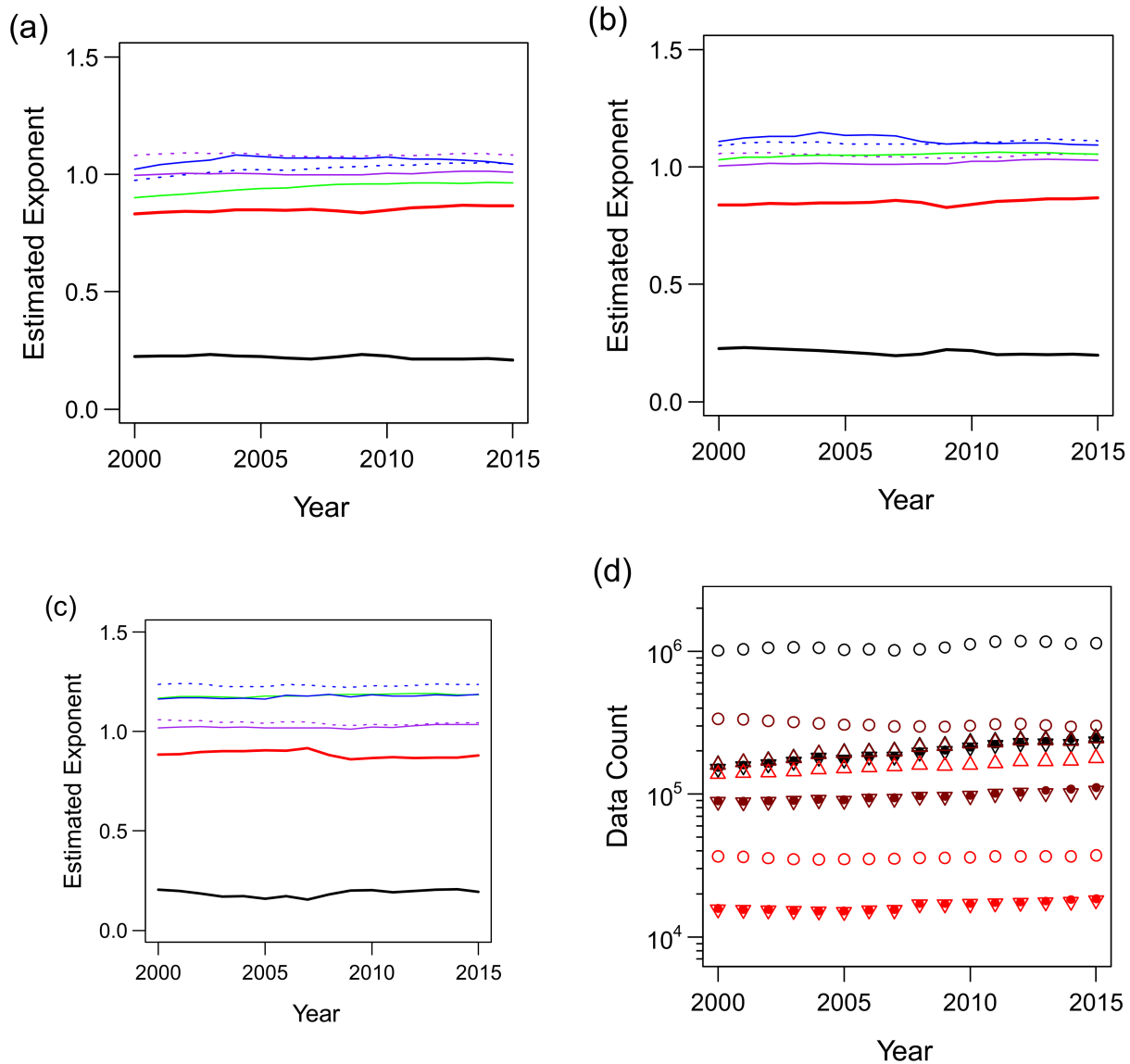


Figure 11. Estimated exponents of the trivariate scaling relations of a against ℓ and s for the period 2000–2015. Estimated exponents γ_3 in $s \propto \ell^{\gamma_3}$ (green), γ_4 in $s \propto \ell^{\gamma_4}$ (blue), γ_5 in $a \propto s^{\gamma_5}$ (purple), and α' (black) and β' (red) in $a \propto \ell^{\alpha'} s^{\beta'}$ are plotted. Additionally, the expected values of γ_4 and γ_5 derived mathematically from α' , β' , and γ_3 in the case of perfect scaling (Equations (10) and (11)) are juxtaposed (blue and purple dashed lines, respectively). (a) All the available data are used to estimate the exponents. (b) Only the data for which the “explanatory” variable is >10 are used. (c) Only the data for which the “explanatory” variable is >100 are used. (d) Plot of the size of samples from which the exponents are estimated against the year. The black, brown, and red points indicate the thresholds of 0 (no exclusion), 10, and 100, corresponding to panels (a–c), respectively. The hollow (\circ) and filled (\bullet) circles and upward (\triangle) and downward (∇) triangles represent the scaling of s vs. ℓ , a vs. ℓ , a vs. s , and a vs. ℓ and s , respectively.

3.6. Scaling Exponents Related to National Economic Conditions

As indicated in the previous section, there were significant changes (i.e., changes beyond the CI) in α and β during the period 1994–2015, although the source of this fluctuation remains to be explored. Here, we only test the hypothesis that national macroeconomic conditions are related to the changes in these scaling exponents.

Interestingly, α and β appeared counter- and pro-cyclical, respectively; i.e., they were apparently positively and negatively correlated to the nominal GDP of the country [40], respectively, as shown in Figure 12a. The GDP was selected here because its fluctuation

cycle is longer than those of other economic indices, such as the Indexes of Business Conditions reported by the Cabinet Office of the Government of Japan [41]. A closer inspection revealed that α was enlarged when the nominal GDP decreased (Figure 12b) and that β decreased almost simultaneously with the GDP, whereas its increase was delayed with respect to the GDP expansion (Figure 12c).

