

# Measuring productivity dispersion: a parametric approach using the Lévy alpha-stable distribution

Jangho Yang <sup>\*1,3</sup>, Torsten Heinrich<sup>2,3,4</sup>, Julian Winkler<sup>3,7</sup>,  
François Lafond<sup>3,5</sup>, Pantelis Koutroumpis<sup>2,3</sup>, and J. Doyne Farmer<sup>2,3,5,6</sup>

<sup>1</sup>*Management Sciences, Faculty of Engineering, University of Waterloo*

<sup>2</sup>*Oxford Martin Programme on Technological and Economic Change, University of Oxford*

<sup>3</sup>*Institute of New Economic Thinking at the Oxford Martin School, University of Oxford*

<sup>4</sup>*Department of Economics and Business Administration, Chemnitz University of Technology*

<sup>5</sup>*Mathematical Institute, University of Oxford*

<sup>6</sup>*Santa Fe Institute*

<sup>7</sup>*Department of Economics, University of Oxford*

## Abstract

It is well-known that value added per worker is extremely heterogeneous among firms, but relatively little has been done to characterize this heterogeneity more precisely. Here we show that the distribution of value-added per worker exhibits heavy tails, a very large support, and consistently features a proportion of negative values, which prevents log transformation. We propose to model the distribution of value added per worker using the four parameter Lévy stable distribution, a natural candidate deriving from the Generalised Central Limit Theorem, and we show that it is a better fit than key alternatives. Fitting a distribution allows us to capture dispersion through the tail exponent and scale parameters separately. We show that these parametric measures of dispersion are at least as useful as interquartile ratios, through case studies on the evolution of dispersion in recent years and the correlation between dispersion and intangible capital intensity.

Keywords: productivity, dispersion, heavy-tails, Lévy alpha stable distribution.

JEL codes: D2, O3, J24, R12

## 1 Introduction

In the last two decades, the availability of micro-data has revealed the tremendous productivity differences between firms, even within sectors at a fairly detailed level (Bartelsman & Doms 2000, Syverson 2011). Measuring productivity dispersion is important in order to understand misallocation (Gopinath et al. 2017), innovation and diffusion (Berlingieri, Blanchenay & Criscuolo 2017, Andrews et al. 2016),

---

\*We would like to acknowledge funding from Baillie Gifford, the Institute for New Economic Thinking at the Oxford Martin School, and the Rebuilding Macroeconomics project, which is funded by the ESRC. We are grateful to Jean-Philippe Bouchaud and José Moran for very useful comments. *Contacts:* j634yang@uwaterloo.ca (Corresponding Author); torsten.heinrich@wiwi.tu-chemnitz.de; julian.winkler@economics.ox.ac.uk; francois.lafond@inet.ox.ac.uk; pantelis.koutroumpis@oxfordmartin.ox.ac.uk; doyne.farmer@inet.ox.ac.uk

business dynamism (Foster et al. 2021), and ultimately the aggregate productivity slowdown (Goldin et al. 2021).

In this paper, we focus on the distribution of value added (VA) per worker at the firm level, a common measure of productivity in the literature<sup>1</sup>. We employ a commercial dataset, *Orbis Europe*. We have access to a comprehensive version, that includes around 23 million European firm-year observations for the period 2006-2017. The distribution of productivity levels, even within industry-country pairs, exhibits heavy tails, and therefore an infinite variance. In practice, this means that measuring the variance in a given sample is not meaningful, as the result is driven by the sample size, rather than reflecting a true moment of the data (see Fig. 1).

A common solution to this problem is to compute the variance of the *logarithm* of productivity (instead of the variance of productivity). For variables that have positive support, such as revenue per worker, this is an acceptable solution, since the distribution of the log values likely exhibits finite variance. However, this solution is problematic for VA per worker, because firm-level datasets often contain a substantial proportion of firms that have *negative* value added - typically of only a few percent, but up to 23% for country-year pairs in our dataset (Table 1). To state the obvious, as far as measuring dispersion is concerned, removing the firms on the left tail is problematic. The prevalence of firms that stay alive despite very poor performance should be a key indicator in studies of misallocation, creative destruction, and business dynamism.

**How can we measure dispersion in a dataset with negative values and heavy tails?** In this paper, we propose a transparent and straightforward approach: first, finding a good parametric model for the distribution of productivity, and second, using the fitted parameters to evaluate dispersion. For instance, if the data was normally distributed, an excellent way to evaluate how dispersion differs across time or sectors would have been to estimate the scale parameter of the normal distribution, that is, the variance. Here, however, we find that the data exhibits heavy tails and substantial asymmetry, so we turn to more complex models. We find that the Lévy alpha-stable distribution gives an excellent fit to the data, and we argue that its parameters - particularly the scale and the tail parameters - provide appropriate metrics of dispersion.

The scale parameter of the Lévy distribution is similar to the standard deviation in a normal distribution. Roughly speaking, it provides information on the span of the support of the distribution where most of the data lies. In contrast, the tail exponent provides information on the prevalence of extremely productive or unproductive firms.

While the Lévy alpha-stable distribution elegantly addresses the issues of heavy tails and negative values, the problem with parametric measures is that the underlying model can be wrong. We address this in two ways. First, we show that the Lévy alpha-stable offers a surprisingly good fit to our productivity data, and that it is better than the main contender, the 5-parameter Subbotin (Asymmetric

---

<sup>1</sup> See for instance Souma et al. (2009a), Aoyama et al. (2015), Andrews et al. (2016), Berlingieri, Blanchenay & Criscuolo (2017), Gu (2019), Ilzetzki & Simonelli (2017), Campbell et al. (2019), Gouin-Bonenfant (2020) and De Loecker et al. (2022). Baily et al. (1996) note that it is a “conceptually preferable measure of labour productivity” to gross output per worker, even though they have to remove negative VA establishments in some of their analyses. Oulton (2000) discards roughly 1% of his sample of UK firms, which have negative value added. Aradanaz-Badia et al. (2017) note that only a third of the UK firms with negative VA in 2010 had exited the market by 2015, which is only 10 percentage points higher than other firms. De Loecker et al. (2022) find a proportion of negative VA UK firms in both Orbis and administrative data. While issues with VA-based measures of productivity have been noted (Campbell et al. 2019, Cunningham et al. 2021), measures based on gross output are less comparable across different sectors, as intermediate consumption is naturally larger in some industries than in others, so that using gross output-based measures creates a need to use industry means to renormalize the values and compare them; this itself comes with its own issues, if only because classification systems are imperfect, lack granularity, and change over time.