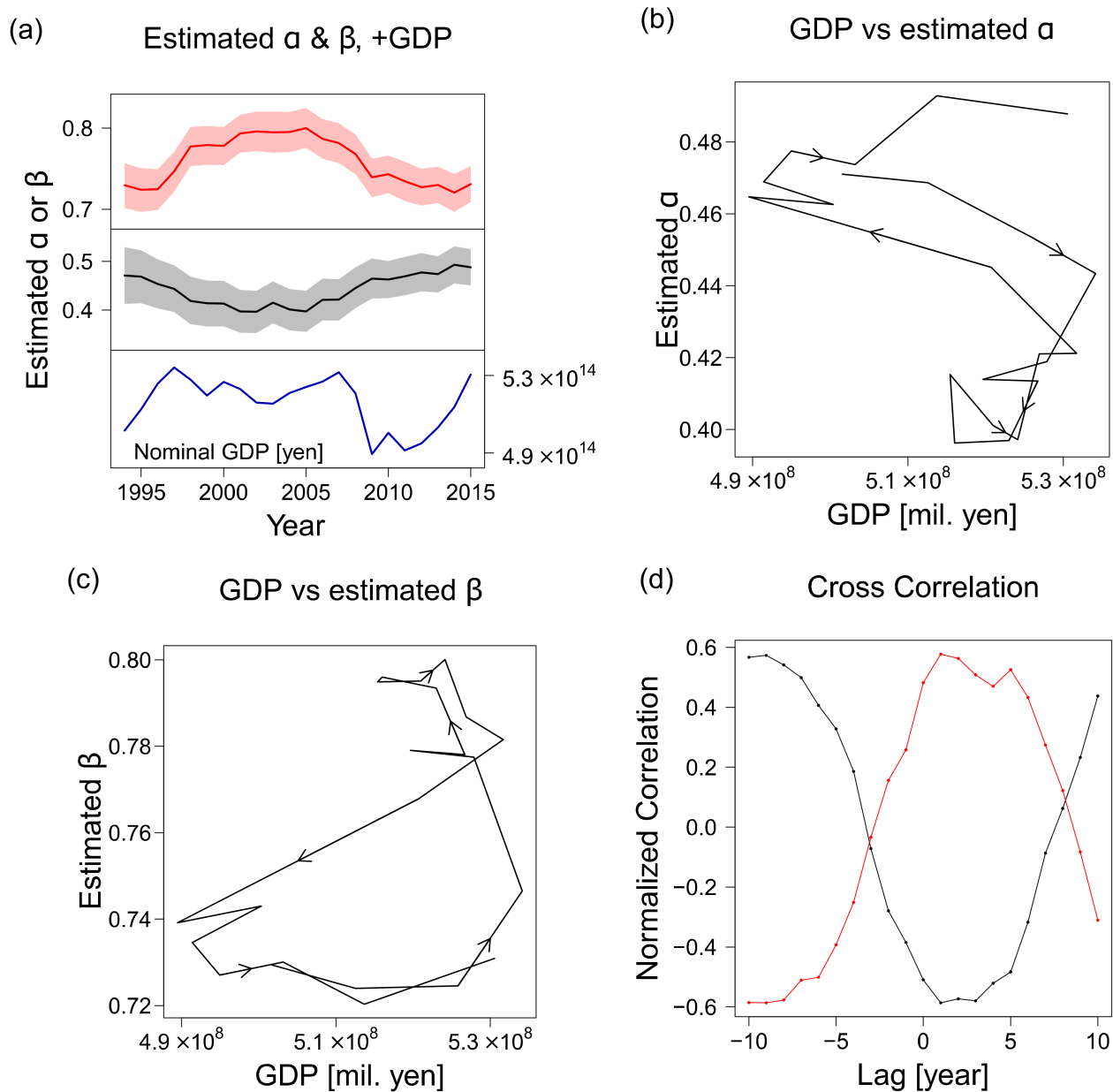


Figure 12. Estimated exponents α (for number of trading partners k) and β (for number of employees ℓ) with the country's nominal GDP for different years in 1994–2015. (a) Estimated exponents α (for number of trading partners k) and β (for number of employees ℓ) and the country's nominal GDP for different years in the period 1994–2015. The bandwidths indicate the 95% confidence interval of the estimation obtained via the bootstrap method. (b) Estimated α with respect to the GDP. The arrows indicate the direction of time evolution. (c) Estimated β with respect to the GDP. (d) Cross correlations between the nominal GDP and α or β are plotted with respect to the time lag.

We calculated the cross-correlation between the national GDP and the scaling exponents to evaluate the delay of the effect of the former on the latter. The normalized cross-correlation, $CC_\tau[x, y]$, was defined by the Pearson's correlation coefficient applied to lagged time-series data $x(t)$ and $y(t + \tau)$ defined for discrete time $t_0 \leq t \in \mathbb{Z} \leq t_{\text{end}}$:

$$CC_\tau[x, y] \equiv \frac{1}{N-1} \sum_{T_\tau} \tilde{x}(t) \tilde{y}(t + \tau) / \sqrt{\frac{1}{N-1} \sum_{T_\tau} \tilde{x}(t)^2} \sqrt{\frac{1}{N-1} \sum_{T_\tau} \tilde{y}(t + \tau)^2} \quad (15)$$

where $T_\tau \equiv \{t \in \mathbb{Z} | t_0 + \tau \leq t \leq t_{\text{end}} + \tau\}$, N represents the length of the time series (the number of elements in truncated set T_τ), and

$$\tilde{x}(t) \equiv x(t) - \frac{1}{N} \sum_{T_\tau} x(t) \text{ and } \tilde{y}(t) \equiv y(t) - \frac{1}{N} \sum_{T_\tau} y(t). \quad (16)$$

As shown in Figure 12d, we plotted the cross correlations, $CC_\tau[\text{GDP}, \hat{\alpha}]$ and $CC_\tau[\text{GDP}, \hat{\beta}]$, with respect to τ (the time lag in years) to quantify the delay of the changes in the estimated α and β values with respect to the GDP. Both the cross correlations peaked at $\tau = 1$ (negatively and positively, respectively), with Pearson correlation coefficients as large as -0.59 and 0.58 . The second-highest peak was present in both cross-correlation series, indicating that the extent of the delay in the exponent changes differed between the increasing GDP and the decreasing GDP. The reversal of the sign at a τ value far from 0 was an artefact of the time window, which covered only slightly more than one economic cycle. The foregoing results suggest that the exponents can be predicted using the GDP of the preceding year. However, we could not eliminate the possibility that the correlations were mere coincidences, as the dataset covered only slightly more than one business cycle. Considering that the cycle was approximately 20 years long, an additional decade of data or a dataset from another country may be needed to verify this trend.

We suspect that the relative importance of inter-firm trades in predicting sales is mechanistically affected by the GDP. It is reasonable that in a recovering economy, selling goods produced with labor is relatively easy, and increasing the number of employees is advantageous for expanding production, which leads to an increase in β . In contrast, in an economic depression, the trading partners to which the firm can sell products may become more crucial for maintaining its sales; in such a situation, weight may be added to α . However, this statement is not supported by the present study and is yet to be verified.

4. Discussion

We analyzed the multivariate scaling relations between firm size measures. First, the marginal distributions of k , ℓ , s , a , and p were examined, and the former four were found to have a power-law tail while the right tail of the latter one did not exhibit a clear power law. Second, we examined the bivariate relations between the size measures to uncover that relations for pairs of k , ℓ , s , and a could all be approximated by an allometric scaling except the k - a relation, while the capital stock, p , was never in an allometric scaling with other size measures. Based on the empirically verified model of allometric scaling, seemingly inconsistent power-law exponents for the right tails of conditional and marginal distributions were theoretically explained. Furthermore, theoretical considerations clarified that the trivariate allometric scaling is related to the transitivity of bivariate allometric scalings. This was exemplified by the comparison of trivariate data of (k, ℓ, s) and (ℓ, s, a) , which conformed to trivariate allometric scaling, with the data of (k, ℓ, a) which did not fit the allometric scaling model. Building on the theoretical development, a method was proposed and applied to the data to estimate the scaling exponents for each year. We observed only small variations when data from different years were compared. The fluctuations of the scaling exponents α and β for the data of (k, ℓ, s) during a 22-year period apparently followed the country's GDP.

We argue that our results might be highly valuable in making the existing research results more robust and clearer. Some reported results regarding corporate finance were not robust when different size measures were used as a control variable indicating corporate size [19], suggesting that a single variable indicating the firm size is not sufficient to control

the “size effects.” An allometric scaling between size measures, e.g., $x \propto y^\gamma$, means that the fluctuation or deviation from the average relation (x/y^γ) is independent of y . Therefore, such an adjusted ratio between size measures in allometric scalings could be added to analyses as a possible factor improving the robustness of the existing methods against the use of alternative corporate size measures, bridging the studies using different size measures of firms as a control variable.

Scaling relations for data of more than three measures of size were not examined in our study. Although it is possible to conceive of such a theoretical development, empirical verification of the statistical model for a high-dimensional space is expected to face a difficulty known as the curse of dimensionality [42]. Furthermore, we do not expect theories regarding four-dimensional allometric scaling to be useful in understanding empirical data, since the allometric scaling was found to be broken in the three-dimensional data of (k, ℓ, a) .