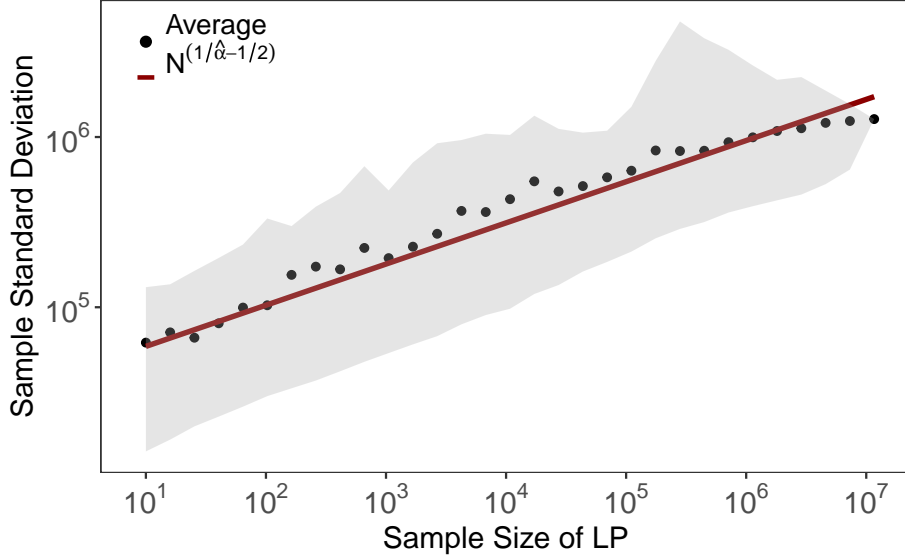


Figure 1: **Measured standard deviation in sub-samples of firm labour productivity.** We construct a distribution by pooling together the productivity levels of firms in France, Italy, Germany, and Spain (which have the same currency), for all years, expressed in thousands. Then, for each subsample size  $N$ , we compute the standard deviation of 100,000 subsamples, and report the average (black dots), and 5<sup>th</sup> and 95<sup>th</sup> percentiles. In Appendix D, we explain that on i.i.d. data with a power law tail, the theoretical scaling between the sample standard deviation and the sample size would be  $N^{\frac{1}{\hat{\alpha}}-1/2}$ , where  $\alpha$  is the tail exponent. The red line shows that this scaling holds empirically here (we use our estimate of the tail parameter  $\hat{\alpha} = 1.33$ , and the intercept is calculated by simulating 100,000 i.i.d. random samples of size 10 from a Lévy alpha-stable with the same parameters as those we estimate on empirical data). Data from Orbis Europe.

Exponential Power, AEP) distribution (Bottazzi & Secchi 2011). We add credibility to the model by using independent tests to confirm that the second moment is infinite. Second, we argue that the Lévy alpha-stable model is plausible *a priori* simply because it is the result of the Generalized Central Limit Theorem (GCLT) (Nolan 2020). If VA per worker can be thought of as the aggregation of several micro-level variables such as the productivity of individual employees, contracts, tasks or routines, then under fairly general conditions the GCLT predicts that firm-level productivity will follow a Lévy alpha-stable distribution.

Having established that the data is well fitted by a Lévy distribution, we can use the estimated parameters to evaluate dispersion. We propose the tail parameter as an indicator of dispersion in the tail, and the ratio of the scale to location parameters (similar to the coefficient of variation ( $\sigma/\mu$ ) for a normal distribution) as an indicator of dispersion in the body of the distribution. We compare these metrics with interquantile ratios, the most widely used non-parametric measures of dispersion, which can be designed to measure dispersion in body ( $Q_{90}/Q_{10}$ ) or in the tail ( $Q_{95}/Q_{50}$ ). While these metrics are all empirically correlated, the correlations are imperfect, so they capture fairly distinct aspects of dispersion.

We then proceed with two case studies: the evolution of dispersion in recent years, and the correlation between dispersion and intangible intensity, measured as intangible capital per worker. We find that the quantile-based and Lévy-based metric give similar results. First, we find that for our period

(2006-2017), there was no overwhelming evidence for a substantial, systematic increase in productivity dispersion in all countries. Second, we find a positive, but fairly noisy correlation between intangible intensity and dispersion. While the results are similar, we show that our parametric measures are superior because they make it possible to keep negative values, and avoid having to select specific quantiles. The results also appear somewhat less noisy with the Lévy metrics than with the quantile-based metrics.

**Literature.** Our results relate to three strands of literature. The first is the large body of literature on heavy tails in economics and finance (Mitchell 1915, Mandelbrot 1960, 1963, Fama 1965, Samuelson 1967, Embrechts et al. 1997, Fagiolo et al. 2008, Gabaix 2011, Acemoglu et al. 2017, Axtell & Guerrero forthcoming), more specifically on the distributions of firm sizes (Ijiri & Simon 1977, Axtell 2001) and firm growth rates (Bottazzi & Secchi 2003, Bottazzi et al. 2007, Bottazzi & Secchi 2011, Schwarzkopf et al. 2010, Holly et al. 2013, Moran et al. 2020). Most related to our study is the work on estimating productivity distributions such as Aoyama et al. (2010) who estimate a Generalized Beta model for labour productivity levels<sup>2</sup>, Souma et al. (2009b) who find a power law tail for labour productivity levels<sup>3</sup>, and Gaffeo (2008, 2011) who estimates a Lévy alpha-stable distribution for the growth rates of sector-level total factor productivity. In the present paper, we focus on *firm-level* data, productivity levels, and productivity measured as *value added per worker*.

Second, our results on productivity levels are relevant to the literature that discusses how productivity dispersion may reflect the misallocation of factors of production (Hsieh & Klenow 2009, Bartelsman et al. 2013, Foster et al. 2021, Gopinath et al. 2017, Haltiwanger et al. 2018). In this paper, we critically discuss measures of dispersion in detail, showing how some statistics may be misleading and what parts of the distribution can affect dispersion. We also suggest that theoretical models should aim to derive a Lévy alpha-stable distribution, where ideally the parameters would be interpreted in terms of misallocation, or other sources<sup>4</sup>.

Third, our results relate to the work of statistical agencies around the world, who are developing procedures to make summary statistics of the micro data publicly available, and in particular industry-level productivity dispersion (Cunningham et al. 2021, Berlingieri, Blanchenay, Calligaris & Criscuolo 2017). A key contribution of this paper is to make a suggestion regarding what statistics should be released, namely 5 quantiles necessary to perform McCulloch (1986)'s quantile estimation of the four Lévy alpha-stable parameters. Typically, readily available software packages use the 5<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 95<sup>th</sup> quantiles, as in McCulloch (1986)'s original paper. We suggest these five quantiles are good candidates, and could be supplemented by more quantiles which can then be used for testing the quality of the fit.

The paper is organized as follows. Section 2 describes the data sources and the basic patterns, including the presence of heavy tails. Section 3 presents the Lévy alpha-stable distribution, the Generalized CLT, the main alternative parametric model, our fitting methods, and shows the quality of the fit. Section 4 then compares the quantile-based and Lévy-based measures of dispersion in practice,

---

<sup>2</sup> Generalized Beta models are defined only over a positive support and thus are not suitable for our data, which contains negative productivity firms.

<sup>3</sup> See also Aoyama et al. (2009), Mizuno et al. (2012) and Aoyama et al. (2015).

<sup>4</sup> There is an important theoretical literature on productivity dispersion, which recognizes that the right tail of the distribution of productivity levels is a power law, and explains this mostly as the result of innovation and imitation processes, see Ghiglino (2012), Lucas & Moll (2014), Perla & Tonetti (2014) and König et al. (2016). The presence of extreme values in empirical data also implies that data cleaning choices may affect substantive research results, as shown for the misallocation literature by Rotemberg & White (2021).

showing the evolution of dispersion over time, and the correlation between industry-level dispersion and intangible capital intensity. Section 5 concludes.

## 2 Data and descriptive statistics

### 2.1 Data sources

In this section, we present the dataset and some basic patterns of the empirical data. We use data from the Orbis Europe database, compiled by Bureau van Dijk, which includes the balance sheets and profit-loss statements for 7 million unique firms across Europe, yielding approximately 23 million firm-year observations from 2006 to 2017.

Unlike other widely used firm-level data such as Compustat and Worldscope, Orbis Europe records a large number of small and medium sized firms (SME) that are often not publicly traded. Table 7 in Appendix A reports the number of observations per country-year.<sup>5</sup>

To avoid double counting, we only use unconsolidated data<sup>6</sup>. We then remove duplicated firm-year pairs, and we exclude self-employed firms, where the distinction between wage and profit is ambiguous. Finally, we regard negative values for total assets, fixed assets, sales, wage, and employment as missing values. For a detailed discussion on data cleaning, see Appendix A.

### 2.2 Construction of variables

For each firm  $i$  in industry  $j$  and year  $t$ , we define real value added ( $Y_{i,t}$ ) as the sum of real labour income ( $W_{i,t}$ ) and real capital income ( $\Pi_{i,t}$ ), that is

$$Y_{i,t} \equiv W_{i,t} + \Pi_{i,t} = \frac{\omega_{i,t}}{p_{j,c,t}^v} + \frac{\pi_{i,t}}{p_{j,c,t}^v}, \quad (1)$$

where  $\omega$  and  $\pi$  are nominal wage and nominal profit, and  $p_{j,c,t}^v$  is the value added deflator of industry  $j$  in country  $c$  at time  $t$ . In Orbis, nominal wage and nominal profit variables are recorded in Cost of Employees (STAF) and Earnings Before Interest, Taxes, Depreciation & Amortization (EBITDA)<sup>7</sup>. The firms' VA is deflated using the industry-level VA deflator from the EUKLEMS database ( $VA\_P$  in Jäger & The Conference Board (2018)). See Appendix A.2 for more details.

<sup>5</sup> Orbis Europe includes many different types of firms, including Private limited company, Joint-stock company, Partnership, Cooperative, Consortium, Foundation, and Public agency/corporation. For a more detailed discussions on the advantages and drawbacks of the Orbis Europe database, see Kalemli-Ozcan et al. (2015).

<sup>6</sup> For some corporate groups that consist of multiple firms, both consolidated and unconsolidated data is provided, i.e. data is listed for the entire group and again for each of the firms.

<sup>7</sup> VA can also be calculated by subtracting intermediate costs (MATE in Orbis) from gross output (OPRE in Orbis). This is the output-based approach to computing VA. However, material costs are less frequently observed compared to wages and earnings in our sample. For example, only 52% of firms in Germany report material costs. In some countries such as the UK and Denmark, no firms report material costs. Therefore, to get the largest possible coverage for VA, we use the income-based value added by adding wages and earnings. We note that the labour productivity variable constructed from the output based VA in our Orbis data tends to overestimate the labour productivity recorded in the national accounts. For example, the output-based productivity in France and Germany is almost twice as high as the labour productivity in the national accounts. Using the output-based labour productivity, only around 1% of firms have negative productivity. Gopinath et al. (2017), who also used Orbis Europe data, construct output-based VA for Spanish firms, but this is justified in their case since around 90% of firms in Spain report material costs, and for Spain the output-based productivity is very close to the productivity recorded in national accounts.

Denoting the number of employees by  $L_{i,t}$  (EMPL in Orbis), our measure of firm-level labour productivity ( $LP_{i,t}$ ) is defined as the ratio of value added to the number of employees,

$$LP_{i,t} = \frac{Y_{i,t}}{L_{i,t}}, \quad (2)$$

and is expressed in units of currency of the country where the firm is located.

Researchers in productivity analysis typically proceed to a data transformation: they analyze the natural logarithm of productivity (Bartelsman & Wolf 2018). However, when output is measured as VA, firm-level labour productivity can be negative. If a firm’s intermediate cost is greater than revenues in a given fiscal year, the firm has negative VA and productivity. From the income perspective of the definition of VA (Wages + Earnings), the sum of earnings and wages can be negative since firms’ losses (negative profits) can be greater than wages.

Empirically, a sizeable share of firms has negative labour productivity. Table 1 shows the proportion (%) of negative observations per country-year in our data. Overall, 5% of the firms in our sample have negative productivity. As a more extreme case, more than a fifth of UK firms had negative productivity in 2006-2008<sup>8</sup>.

Table 1: Proportion (%) of negative observations per country-year.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Belgium	2.58	1.33	1.56	1.62	1.49	1.44	1.55	1.59	1.53	1.49	1.53	1.58
Bulgaria	0.00	9.33	4.02	5.79	5.86	7.45	7.60	7.02	6.35	6.10	5.47	5.13
Croatia	6.91	4.46	4.60	5.66	6.18	5.27	5.79	4.54	4.00	3.55	2.86	0.00
Czech Republic	7.10	4.89	5.32	6.49	6.26	6.35	6.33	6.19	5.48	4.52	4.30	3.69
Denmark	0.00	0.00	0.00	0.00	0.00	0.00	2.51	2.98	2.83	2.51	2.91	3.04
Estonia	0.00	2.90	4.09	6.09	4.65	3.25	3.11	3.12	3.37	3.38	3.30	2.75
Finland	2.09	1.60	1.83	1.99	1.84	2.03	1.91	1.95	1.94	2.08	2.08	1.90
France	1.19	1.10	1.31	1.27	1.13	1.16	1.32	1.46	1.43	1.60	1.81	1.89
Germany	2.56	1.96	2.48	2.43	2.14	2.17	2.16	2.00	2.39	2.48	2.31	0.00
Hungary	0.00	4.11	3.11	7.01	6.61	6.68	7.18	5.76	5.16	4.87	4.47	3.62
Italy	4.64	3.08	3.64	4.16	3.68	3.78	5.00	4.74	4.32	3.84	3.16	2.79
Poland	3.77	4.01	4.78	6.11	3.72	3.95	8.12	7.42	0.00	0.00	0.00	2.49
Portugal	8.62	6.06	6.51	6.12	5.81	6.87	8.34	7.87	7.86	6.47	5.57	4.80
Romania	0.00	7.58	8.52	10.73	11.73	11.46	12.31	12.47	10.40	7.34	6.39	0.00
Slovakia	0.00	4.44	4.55	7.64	7.11	7.23	7.20	7.13	6.05	4.10	4.21	0.00
Slovenia	0.00	0.00	0.00	1.50	1.69	1.81	1.88	1.56	1.44	1.35	1.25	1.15
Spain	3.78	2.92	3.60	3.90	3.65	4.04	4.41	4.05	3.33	2.94	2.57	2.07
Sweden	3.82	2.16	2.41	2.27	2.13	2.18	2.10	2.10	2.07	2.08	2.15	2.35
United Kingdom	20.98	21.72	23.49	11.31	4.14	4.04	3.84	3.52	3.57	3.73	3.92	0.00