The current analyses are based on aggregated data from all industries. It is highly possible that firms in different industries or markets have different scaling relations. Nevertheless, our results would still serve as a basis for investigating the characteristics of different industries and markets, since the results can be regarded as representing average behaviors of firms across different sectors, calculated from nearly the most comprehensive dataset. Therefore, it would be a promising research direction to study the uniqueness of different groups of firms relative to the average behaviors.

It is our current expectation that similar results could be obtained if it were possible to conduct the same analysis on different datasets of business firms operating in areas other than Japan. Tests of the reproducibility of our results would be highly valuable for understanding the diversity and universality of business firms. While analysis regarding the sales, total assets, and employees of firms can be performed with minimal effort owing to the abundance of information, it would be more difficult to include the number of trading partners because data related to trading between firms are often missing in existing datasets. Text mining of published corporate reports may be a promising approach for compiling datasets of inter-firm trading [43]. We also encourage the application of our method to other systems that exhibit scale-invariant relative fluctuations, such as metropolises and cities [44,45].

We confirmed that the fluctuation of s —conditional to k and l and relative to the median—was invariant, directly verifying the hypothesis proposed in the previous work [23]. This scale-invariance of the relative fluctuation is highly suggestive of the similarity between firms of different sizes and the renormalizability of internal organizations or inter-firm relationships with regard to complex networks [46,47].

Verifying previously hypothesized stylized facts and checking the mathematical consistency of independently known power-law exponents help to validate or refute possible theories regarding business firms. Numerous models [35,48–61] have been suggested to explain only a few stylized facts relating to firms, and it appears that new criteria are needed to select the models empirically. Thus, our results for multivariate scaling should be incorporated into subsequent theoretical considerations. Theories for the scaling relations in other complex systems, such as animal bodies and cities, may benefit from this study because fractal-like hierarchical organization is a pervasive design in such systems (including business firms) [15,32,33,62]. However, the data are far more abundant for firms than for other systems. The multidimensional notion of size may introduce prospects for devising general theories and modeling principles for such complex systems.

Supplementary Materials: The following are available online at <https://www.mdpi.com/1099-4300/23/2/168/s1>, Table S1: Descriptive statistics of the final data of firm size measures for different years.

Author Contributions: Conceptualization, M.T., S.H., H.T., and Y.K.; methodology, Y.K., H.T., and S.H.; software, Y.K.; validation, Y.K., H.T., and M.T.; formal analysis, Y.K.; investigation, Y.K.; resources, M.T.; data curation, Y.K. and M.T.; writing—original draft preparation, Y.K.; writing—review and editing, S.H., M.T., and H.T.; visualization, Y.K.; supervision, M.T., S.H., and H.T.; project

administration, M.T.; funding acquisition, M.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Center for TDB Advanced Data Analysis and Modeling, Tokyo Institute of Technology, and partially funded by (i) the Grant-in-Aid for Scientific Research (B), JSPS, grant number 26310207; (ii) JST, Strategic International Collaborative Research Program (SICORP) on the topic of “ICT for a Resilient Society” by Japan and Israel; (iii) MEXT, as “Exploratory Challenges on Post-K computer (Study on multilayered multiscale spacetime simulations for social and economical phenomena)”.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from Teikoku Databank, Ltd. (Japan) and are available from the authors with the permission of Teikoku Databank, Ltd. (Japan).

Acknowledgments: We thank O. Levy (Bar-Ilan University) for discussions.

Conflicts of Interest: Teikoku Databank, Ltd. supported our research by providing the data related to Japanese business firms and by financially supporting the Center for TDB Advanced Data Analysis and Modeling, Tokyo Institute of Technology for academic research purposes. Teikoku Databank, Ltd. did not participate in the research or the preparation of the manuscript, except for the data collection.

Appendix A. Rationales for Estimation of Scaling Exponents

Here, we consider the rationales for the estimations of the scaling exponents using the standard regression analysis.

We begin with the bivariate scaling relations, which were formulated with Equations (1)–(3) or (4)–(6) and confirmed to be present in our data:

$$P(x|y) = \tilde{P}_{x|y}(x/y^\gamma)/y^\gamma,$$

where $P(x|y)$ represents the probability density of x conditional on y , $\tilde{P}_{x|y}$ represents a probability density function, and γ is a positive constant. When variable \tilde{x} is defined as $\tilde{x} \equiv x/y^\gamma$, we have

$$P(\tilde{x}|y) = \tilde{P}_{x|y}(\tilde{x}),$$

given that the probability density should satisfy normalization $\int P(\tilde{x}|y)d\tilde{x} = 1$. Variable \tilde{x} does not depend on the y -value; thus, \tilde{x} is independent of y .

This indicates that exponent γ can be estimated by finding the value that maximizes the degree of independence of \tilde{x} and y . One of the simplest indices for measuring the dependence between stochastic variables is the Pearson’s product–moment correlation coefficient. We can obtain $\hat{\gamma}$, i.e., the estimated value of γ , as follows:

$$\hat{\gamma} = \arg \min_{\gamma} \left(\text{Cor} \left[\log \left(\frac{x}{y^\gamma} \right), \log y \right] \right)^2, \quad (\text{A1})$$

where $\text{Cor}[\cdot, \cdot]$ is the correlation coefficient of the two terms. We apply the log-transformation to the raw data because the distribution of x conditional on y is heavy-tailed, as shown in Figure 4d–f. In such heavy-tailed distributions, a few extreme values can substantially affect the correlation coefficient compared with the case of the normal or exponential distribution.

The minimization of the absolute value of correlation coefficients amounts to the linear least-squares regression of $\log x$ against $\log y$. Indeed, when the correlation coefficient is zero, the covariance is also zero, and if we apply transformations $x \leftarrow x/\bar{x}$ and $y \leftarrow y/\bar{y}$, where \bar{x} and \bar{y} represent the geometric means of x and y , respectively,

$$(N-1) \times \text{Cov} \left[\log \left(\frac{x}{y^\gamma} \right), \log y \right] \equiv \sum_i \log y_i \times (\log x_i - \gamma \log y_i) = 0$$

or

$$\sum_i \log x_i \times \log y_i = \gamma \sum_i (\log y_i)^2. \quad (\text{A2})$$

Here, $\text{Cov}[\cdot, \cdot]$ represents the covariance, N represents the number of samples, and x_i and y_i represent the i^{th} samples of x and y , respectively. When the residual sum of squares is minimized,

$$\frac{\partial}{\partial \gamma} \sum_i (\log x_i - \gamma \log y_i)^2 = 0$$

should hold. This is indeed satisfied by Equation (A2).

Similar considerations are valid for the estimation of exponents in multivariate scaling. Assuming Equation (8) and defining $\tilde{s} \equiv s/k^\alpha l^\beta$, the probability function of \tilde{s} is constant regardless of k and l : $P(\tilde{s}|k, l) = \tilde{P}_{s|k,l}(\tilde{s})$. Therefore, α and β can be estimated as follows:

$$(\hat{\alpha}, \hat{\beta}) = \arg \min_{(\alpha, \beta)} \left[\left(\text{Cor} \left[\log \left(\frac{s}{k^\alpha l^\beta} \right), \log k \right] \right)^2 + \left(\text{Cor} \left[\log \left(\frac{s}{k^\alpha l^\beta} \right), \log l \right] \right)^2 \right]. \quad (\text{A3})$$

When the correlation coefficients are zero so that the right-hand side of Equation (A3) is minimal and when we apply transformations $s \leftarrow s/\bar{s}$, $k \leftarrow k/\bar{k}$, and $l \leftarrow l/\bar{l}$, where \bar{x} represents the geometric mean of the variable x ,

$$\sum_i (\log s_i - \alpha \log k_i - \beta \log l_i) \times \log k_i = \sum_i (\log s_i - \alpha \log k_i - \beta \log l_i) \times \log l_i = 0,$$

where k_i , l_i , and s_i represent the i^{th} samples of k , l , and s , respectively. Under this condition,

$$\frac{\partial}{\partial \alpha} \sum_i (\log s_i - \alpha \log k_i - \beta \log l_i)^2 = 0; \quad \frac{\partial}{\partial \beta} \sum_i (\log s_i - \alpha \log k_i - \beta \log l_i)^2 = 0. \quad (\text{A4})$$

Therefore, the estimation in Equation (A3) can be equated to the linear least-squares regression of $\log s$ against $\log k$ and $\log l$ without the interaction term.