This table records the proportion (%) of negative observations of value added per country-year for all 19 countries.

### 2.3 Characteristic patterns of the distributions

To motivate Lévy alpha-stable distributions as a model for  $LP$ , we first show qualitatively that the distributions are heavy-tailed and asymmetric.

<sup>8</sup> Other authors find smaller percentages for the UK, but there is systematically at least a few percents of negative VA firms. This is close to 10% in Haldane (2017, Chart 17). In employment weighted distributions, this is around 4-5% (Office for National Statistics 2022). De Loecker et al. (2022) find only ‘a few’ in Orbis, but remove many sectors and firms with less than 10 employees.

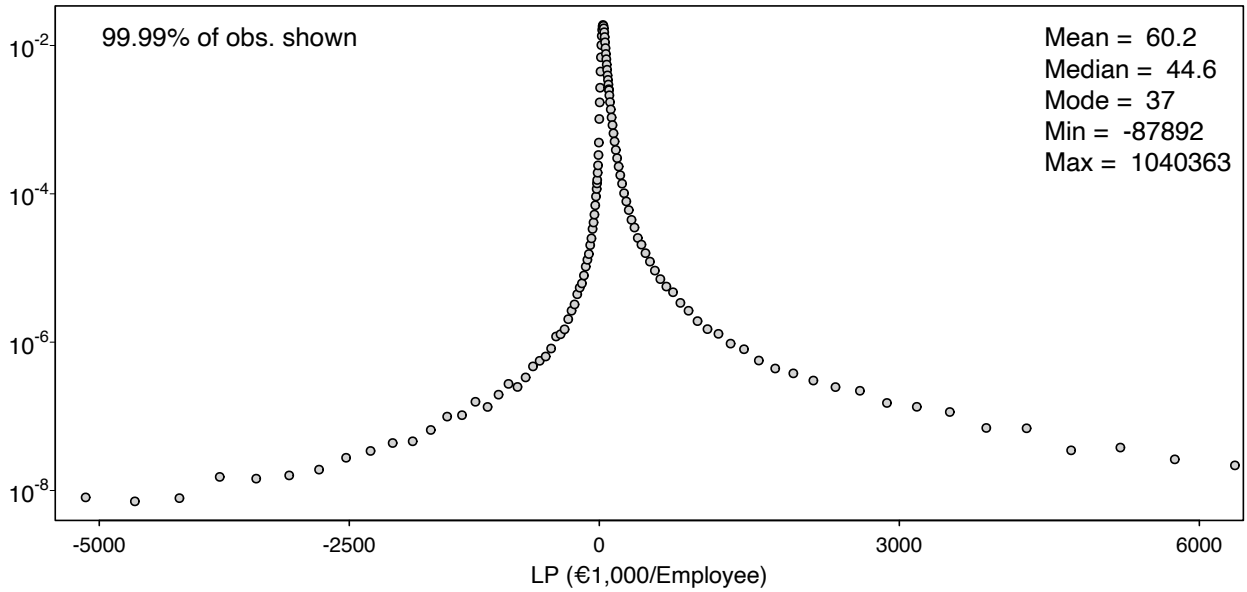


Figure 2: **Distribution of labour productivity in France, 2006-2015.** To display as many data points in the extreme tails as possible, we use log binning for the bottom and top 10% productivity. The plot shows 99.99% of the entire data.

Fig. 2 shows the distribution of labour productivity for France for the years covered by the data on a semi-log scale. An important observation is that even though the chart shows 99.99% of observations, the tails appear relatively “well behaved”. Once the bins mid-points are chosen appropriately, there are no large fluctuations and one could draw a fairly smooth line through the dots. This suggests that the common practice of windsorizing (removing e.g. 1% of observations in each tail) removes observations that are actually well-behaved statistically, rather than being unexpected or strange “outliers” (see e.g. Nolan (2020, p.196-197) for a discussion).

We note the following five general characteristics of the empirical distribution of the labour productivity. First, the distribution appears unimodal. Second, the support of the distribution is very large. Third, the distribution is asymmetric, with a pronounced right skewness. Fourth, the distribution exhibits slowly decaying tails. And fifth, comparing the results for different years (not shown here), we find that the shape of the distribution is very persistent, as one would expect.

We now focus on one of the most important of these features: heavy tails.

## 2.4 Testing for power law tails and infinite variance

We say that the tail of a distribution follows a power-law if its cumulative distribution function  $F(x)$  takes the form

$$1 - F(x) \sim L(x)x^{-\alpha}, \quad x \rightarrow \infty, \quad (3)$$

where  $L(x)$  is a slowly varying function, and  $\alpha$  is the tail exponent. This tail exponent regulates the rate with which the size of extreme values decreases as we increase the sample of extreme values. Intuitively, the smaller the tail exponent, the slower the frequency of a large event decreases as we consider increasingly extreme values. Very extreme values will be relatively for frequent than a scenario with a higher tail exponent. In practice, it also determines what moments are finite; any moment greater

than  $\alpha$  will be infinite, while any moment less than  $\alpha$  remains finite. Some authors use the term ‘heavy’ tails for the case where  $\alpha < 2$ , as in this case the variance is infinite.

Here we propose four methods to determine whether labour productivity distributions have heavy tails: (modified) QQ-plots, scaling of the sample standard deviation with sample size, direct estimates of the tail exponent based on extreme value distributions, and [Trapani \(2016\)](#)’s test for infinite moments.

**QQ-plots** [Resnick \(2007\)](#) shows that QQ (Quantile-Quantile) plots for location-scale families can be used not only to provide visual evidence on the compatibility of a specific functional form with the data, but also to estimate the parameters. This is convenient because if  $X$  is a random variable drawn from a Pareto distribution,

$$P(X > x) = \left(\frac{x}{k}\right)^{-\alpha}, \quad x > k, \quad (4)$$

where  $k$  is the threshold, and  $\alpha$  is the tail exponent, the distribution of  $\log X$  is a location-scale family. [Resnick \(2007\)](#) shows that the values

$$\left\{ \left( -\log \left( 1 - \frac{i}{N+1} \right); \log X_{(i)} \right), 1 \leq i \leq N \right\}, \quad (5)$$

where  $X_{(i)}$  is the  $i^{\text{th}}$  order statistic, would lie on a line with intercept  $\log k$  and slope  $1/\alpha$  under the null (Eq. 4). To implement this procedure, we need to choose a cutoff  $k$ , or, alternatively, the sample size  $N$ . If we believe that the data has a Pareto tail but only as  $x \rightarrow \infty$ , we would prefer to choose the highest possible cutoff. However, this implies less data and thus more variance in the resulting estimates, so there is a trade-off.

[Fig 3](#) shows the log of top and bottom 0.5% observations of firm labour productivities. The data-points by country roughly follow straight lines, which is good evidence that the tails are Pareto. Moreover, the slopes appear to lie between 1 and 1/2, suggesting  $1 < \alpha < 2$ .

One point to note is that the slopes for negative productivity appear slightly steeper than for positive productivity, suggesting a degree of tail asymmetry that is not in line with the Lévy hypothesis, where the tail exponent is the same on each side (Section 3.1.1). Indeed, as shown in [Fig 7](#), our fits may have a perceptible tendency to underestimate the left tail and overestimate the right tail. But overall, the data is substantially right-skewed, and this is well captured by the asymmetry parameter of the Lévy distribution, so the overall fit remains very good.

**Scaling of the sample standard deviation with sample size** If the distribution has power-law tail with  $\alpha < 2$ , the variance and higher-order moments are infinite. Due to this property, in finite samples, the larger the sample size, the higher the chance that an extreme event is drawn, leading to a larger sample standard deviation. More precisely, the sample standard deviation scales with sample size  $N$  as  $N^{\frac{1}{\alpha}-\frac{1}{2}}$  (see [Appendix D](#) for a heuristic derivation).

[Figure 4](#) shows the scaling of the sample standard deviation of labour productivity for different sample sizes. It is clear that for most countries, the sample standard deviation gets larger as the sample size increases. For four selected countries, we show the theoretical scaling derived by estimating the power law exponent as the tail parameter of a Lévy alpha-stable distribution. The scaling of the sample standard deviation holds well empirically, although with a possible plateauing as  $n \rightarrow \infty$  for some countries, such as Spain. Overall, this provides good evidence not only for power law tails, but also



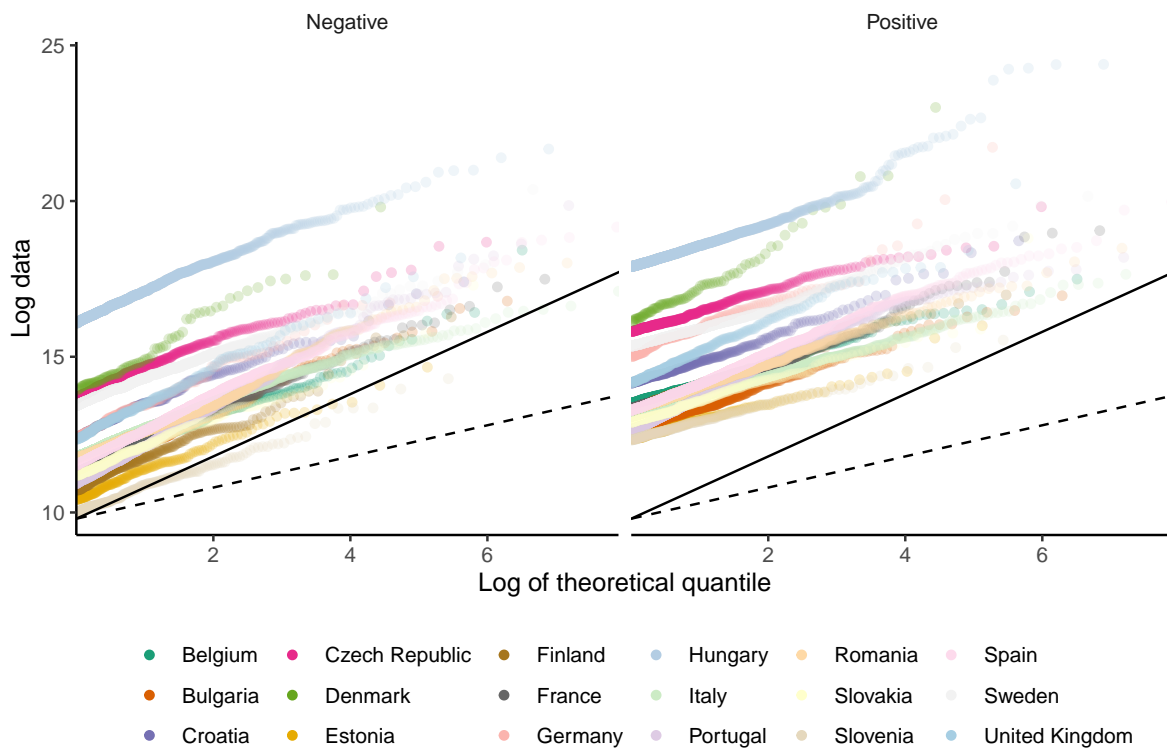


Figure 3: **QQ-plots for positive and negative labour productivity tails, 2015.** Quantiles of the labour productivity for the top and bottom 0.5% of firms in 18 European countries are very close to the quantiles of a Pareto distribution. A Pareto tail with a exponent  $\alpha$  would appear as a line with slope  $a/\alpha$  (Resnick 2007). A line with a slope of one (solid black), and a line with a slope of 1/2 (dashed) are plotted for visual comparison.

for the idea that the Lévy alpha-stable distribution can be a good model to retrieve the value of the tail index.