Although it is more typical to remove the correlations between “explanatory” variables in the regression analysis, the foregoing method provides a result equivalent to such an orthogonalized version of regression. Let us consider the regression of $\log s$ against $\log k$ and $\log[l/k^{\gamma_1}]$, where γ_1 is a constant determined empirically ($\gamma_1 \approx 1.0$, as shown in Figure 10c). The variables are again normalized with transformation $x \leftarrow x/\bar{x}$, where \bar{x} represents the geometric mean of the variable x . Here, exponents α' and β' are intended to fulfill the multivariate scaling, $s \propto k^{\alpha'} \left(\frac{l}{k^{\gamma_1}} \right)^{\beta'}$, and are thus estimated as follows:

$$(\hat{\alpha}', \hat{\beta}') = \arg \min_{\alpha', \beta'} \left[\text{Cor} \left[\log \left(\frac{s}{k^{\alpha'} (l/k^{\gamma_1})^{\beta'}} \right), \log k \right] \right]^2 + \text{Cor} \left[\log \left(\frac{s}{k^{\alpha'} (l/k^{\gamma_1})^{\beta'}} \right), \log(l/k^{\gamma_1}) \right] \right]^2.$$

When the correlations are zero,

$$\sum_i (\log s_i - \alpha' \log k_i - \beta' (\log l_i - \gamma_1 \log k_i)) \cdot \log k_i = 0; \quad (\text{A5})$$

$$\sum_i (\log s_i - \alpha' \log k_i - \beta' (\log l_i - \gamma_1 \log k_i)) \cdot (\log l_i - \gamma_1 \log k_i) = 0. \quad (\text{A6})$$

By adding Equation (A6) to Equation (A5) multiplied by γ_1 , we obtain

$$\sum_i (\log s_i - \alpha' \log k_i - \beta' (\log l_i - \gamma_1 \log k_i)) \cdot \log l_i = 0. \quad (\text{A7})$$

Equations (A5) and (A7) are satisfied with $\alpha' = \alpha + \beta\gamma_1$ and $\beta' = \beta$ when Equation (A4) holds. Therefore, the estimated values of the exponents in a formal regression analysis can be derived from the regression with the “explanatory” variables (not orthogonalized). Note that α' is the mathematically expected value of γ_2 , as indicated by Equation (A16) in Appendix B.

Appendix B. Derivation of Asymptotic Power-Law or Scaling Exponents

Here, we derive the relations between asymptotic exponents for power-law distributions and scalings assuming the bi- and trivariate scaling relations. This allows us

to check the consistency of results of regression analyses. First, given Equations (1)–(3), we derive the asymptotic power-law exponent of the marginal distribution, $P(x)$, of the “explained” variable, x . Second, given Equations (1) and (8), we calculate the asymptotic power-law exponent of the marginal distribution of s . Third, we show that the bivariate scaling exponents, γ_2 and γ_3 (i.e., those of sales s against number of trading partners k and against number of employees ℓ , respectively), can be determined from α , β , and γ_1 .

Because Equations (1)–(3) have the general form:

$$P(x|y) = \tilde{P}_{x|y}(x/y^\gamma)/y^\gamma \quad (\text{A8})$$

we can derive the power-law exponent at the tail of the density function of the marginal distribution, $P_x(x) \equiv P(x)$, for variable x given the following: positive scaling exponent γ , $\tilde{P}_{x|y}$ (the universal distribution function of x conditional on y), and marginal distribution $P_y(y) \equiv P(y)$. Assuming that $\tilde{P}_{x|y}$ and P_y have power-law right tails of exponents $\delta_{x|y}$ and δ_y , respectively, we approximate the functions as

$$\tilde{P}_{x|y}(\tilde{x}) = \begin{cases} c_{x|y}\tilde{x}^{-\delta_{x|y}} & (\tilde{x} \geq 1) \\ d_{x|y}(\tilde{x}) & (0 \leq \tilde{x} < 1) \end{cases} \quad (\text{A9})$$

and

$$P_y(y) = \begin{cases} c_y y^{-\delta_y} & (y \geq 1) \\ d_y(y) & (0 \leq y < 1) \end{cases} \quad (\text{A10})$$

where $c_{x|y}$ and c_y are positive constants, $\delta_{x|y}$ and δ_y represent the power-law exponents larger than 1, and $d_{x|y}$ and d_y represent the distribution functions for the smaller side. This approximation is based on the actual distributions of the size variables (Figures 2a–d, 4d–f and 5d–f). We set the threshold beyond which the distribution follows the power law to unity because the same results are easily derived for the asymptotic value of exponents with linear transformations as long as the threshold is positive. By applying Bayes’ theorem and assuming that $x > 1$ (because we are interested in the “tails” or asymptotic behaviors as $x \rightarrow \infty$), we derive $P_x(x)$, i.e., the marginal distribution of x , from $P(x, y)$, i.e., the joint distribution of x and y :

$$P_x(x) = \int_0^\infty P(x, y) dy = \int_0^\infty P(x|y)P(y) dy = (I_1 - C)x^{-\delta_{x|y}} + (I_2 + C)x^{-1-(\delta_y-1)/\gamma},$$

where I_1 , I_2 , and C are constants respectively defined as

$$I_1 = c_{x|y} \int_0^1 y^{\gamma(\delta_{x|y}-1)} d_y(y) dy,$$

$$I_2 = c_y \int_1^\infty \tilde{y}^{-\gamma-\delta_y} d_{x|y}\left(\frac{1}{\tilde{y}^\gamma}\right) d\tilde{y},$$

$$C = c_{x|y}c_y / \left(\gamma(\delta_{x|y}-1) - \delta_y + 1\right)$$

and substitution $\tilde{y} = y/x^{1/\gamma}$ is applied in the last integral term. Therefore, the asymptotic behavior of $P_x(x)$ at positive infinity is determined by whether $\delta_{x|y}$ or $1 + (\delta_y - 1)/\gamma$ is smaller: as $x \rightarrow \infty$, the behavior of P_x is approximated as

$$P_x(x) \propto x^{-\delta_x} \quad (\text{A11})$$

where

$$\delta_x = \min\left[\delta_{x|y}, 1 + \frac{\delta_y - 1}{\gamma}\right] \quad (\text{A12})$$

The change in $\delta_{x|y}$ does not affect the power-law exponent as long as δ_y is sufficiently large; this is consistent with the intuitive explanation (Figure 7) presented in the main text.

Although rather complex, a similar strategy works for the case of multivariate scaling. Let us continue with the foregoing notations, except that x , y , and γ are replaced with ℓ , k , and γ_1 , respectively, and Equation (8) is assumed as well as the following:

$$\tilde{P}_{s|k,l}(\tilde{s}) = \begin{cases} c_{s|k,l} \tilde{s}^{-\delta_{s|k,l}} & (\tilde{s} \geq 1) \\ d_{s|k,l}(\tilde{s}) & (0 \leq \tilde{s} < 1) \end{cases} \quad (A13)$$

where $c_{s|k,l}$ is a positive constant, $\delta_{s|k,l}$ represents a power-law exponent larger than 1, and $d_{s|k,l}$ represents the distribution function for the smaller side. Again, the threshold of the functional change of $\tilde{P}_{s|k,l}$ is set to unity without loss of generality. The marginal distribution of s , i.e., $P_s(s)$, for $s > 1$ is

$$\begin{aligned} P_s(s) &= \int_0^\infty \int_0^\infty P(k, l, s) dl dk = \int_0^\infty \int_0^\infty P(s|k, l) P(k, l) dl dk \\ &= \int_0^\infty \int_0^\infty P(s|k, l) P(l|k) P(k) dl dk. \end{aligned}$$

This can be evaluated as the sum of the following 10 integral terms:

$$\int_0^1 \int_0^{k^{\gamma_1}} P(s|k, l) P(l|k) P(k) dl dk = \left[c_{s|k,l} \int_0^1 k^{(\alpha+\beta\gamma_1)(\delta_{s|k,l}-1)} d_k(k) \int_0^1 \tilde{l}^{\beta(\delta_{s|k,l}-1)} d_{l|k}(\tilde{l}) d\tilde{l} dk \right] s^{-\delta_{s|k,l}},$$

where substitution $\tilde{l} = l/k^{\gamma_1}$ is applied;

$$\int_0^1 \int_{k^{\gamma_1}}^{\left(\frac{s}{k}\right)^{\frac{1}{\beta}}} P(s|k, l) P(l|k) P(k) dl dk = \frac{c_{l|k} c_{s|k,l}}{E_1} \left(\left[\int_0^1 k^{(\alpha/\beta+\gamma_1)(\delta_{l|k}-1)} d_k(k) dk \right] s^{-1-(\delta_{l|k}-1)/\beta} - \left[\int_0^1 k^{(\alpha+\beta\gamma_1)(\delta_{s|k,l}-1)} d_k(k) dk \right] s^{-\delta_{s|k,l}} \right)$$

where $E_1 = \beta(\delta_{s|k,l} - 1) - \delta_{l|k} + 1$;

$$\int_0^1 \int_{\left(\frac{s}{k}\right)^{\frac{1}{\beta}}}^\infty P(s|k, l) P(l|k) P(k) dl dk = \left[c_{l|k} \int_0^1 k^{(\alpha/\beta+\gamma_1)(\delta_{l|k}-1)} d_k(k) dk \int_1^\infty d_{s|k,l} \left(\frac{1}{\tilde{l}^\beta} \right) \tilde{l}^{-\delta_{l|k}-\beta} d\tilde{l} \right] s^{-1-(\delta_{l|k}-1)/\beta},$$

where substitution $\tilde{l} = (k^{\alpha/\beta}/s^{1/\beta}) \times l$ is applied;

$$\int_1^{s^{1/(\alpha+\beta\gamma_1)}} \int_0^{k^{\gamma_1}} P(s|k, l) P(l|k) P(k) dl dk = \frac{c_k c_{s|k,l}}{E_2} \left[\int_0^1 \tilde{l}^{\beta(\delta_{s|k,l}-1)} d_{l|k}(\tilde{l}) d\tilde{l} \right] \left(s^{-1-(\delta_k-1)/(\alpha+\beta\gamma_1)} - s^{-\delta_{s|k,l}} \right),$$

where substitution $\tilde{l} = l/k^{\gamma_1}$ is applied and $E_2 = (\alpha + \beta\gamma_1)(\delta_{s|k,l} - 1) - \delta_k + 1$;

$$\int_1^{s^{\frac{1}{\alpha+\beta\gamma_1}}} \int_{k^{\gamma_1}}^{\left(\frac{s}{k}\right)^{\frac{1}{\beta}}} P(s|k, l) P(l|k) P(k) dl dk = \frac{c_k c_{l|k} c_{s|k,l}}{E_1} \left(\left(\frac{1}{E_3} - \frac{1}{E_2} \right) s^{-1-(\delta_k-1)/(\alpha+\beta\gamma_1)} - \frac{1}{E_3} s^{-1-(\delta_{l|k}-1)/\beta} + \frac{1}{E_2} s^{-\delta_{s|k,l}} \right)$$

where $E_3 = (\alpha/\beta + \gamma_1)(\delta_{l|k} - 1) - \delta_k + 1$;

$$\int_1^{s^{1/(\alpha+\beta\gamma_1)}} \int_{(s/k^\alpha)^{1/\beta}}^\infty P(s|k, l) P(l|k) P(k) dl dk = \left[\frac{c_k c_{l|k}}{E_3} \int_1^\infty d_{s|k,l} \left(\frac{1}{\tilde{l}^\beta} \right) \tilde{l}^{-\delta_{l|k}-\beta} d\tilde{l} \right] \left(s^{-1-(\delta_k-1)/(\alpha+\beta\gamma_1)} - s^{-1-(\delta_{l|k}-1)/\beta} \right),$$

where substitution $\tilde{l} = (k^{\alpha/\beta}/s^{1/\beta}) \times l$ is applied;

$$\int_0^{s^{\gamma_1/(\alpha+\beta\gamma_1)}} \int_{s^{1/(\alpha+\beta\gamma_1)}}^{(s/l^\beta)^{1/\alpha}} P(s|k, l) P(l|k) P(k) dk dl = \left(c_k c_{s|k,l} I_1^3 \right) s^{-1-(\delta_k-1)/(\alpha+\beta\gamma_1)},$$

where

$$I_1^3 = \int_0^1 \tilde{l}^{-1+(\delta_{s|k,l}-1)(\alpha+\beta\gamma_1)/\gamma_1-(\delta_k-1)/\gamma_1} \int_{\tilde{l}^{-1/\gamma_1}}^{\tilde{l}^{-(\alpha+\beta\gamma_1)/\alpha\gamma_1}} \tilde{k}^{\alpha(\delta_{s|k,l}-1)-\gamma_1-\delta_k} d_{l|k} \left(\frac{1}{\tilde{k}^{\gamma_1}} \right) d\tilde{k} d\tilde{l}$$

and substitutions $\tilde{k} = l^{-1/\gamma_1} \times k$ and $\tilde{l} = s^{-\gamma_1/(\alpha+\beta\gamma_1)} \times l$ are applied;

$$\int_0^{s^{\gamma_1/(\alpha+\beta\gamma_1)}} \int_{(s/l^\beta)^{1/\alpha}}^{\infty} P(s|k, l)P(l|k)P(k)dkdl = \left(c_k I_2^3\right) s^{-1-(\delta_k-1)/(\alpha+\beta\gamma_1)},$$

where

$$I_2^3 = \int_0^1 \tilde{l}^{(\gamma_1+\delta_k-1)\beta/\alpha} \int_1^{\infty} \frac{1}{\tilde{k}^\alpha} d_{s|k,l} \left(\frac{1}{\tilde{k}^\alpha} \right) \frac{1}{\tilde{k}^{\gamma_1}} d_{s|k,l} \left(\frac{\tilde{l}^{(\alpha+\beta\gamma_1)/\alpha}}{\tilde{k}^{\gamma_1}} \right) \tilde{k}^{-\delta_k} d\tilde{k}d\tilde{l}$$

and substitutions $\tilde{k} = (l^\beta/s)^{1/\alpha} \times k$ and $\tilde{l} = s^{-\gamma_1/(\alpha+\beta\gamma_1)} \times l$ are applied;

$$\int_{s^{\gamma_1/(\alpha+\beta\gamma_1)}}^{\infty} \int_{l^{1/\gamma_1}}^{\infty} P(s|k, l)P(l|k)P(k)dkdl = \left(c_k I_3^3\right) s^{-1-(\delta_k-1)/(\alpha+\beta\gamma_1)},$$