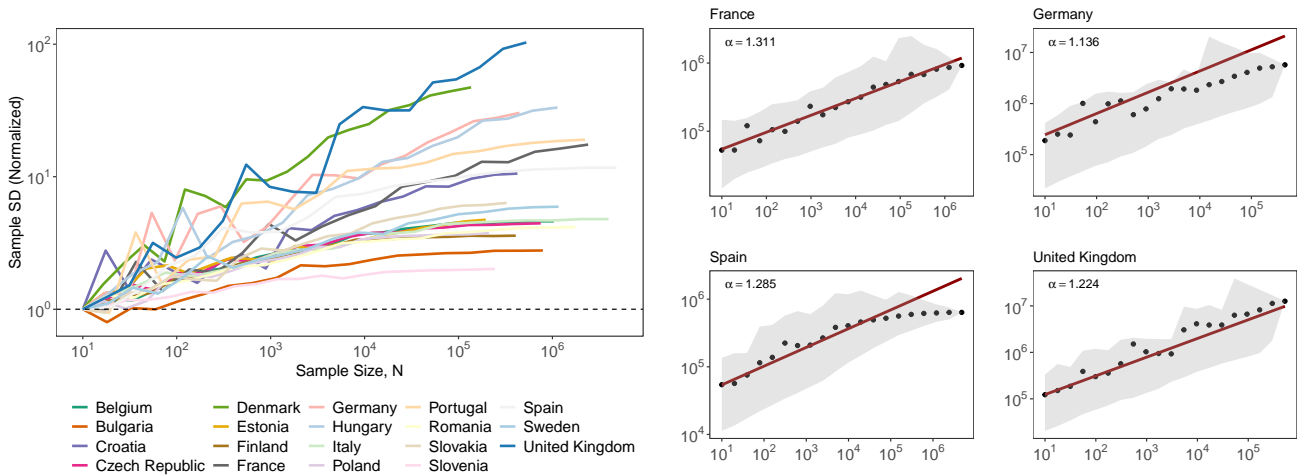


Figure 4: **Measured standard deviation in sub-samples of firm labour productivity, country samples.** We construct a distribution by pooling together the productivity levels of firms in each of 19 countries, for all years. Then, for each subsample size  $N$ , we compute the standard deviation of each of 1,000 subsamples. The plot on the left shows the average of standard deviations for each of all 19 countries, while the plot on the right shows the same scaling for four selected countries, with the linear scaling calculated as  $N^{\frac{1}{\hat{\alpha}} - \frac{1}{2}}$  (see Appendix D), where  $\hat{\alpha}$  is estimated by fitting the Lévy distribution. The shaded area corresponds to the values of the sample standard deviation that fall between the 5<sup>th</sup> and 95<sup>th</sup> percentiles.

**Estimating tail exponents** Another way to test for infinite second moment is to estimate the tail exponent and see if it is lower than 2. Typically, one considers data from the tail only; that is, order statistics of up to order  $k$ , which makes it possible to estimate tail behaviour and determine the finiteness of moments, independently of the behaviour of the rest of the distribution. An important issue is that one has to choose a value  $k^*$  which determines which data is used to estimate the tail parameter. Usually, the tails are influenced in a non-negligible way by the slowly varying function  $L(x)$  (see Eq. 3), and therefore the choice of  $k^*$  may be difficult and lead to biased estimates of  $\alpha$ .

An early and popular method for estimating tail exponents is the Hill estimator, but as noted in Resnick (2007), the Hill estimator provides very different estimates of the tail depending on which value of  $k^*$  we chose (“Hill Horror Plots”). Because of these well-known issues for regularly varying distributions (Eq. 3) that are somewhat far from pure power laws, several estimators of the tail exponents have been developed, and tested for cases where the slowly varying function is non-negligible. Here we use the estimators described and implemented by Voitalov et al. (2019). Voitalov et al. (2019)’s implementation includes an automatic double bootstrapping procedure for picking  $k^*$ . We use their package with default values. In addition, we also use the popular Hill estimator described in Clauset et al. (2009), which finds  $k^*$  by minimizing the Kolmogorov-Smirnov statistic (computed assuming that the true model is a pure power law). We add the estimate of  $\alpha$  based on the Lévy distribution as well.

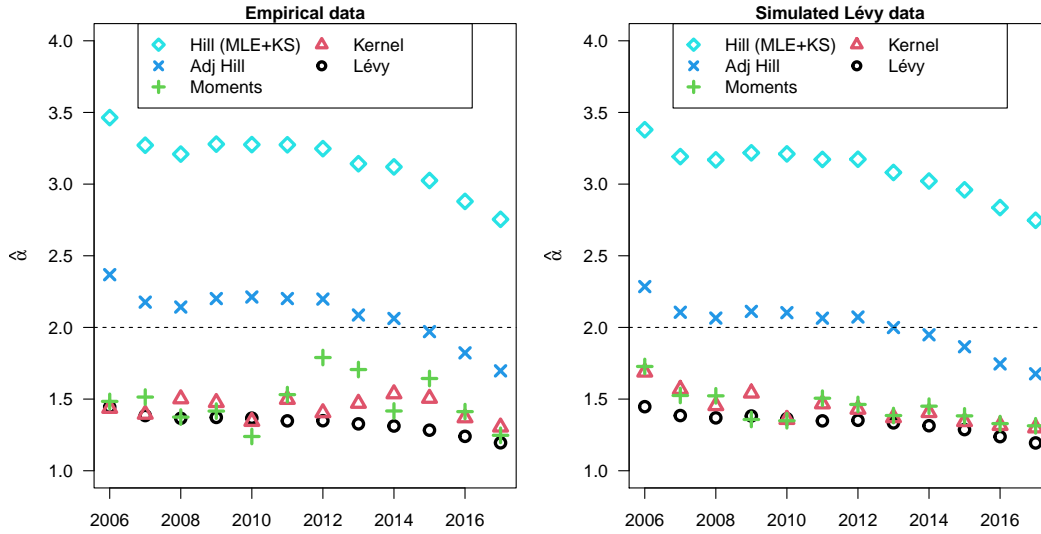


Figure 5: **Estimated values of the tail exponent  $\hat{\alpha}$** , using pooled data from France for each year separately, and using 5 different estimators. The left chart shows  $\hat{\alpha}$  for the empirical data. The right chart shows  $\hat{\alpha}$  for data drawn from a Lévy-stable distribution with parameters equal to those estimated on the empirical data. The left chart shows that 3 estimators suggest infinite variance ( $\hat{\alpha} < 2$ ), while the two Hill-based estimators suggest finite variance on empirical data, but would also fail to capture infinite variance if the data were indeed Lévy-distributed (right). We conclude that labour productivity has indeed a power law tail with  $\alpha < 2$ .

Figure 5 (left) shows the results using data from France for each year separately, and strongly suggests that  $\alpha < 2$ . Three out of 5 estimators suggest similar values of  $\alpha \approx 1.4$  (dots for 2010 are missing for two estimators as the routines sometimes fail). The two estimators that suggest  $\alpha > 2$  are Hill-based estimators, which are known to produce poor results for Lévy distributed data (Resnick 2007). To further confirm the plausibility of these estimates, for each year, we simulate one sample of Lévy distributed data using the empirical sample size, and using parameters for the Lévy distribution estimated from the empirical data. In this case (right panel), the 5 estimators behave almost exactly as in the empirical data, suggesting that the Lévy distribution is a plausible model, and reinforcing the previous conclusion that the second moment is infinite.