where

$$I_3^3 = \int_1^{\infty} \tilde{l}^{-1-(\alpha+\beta\gamma_1+\delta_k-1)/\gamma_1} \int_1^{\infty} \frac{1}{\tilde{k}^\alpha} d_{s|k,l} \left(\frac{1}{\tilde{k}^\alpha \tilde{l}^{(\alpha+\beta\gamma_1)/\gamma_1}} \right) \frac{1}{\tilde{k}^{\gamma_1}} d_{s|k,l} \left(\frac{1}{\tilde{k}^{\gamma_1}} \right) \tilde{k}^{-\delta_k} d\tilde{k}d\tilde{l}$$

and substitutions $\tilde{k} = l^{-1/\gamma_1} \times k$ and $\tilde{l} = s^{-\gamma_1/(\alpha+\beta\gamma_1)} \times l$ are applied;

$$\int_{s^{\gamma_1/(\alpha+\beta\gamma_1)}}^{\infty} \int_{s^{1/(\alpha+\beta\gamma_1)}}^{l^{1/\gamma_1}} P(s|k, l)P(l|k)P(k)dkdl = \left(c_k c_{l|k} I_4^3\right) s^{-1-(\delta_k-1)/(\alpha+\beta\gamma_1)},$$

where

$$I_4^3 = \int_1^{\infty} \tilde{l}^{-1-(\alpha+\beta\gamma_1+\delta_k-1)/\gamma_1} \int_{\tilde{l}^{-1/\gamma_1}}^1 \frac{1}{\tilde{k}^\alpha} d_{s|k,l} \left(\frac{1}{\tilde{k}^\alpha \tilde{l}^{(\alpha+\beta\gamma_1)/\gamma_1}} \right) \tilde{k}^{\gamma_1(\delta_{l|k}-1)-\delta_k} d\tilde{k}d\tilde{l}$$

and substitutions $\tilde{k} = l^{-1/\gamma_1} \times k$ and $\tilde{l} = s^{-\gamma_1/(\alpha+\beta\gamma_1)} \times l$ are applied. The sum of these 10 terms has the form

$$P_s(s) = A_1 \times s^{-\delta_{s|k,l}} + A_2 \times s^{-1-\frac{\delta_{l|k}-1}{\beta}} + A_3 \times s^{-1-\frac{\delta_k-1}{\alpha+\beta\gamma_1}},$$

where A_1 , A_2 , and A_3 are constants independent of s . Therefore, the asymptotic behavior of P_s as $s \rightarrow \infty$ is expressed as follows:

$$P_s(s) \propto s^{-\delta_s}, \quad (A14)$$

where δ_s is the smallest of the three exponents:

$$\delta_s = \min \left[\delta_{s|k,l}, 1 + \frac{\delta_{l|k}-1}{\beta}, 1 + \frac{\delta_k-1}{\alpha+\beta\gamma_1} \right] \quad (A15)$$

Finally, we evaluate the values of exponents γ_2 and γ_3 using α , β , and γ_1 , given Equations (1) and (7). Then, the joint distribution of k , ℓ , and s is given as follows:

$$P(k, l, s) = P(s|k, l)P(l|k)P(k) = \frac{1}{k^\alpha l^\beta} \tilde{P}_{s|k,l} \left(\frac{s}{k^\alpha l^\beta} \right) \frac{1}{k^{\gamma_1}} \tilde{P}_{l|k} \left(\frac{l}{k^{\gamma_1}} \right) P_k(k).$$

We can obtain the joint distribution of k and s by integrating this probability with respect to ℓ .

$$P(k, s) = \int_{-\infty}^{\infty} P(k, l, s) dl = \int_0^{\infty} \frac{1}{k^\alpha l^\beta} \tilde{P}_{s|k,l} \left(\frac{s}{k^\alpha l^\beta} \right) \frac{1}{k^{\gamma_1}} \tilde{P}_{l|k} \left(\frac{l}{k^{\gamma_1}} \right) P_k(k) dl.$$

Comparing $P(k = k_0, s)$ and $P(k = k_1, s)$ we have

$$\begin{aligned}
P(k = k_1, s) &= \int_0^\infty \frac{1}{k_1^{\alpha\beta}} \tilde{P}_{s|k,l} \left(\frac{s}{k_1^{\alpha\beta}} \right) \frac{1}{k_1^{\gamma_1}} \tilde{P}_{l|k} \left(\frac{l}{k_1^{\gamma_1}} \right) P_k(k_1) dl \\
&= \left(\frac{k_0}{k_1} \right)^{\alpha+\beta\gamma_1} \times \frac{P_k(k_1)}{P_k(k_0)} \int_0^\infty \frac{1}{k_0^{\alpha\beta}} \tilde{P}_{s|k,l} \left(\frac{s}{k_0^{\alpha\beta}} \times \left(\frac{k_0}{k_1} \right)^{\alpha+\beta\gamma_1} \right) \frac{1}{k_0^{\gamma_1}} \tilde{P}_{l|k} \left(\frac{\tilde{l}}{k_0^{\gamma_1}} \right) P_k(k_0) d\tilde{l} \\
&= \left(\frac{k_0}{k_1} \right)^{\alpha+\beta\gamma_1} \times \frac{P_k(k_1)}{P_k(k_0)} \times P \left(k = k_0, s = \left(\frac{k_0}{k_1} \right)^{\alpha+\beta\gamma_1} s \right)
\end{aligned}$$

Therefore,

$$P(s|k = k_1) = \frac{1}{k_1^{\alpha+\beta\gamma_1}} \tilde{P}_{s|k} \left(\frac{s}{k_1^{\alpha+\beta\gamma_1}} \right),$$

where

$$\tilde{P}_{s|k}(\tilde{s}) = \frac{k_0^{\alpha+\beta\gamma_1}}{P_k(k_0)} \times P \left(k = k_0, s = k_0^{\alpha+\beta\gamma_1} \times \tilde{s} \right).$$

Comparing the foregoing result with Equation (2), we have

$$\gamma_2 = \alpha + \beta\gamma_1 \quad (\text{A16})$$

Similarly, the joint distribution of ℓ and s can be evaluated as

$$P(l = l_0, s) = \int_{-\infty}^\infty P(k, l = l_0, s) dk = \int_0^\infty \frac{1}{k^\alpha l_0^\beta} \tilde{P}_{s|k,l} \left(\frac{s}{k^\alpha l_0^\beta} \right) \frac{1}{k^{\gamma_1}} \tilde{P}_{l|k} \left(\frac{l}{k^{\gamma_1}} \right) P_k(k) dk.$$

If we assume a power-law distribution of k similar to Equation (A10), i.e.,

$$P_k(k) = \begin{cases} c_k k^{-\delta_k} & (y \geq \theta_k) \\ d_k(k) & (0 \leq y < \theta_k) \end{cases}$$

where θ_k is a positive constant of the threshold, then by comparing $P(l = l_0, s)$ with $P(l = l_1, s)$ we obtain

$$\begin{aligned}
P(l = l_1, s) &= \left(\frac{l_0}{l_1} \right)^{(\gamma_1-1+\alpha+\beta\gamma_1)/\gamma_1} \left[\left(\frac{l_0}{l_1} \right)^{\delta_k/\gamma_1} P \left(l = l_0, \left(\frac{l_0}{l_1} \right)^{(\alpha+\beta\gamma_1)/\gamma_1} s \right) \right. \\
&\quad \left. - \int_0^{\theta_k} \frac{1}{k^\alpha l_0^\beta} \tilde{P}_{s|k,l} \left(\frac{s}{k^\alpha l_0^\beta} \times \left(\frac{l_0}{l_1} \right)^{(\alpha+\beta\gamma_1)/\gamma_1} \right) \frac{1}{k^{\gamma_1}} \tilde{P}_{l|k} \left(\frac{l_0}{k^{\gamma_1}} \right) P_k(\tilde{k}) d\tilde{k} \right] \\
&\quad + \int_0^{\theta_k(l_1/l_0)^{1/\gamma_1}} \frac{1}{k^\alpha l_1^\beta} \tilde{P}_{s|k,l} \left(\frac{s}{k^\alpha l_1^\beta} \right) \frac{1}{k^{\gamma_1}} \tilde{P}_{l|k} \left(\frac{l_1}{k^{\gamma_1}} \right) P_k(k) dk.
\end{aligned}$$