The left panel of Figure 5 also helps to make an important point: if the data is *not* Lévy distributed, but is a regularly varying distribution with a tail exponent that is correctly estimated by the Kernel or Moments estimators, then estimating the exponent using simply 5 quantiles and assuming a Lévy distribution gives relatively good results, in the case of our data. This is important, because statistical offices holding detailed micro-data may not want to release the order statistics that are necessary to run the Kernel or Moments estimators, but they may be willing to release 5 quantiles, particularly since the lowest/highest quantiles ( $5^{th}$  and  $95^{th}$ ) are likely to be uninformative about the situation of specific firms (in contrast to, say, the  $99.9^{th}$  quantile).

**Trapani’s test** As a final check, we implemented a more formal statistical test due to Trapani (2016). The testing procedure exploits the fact that non-finite sample moments diverge with sample size. A full overview of the testing procedure and the detailed results are available in Appendix F.

In the vast majority of both the country-year and country-industry samples, we cannot reject the hypothesis that the second moment of labour productivity is infinite. For example, for 84% of country-

year samples, Trapani’s finite moment test failed to reject the null hypothesis of the infinite second moment at the 5% confidence level.

We also perform Trapani’s test for country-year-industry samples at the level of two-digit NACE codes. For 55% of country-industry samples, Trapani’s finite moment test failed to reject the null hypothesis of the infinite second moment at the 5% confidence level. For the largest 5 European countries (France, Germany, Italy, Spain, and the UK), the same result holds for nearly 70% of samples<sup>9</sup>.

In Appendix F, we evaluate the performance of the test, and find that it tends to over-reject the null, sometimes substantially; this is unsurprising since we make a conservative choice for the parameters of the test. As a result, despite somewhat mixed results on the fine-grained sub-samples, we conclude that in a large majority of cases, the second moment of the distribution of VA per worker is likely to be infinite.

### 3 Models and estimation

This section provides a detailed discussion of the key properties of the Lévy alpha-stable distribution, and introduces the estimation methods. For details on the competing distributional models (the Asymmetric Exponential Power (AEP) with 4 or 5 parameters), see Appendix C and Bottazzi & Secchi (2011).

#### 3.1 The Lévy alpha-stable distribution

The Lévy alpha-stable distribution is a natural candidate for many distributions exhibiting heavy tails, as it emerges from Generalized Central Limit Theorem (GCLT). We first describe the key properties of the Lévy alpha-stable distribution, and then discuss the GCLT.

##### 3.1.1 Characteristics of the Lévy alpha-stable distribution

The Lévy alpha-stable distribution is a four-parameter distribution with parameters  $\alpha$  (tail exponent),  $\beta$  (skewness),  $\gamma$  (scale), and  $\delta$  (location). The density function exists in closed form only in a few special cases such as  $\alpha = 2$  (Gaussian),  $\{\alpha = 1, \beta = 0\}$  (Cauchy), and  $\{\alpha = 1/2, \beta = 1\}$  (standard Lévy). Out of several alternatives, we use Nolan’s (1998)  $S_0$  parametrization, which has the characteristic function

$$\varphi(t) \equiv E[\exp(itX)] = \begin{cases} \exp\left(-|\gamma t|^\alpha \left[1 + i\beta \tan\left(\frac{\pi\alpha}{2}\right) \operatorname{sgn}(t) (|\gamma t|^{1-\alpha} - 1)\right] + i\delta t\right) & \alpha \neq 1, \\ \exp\left(-\gamma |t| \left[1 + i\beta \frac{2}{\pi} \operatorname{sgn}(t) \log(\gamma |t|)\right] + i\delta t\right) & \alpha = 1, \end{cases} \quad (6)$$

where  $i$  is the imaginary unit,  $t \in R$  is the argument of the characteristic function, and  $\operatorname{sgn}()$  is the sign function. This particular parametrization is useful for numerical work and statistical inference since the characteristic function is continuous in all four parameters.

---

<sup>9</sup> We also estimated the tail exponent using Voitalov et al. (2019)’s estimators at the country-sector-year level and found similar results. While the majority of cases have a tail exponent lower than 2, there are many cases having an exponent between 2 and 3, and many cases having too few values to give reliable estimates. In Appendix E, we show examples of the scaling (or lack thereof) of the empirical standard deviation with sample size at the country-industry-year level. Of course, part of the dispersion on the pooled distributions is due to pooling heterogenous subsamples, and we acknowledge that we cannot validate the infinite moment hypothesis (or, for that matter, the Lévy hypothesis) on detailed subsamples as clearly as on the pooled distribution.