Here, substitution $\tilde{k} = (l_0/l_1)^{1/\gamma_1} \times k$ is used. Clearly, the two integral terms in this strict relation becomes negligible as $s \rightarrow \infty$. Under an additional condition that Equation (A11) holds for l with $-\delta_l = -1 - (\delta_k - 1)/\gamma_1$

$$P(s|l = l_1) = \frac{P(l = l_1, s)}{P(l = l_1)} \sim \frac{1}{l_1^{(\alpha+\beta\gamma_1)/\gamma_1}} \tilde{P}_{s|l} \left(\frac{s}{l_1^{(\alpha+\beta\gamma_1)/\gamma_1}} \right)$$

as $l_1 \rightarrow \infty$ and $s/l_1^{(\alpha+\beta\gamma_1)/\gamma_1} \rightarrow \infty$ where

$$\tilde{P}_{s|l}(\tilde{s}) = l_0^{(\alpha+\beta\gamma_1)/\gamma_1} \times P \left(l = l_0, s = l_0^{(\alpha+\beta\gamma_1)/\gamma_1} \times \tilde{s} \right).$$

Comparing this result with Equation (3), we have

$$\gamma_3 = \frac{\alpha + \beta\gamma_1}{\gamma_1}. \quad (\text{A17})$$

for the right tail of s in a limited condition.

References

1. Rajan, R.G.; Zingales, L. What Do We Know about Capital Structure? Some Evidence from International Data. *J. Financ.* **1995**, *50*, 1421–1460. [\[CrossRef\]](#)
2. Vijh, A.M.; Yang, K. Are Small Firms Less Vulnerable to Overpriced Stock Offers? *J. Financ. Econ.* **2013**, *110*, 61–86. [\[CrossRef\]](#)
3. Baker, G.P.; Hall, B.J. CEO Incentives and Firm Size. *J. Labor Econ.* **2004**, *22*, 767–798. [\[CrossRef\]](#)
4. Dekker, I.; Barling, J.; Kelloway, E.K. Work Force Size and Multifaceted Job Satisfaction: A Cross-National Study. *J. Soc. Psychol.* **1996**, *136*, 201–208. [\[CrossRef\]](#)
5. Oi, W.Y.; Idson, T.L. Firm Size and Wages. In *Handbook of Labor Economics*; Ashenfelter, O.C., Card, D., Eds.; North-Holland: Amsterdam, The Netherlands, 1999; pp. 2165–2214. [\[CrossRef\]](#)
6. Kumar, K.; Rajan, R.; Zingales, L. *What Determines Firm Size?* National Bureau of Economic Research: Cambridge, MA, USA, 1999. [\[CrossRef\]](#)
7. Curry, B.; George, K.D. Industrial Concentration: A Survey. *J. Ind. Econ.* **1983**, *31*, 203–255. [\[CrossRef\]](#)
8. Davies, S.W.; Lyons, B.R. Seller Concentration: The Technological Explanation and Demand Uncertainty. *Econ. J.* **1982**, *92*, 903–919. [\[CrossRef\]](#)
9. Ross, N.S. Management and the Size of the Firm. *Rev. Econ. Stud.* **1952**, *19*, 148–154. [\[CrossRef\]](#)
10. Gibrat, R. *Les Inégalités Économiques*; Recueil Sirey: Paris, France, 1931.
11. Quandt, R.E. On the Size Distribution of Firms. *Am. Econ. Rev.* **1966**, *56*, 416–432.
12. Okuyama, K.; Takayasu, M.; Takayasu, H. Zipf's Law in Income Distribution of Companies. *Phys. A Stat. Mech. Its Appl.* **1999**, *269*, 125–131. [\[CrossRef\]](#)
13. Cirillo, P.; Hüslér, J. On the Upper Tail of Italian Firms' Size Distribution. *Phys. A Stat. Mech. Its Appl.* **2009**, *388*, 1546–1554. [\[CrossRef\]](#)
14. Ogwang, T. Power Laws in Top Wealth Distributions: Evidence from Canada. *Empir. Econ.* **2011**, *41*, 473–486. [\[CrossRef\]](#)
15. Fix, B. Energy and Institution Size. *PLoS ONE* **2017**, *12*, e0171823. [\[CrossRef\]](#)
16. Ijiri, Y.; Simon, H.A. *Skew Distributions and the Sizes of Business Firms*; North-Holland: Amsterdam, The Netherlands, 1977.
17. Sutton, J. Gibrat's Legacy. *J. Econ. Lit.* **1997**, *35*, 40–59.
18. Coad, A. *The Growth of Firms: A Survey of Theories and Empirical Evidence*; Edward Elgar: Cheltenham, UK, 2009.
19. Dang, C.; (Frank) Li, Z.; Yang, C. Measuring Firm Size in Empirical Corporate Finance. *J. Bank. Financ.* **2018**, *86*, 159–176. [\[CrossRef\]](#)
20. Kimberly, J.R. Organizational Size and the Structuralist Perspective: A Review, Critique, and Proposal. *Adm. Sci. Q.* **1976**, *21*, 571–597. [\[CrossRef\]](#)
21. Bujaki, M.L.; Richardson, A.J. A Citation Trail Review of the Uses of Firm Size in Accounting Research. *J. Account. Lit.* **1997**, *16*, 1–27.
22. Smyth, D.J.; Boyes, W.J.; Peseau, D.E. The Measurement of Firm Size: Theory and Evidence for the United States and the United Kingdom. *Rev. Econ. Stat.* **1975**, *57*, 111–114. [\[CrossRef\]](#)
23. Watanabe, H.; Takayasu, H.; Takayasu, M. Relations between Allometric Scalings and Fluctuations in Complex Systems: The Case of Japanese Firms. *Phys. A Stat. Mech. Its Appl.* **2013**, *392*, 741–756. [\[CrossRef\]](#)
24. West, G.B. *Toward a Science of Companies*; Penguin: New York, NY, USA, 2017.
25. Kleiber, M. Body Size and Metabolic Rate. *Physiol. Rev.* **1947**, *27*, 511–541.
26. Stahl, W.R. Organ Weights in Primates and Other Mammals. *Science* **1965**, *150*, 1039–1042. [\[CrossRef\]](#)
27. Schmidt-Nielsen, K. *Scaling: Why Is Animal Size So Important?* Cambridge University Press: Cambridge, UK, 1984.
28. Savage, V.M.; Gillooly, J.F.; Woodruff, W.H.; West, G.B.; Allen, A.P.; Enquist, B.J.; Brown, J.H. The Predominance of Quarter-Power Scaling in Biology. *Funct. Ecol.* **2004**, *18*, 257–282. [\[CrossRef\]](#)
29. Arrhenius, O. Species and Area. *J. Ecol.* **1921**, *9*, 95–99. [\[CrossRef\]](#)
30. Rosenzweig, M.L. *Species Diversity in Space and Time*; Cambridge University Press: Cambridge, UK, 1995.
31. Bettencourt, L.M.A.; Lobo, J.; Helbing, D.; Kuhnert, C.; West, G.B. Growth, Innovation, Scaling, and the Pace of Life in Cities. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7301–7306. [\[CrossRef\]](#)
32. Li, X.; Wang, X.; Zhang, J.; Wu, L. Allometric Scaling, Size Distribution and Pattern Formation of Natural Cities. *Palgrave Commun.* **2015**, *1*, 15017. [\[CrossRef\]](#)
33. West, G.B.; Brown, J.H.; Enquist, B.J. A General Model for the Origin of Allometric Scaling Laws in Biology. *Science* **1997**, *276*, 122–126. [\[CrossRef\]](#)
34. Do Yi, S.; Noh, J.D.; Minnhagen, P.; Song, M.Y.; Chon, T.S.; Kim, B.J. Human Bipedalism and Body-Mass Index. *Sci. Rep.* **2017**, *7*, 1–8. [\[CrossRef\]](#)
35. Stanley, M.H.R.; Amaral, L.A.N.; Buldyrev, S.V.; Havlin, S.; Leschhorn, H.; Maass, P.; Salinger, M.A.; Stanley, H.E. Scaling Behaviour in the Growth of Companies. *Nature* **1996**, *379*, 804–806. [\[CrossRef\]](#)
36. Tsuruta, D. SME Policies as a Barrier to Growth of SMEs. *Small Bus. Econ.* **2020**, *54*, 1067–1106. [\[CrossRef\]](#)
37. Kobayashi, Y.; Takayasu, H.; Havlin, S.; Takayasu, M. Time Evolution of Companies towards a Stable Scaling Curve Obtained from Flow Diagrams in Three-Dimensional Phase Space. *New J. Phys.* **2019**, *21*, 043038. [\[CrossRef\]](#)
38. Klingenberg, C.P. Size, Shape, and Form: Concepts of Allometry in Geometric Morphometrics. *Dev. Genes Evol.* **2016**, *226*, 113–137. [\[CrossRef\]](#)

39. Davison, A.C.; Hinkley, D.V. *Bootstrap Methods and Their Application*; Cambridge University Press: Cambridge, UK, 1997.
40. Cabinet Office Government of Japan. National Accounts for 2015 (2008SNA, Benchmark Year = 2011). Available online: http://www.esri.cao.go.jp/en/sna/data/kakuhou/files/2015/29annual_report_e.html (accessed on 20 October 2017).
41. Cabinet Office Government of Japan. Indexes of Business Conditions. Available online: <http://www.esri.cao.go.jp/en/stat/di/di-e.html> (accessed on 21 October 2017).
42. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*; Springer Series in Statistics; Springer: New York, NY, USA, 2009. [[CrossRef](#)]
43. Viegas, E.; Goto, H.; Takayasu, M.; Takayasu, H.; Jensen, H.J. Assembling Real Networks from Synthetic and Unstructured Subsets: The Corporate Reporting Case. *Sci. Rep.* **2019**, *9*, 11075. [[CrossRef](#)]
44. Alves, L.G.A.; Ribeiro, H.V.; Lenzi, E.K.; Mendes, R.S. Distance to the Scaling Law: A Useful Approach for Unveiling Relationships between Crime and Urban Metrics. *PLoS ONE* **2013**, *8*, e69580. [[CrossRef](#)]
45. Alves, L.G.A.; Ribeiro, H.V.; Lenzi, E.K.; Mendes, R.S. Empirical Analysis on the Connection between Power-Law Distributions and Allometries for Urban Indicators. *Phys. A Stat. Mech. Its Appl.* **2014**, *409*, 175–182. [[CrossRef](#)]
46. Song, C.; Havlin, S.; Makse, H.A. Self-Similarity of Complex Networks. *Nature* **2005**, *433*, 392–395. [[CrossRef](#)]
47. Rozenfeld, H.D.; Song, C.; Makse, H.A. Small-World to Fractal Transition in Complex Networks: A Renormalization Group Approach. *Phys. Rev. Lett.* **2010**, *104*, 025701. [[CrossRef](#)]
48. Fu, D.; Pammolli, F.; Buldyrev, S.V.; Riccaboni, M.; Matia, K.; Yamasaki, K.; Stanley, H.E. The Growth of Business Firms: Theoretical Framework and Empirical Evidence. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 18801–18806. [[CrossRef](#)]
49. Alfarano, S.; Milaković, M.; Irle, A.; Kauschke, J. A Statistical Equilibrium Model of Competitive Firms. *J. Econ. Dyn. Control* **2012**, *36*, 136–149. [[CrossRef](#)]
50. Gallegati, M.; Guilioni, G.; Kichiji, N. Complex Dynamics and Financial Fragility in an Agent Based Model. In Proceedings of the 2003 International Conference on Computational Science and Its Applications (ICCSA'03), Montreal, QC, Canada, 18–21 May 2003; Volume 6, pp. 770–779. [[CrossRef](#)]
51. Takayasu, M.; Watanabe, H.; Takayasu, H. Generalised Central Limit Theorems for Growth Rate Distribution of Complex Systems. *J. Stat. Phys.* **2014**, *155*, 47–71. [[CrossRef](#)]
52. Buldyrev, S.V.; Nunes Amaral, L.A.; Havlin, S.; Leschhorn, H.; Maass, P.; Salinger, M.A.; Eugene Stanley, H.; Stanley, M.H.R. Scaling Behavior in Economics: II. Modeling of Company Growth. *J. Phys. I* **1997**, *7*, 635–650. [[CrossRef](#)]
53. Buldyrev, S.V.; Pammolli, F.; Riccaboni, M.; Yamasaki, K.; Fu, D.F.; Matia, K.; Stanley, H.E. A Generalized Preferential Attachment Model for Business Firms Growth Rates: II. Mathematical Treatment. *Eur. Phys. J. B* **2007**, *57*, 131–138. [[CrossRef](#)]
54. Wright, I. The Social Architecture of Capitalism. *Phys. A Stat. Mech. Its Appl.* **2005**, *346*, 589–620. [[CrossRef](#)]
55. Riccaboni, M.; Pammolli, F.; Buldyrev, S.V.; Ponta, L.; Stanley, H.E. The Size Variance Relationship of Business Firm Growth Rates. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 19595–19600. [[CrossRef](#)]
56. Mondani, H.; Holme, P.; Liljeros, F. Fat-Tailed Fluctuations in the Size of Organizations: The Role of Social Influence. *PLoS ONE* **2014**, *9*, e100527. [[CrossRef](#)]
57. Malcai, O.; Biham, O.; Solomon, S. Power-Law Distributions and Lévy-Stable Intermittent Fluctuations in Stochastic Systems of Many Autocatalytic Elements. *Phys. Rev. E* **1999**, *60 Pt A*, 1299–1303. [[CrossRef](#)]
58. Sutton, J. The Variance of Firm Growth Rates: The “scaling” Puzzle. *Physica A* **2002**, *312*, 577–590. [[CrossRef](#)]
59. Wyart, M.; Bouchaud, J.P. Statistical Models for Company Growth. *Physica A* **2003**, *326*, 241–255. [[CrossRef](#)]
60. Bottazzi, G.; Secchi, A. Why Are Distributions of Firm Growth Rates Tent-Shaped? *Econ. Lett.* **2003**, *80*, 415–420. [[CrossRef](#)]
61. Bottazzi, G.; Secchi, A. A Stochastic Model of Firm Growth. *Physica A* **2003**, *324*, 213–219. [[CrossRef](#)]
62. Batty, M.; Longley, P. *Fractal Cities: A Geometry of Form and Function*; Academic Press: London, UK, 1994.