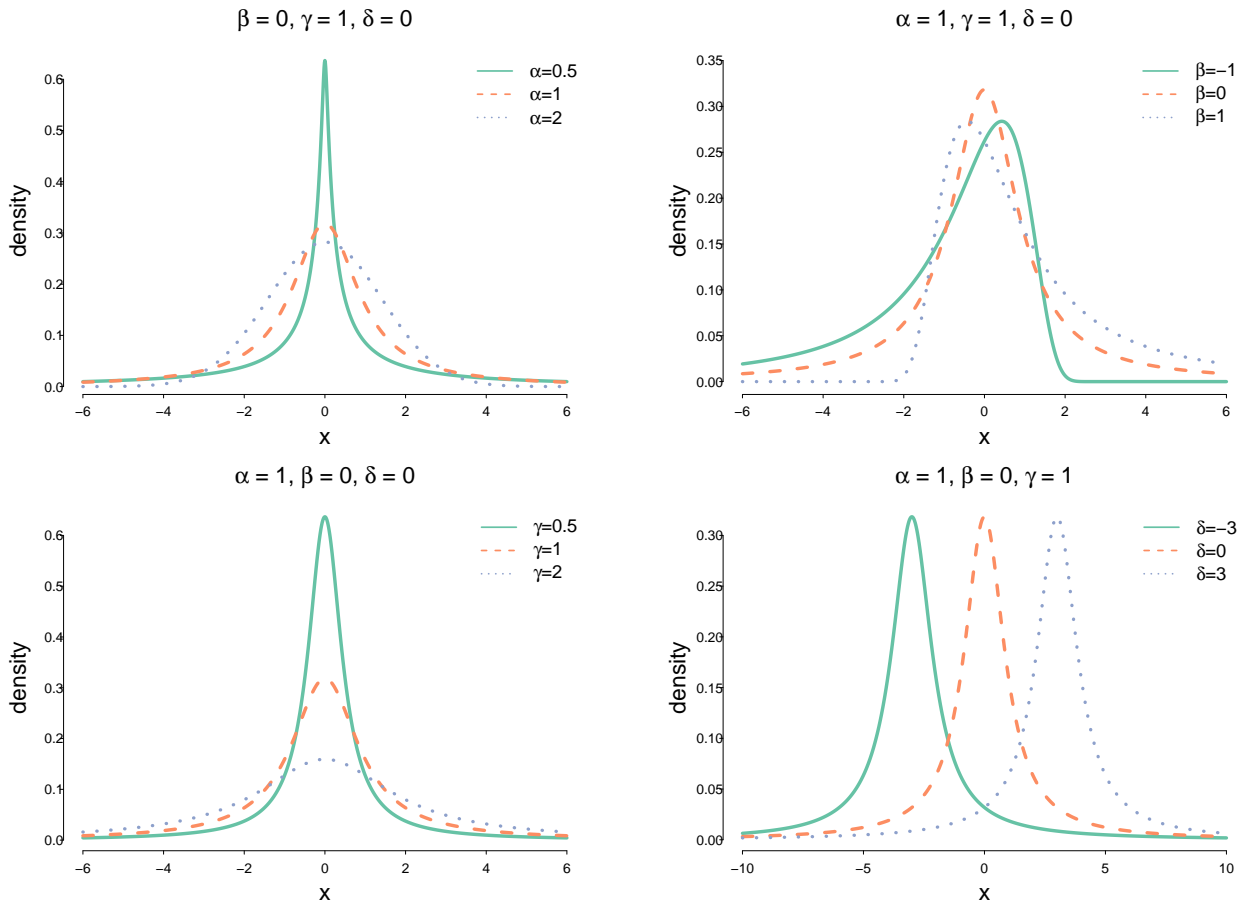


Figure 6: **Probability densities of Lévy alpha-stable distributions for different parameter values.** The top left plot varies the  $\alpha$  parameter (tail exponent), which determines how heavy the tail is. The top right plot varies the  $\beta$  parameter (skew parameter). The bottom left plot varies the  $\gamma$  parameter, which determines the ‘scale’, or ‘width’, of the distribution. The last plot varies the  $\delta$  parameter, which shifts the location of the modal value of the distribution.

Fig 6 shows the probability density function of the Lévy alpha-stable distribution, with three different values for each of the four parameters. Starting from the top left, the parameter  $\alpha \in (0, 2]$  is also called the tail exponent because the tail of the distribution decays as a power law with exponent  $\alpha$ , that is,  $P(X > x) \approx Cx^{-\alpha}$ . This implies that scaling the value of  $x$  by a factor  $h$  just scales the tail probability by  $h^{-\alpha}$ , since  $h^{-\alpha}P(X > x) \approx C(hx)^{-\alpha}$ . For example, if  $\alpha = 1$ , values of  $x$  that are twice as large (i.e.,  $h = 2$ ) as some reference  $\tilde{x}$  will be half as common, no less. As a result, with power law tails extreme values are very common and can dominate certain moments of the distribution. Thus, the tail parameter  $\alpha$  is an important indicator of dispersion. The lower  $\alpha$ , the thicker the tail (the more “tail dispersion” there is). However, note that while a smaller  $\alpha$  indicates more dispersion because of a heavier tail, a smaller  $\alpha$  also leads to a higher concentration of values close to the mode, and thus, in this sense, to a lower dispersion in the body of the distribution.

The top right panel shows the effects of the “skew” or “asymmetry” parameter  $\beta \in [-1, 1]$ . For  $\beta = 0$ , the distribution is symmetric, for  $\beta > 0$  right-skewed, and for  $\beta < 0$  it is left-skewed (except when  $\alpha = 2$ , as the skew parameter  $\beta$  vanishes in that case). In practice, we will find that productivity distributions are right skewed.

The bottom-left panel shows the effect of the scale parameter  $\gamma \in [0, +\infty]$ . The higher the scale parameter  $\gamma$ , the wider the body of the distribution. The parameter  $\gamma$  is not dimensionless; it is expressed in the same units as the data. We can get more intuition on the role of  $\gamma$  by considering special cases (Nolan 2020, p.168). First,  $\gamma$  is highly related to the (25-75) Inter Quartile Range; in fact, when  $\beta$  is small and  $\alpha$  is large (say,  $\beta = 0.5$  and  $\alpha = 1.3$ ), The interquartile range divided by 2 is a good estimator of  $\gamma$ . Second, for large enough  $\alpha$ , the re-centred data (so that  $\delta = 0$ ) can be used to estimate  $\gamma$  as the median of the absolute values.

Finally, the bottom right panel shows the effect of the location parameter  $\delta \in (-\infty, +\infty)$ , which is fairly intuitive even though it is not in general exactly related to key quantities such as the mode, mean, or median. For symmetric stable distributions, the median is equal to the location, and for asymmetric distributions the median remains close to the location as long as  $|\beta|$  is not too large and  $\alpha$  is not too small (Nolan 2020, p.168). When  $\alpha > 1$ , the sample mean is actually a consistent (but slowly converging) estimator of  $\delta$  (Nolan 2020, p.174). Nolan (2020, p.92) defines an  $\mathbf{S}_2$  parametrization specifically so that  $\delta$  is the mode.

We can gain further intuition by considering special cases: For  $\alpha = 2$ , the distribution becomes a Gaussian with variance  $2\gamma^2$  and mean  $\delta$ . For  $\alpha = 1, \beta = 0$  the distribution becomes Cauchy with scale parameter  $\gamma$  and location parameter  $\delta$ . The role of  $\delta$  and  $\gamma$  as location and scale parameters can be seen clearly by writing the Cauchy density as  $f(x) \propto \left[1 + \left(\frac{x-\delta}{\gamma}\right)^2\right]^{-1}$ . More generally, if  $X \sim \mathbf{S}_0(\alpha, \beta, \gamma, \delta)$ , then  $(X - \delta)/\gamma \sim \mathbf{S}_0(\alpha, \beta, 1, 0)$ .

One of the key characteristics of the Lévy alpha-stable distribution is that  $\gamma$  and  $\alpha$  measure qualitatively different aspects of dispersion, providing a richer perspective on observed empirical patterns. In discussing the estimation results in the following sections, we will compare the estimated  $\alpha$  and  $\gamma$  with other conventional measures of dispersion, to demonstrate the strength of parametric measures of dispersion based on the Lévy alpha-stable distribution in the case of productivity. But before, let us explain that alpha stable distributions are (in general) likely to arise, due to the GCLT.

### 3.1.2 The Generalized Central Limit Theorem

Consider the sum of i.i.d. variables (with a zero mean, without loss of generality). If they have a finite variance, then this sum divided by  $N^{1/2}$  tends in distribution to a Gaussian distribution - this is the CLT. If we are unwilling to assume that the i.i.d. variables have finite variance, the Generalized CLT states that the sum should be normalized  $N^{1/\alpha}$ , where  $\alpha$  is the tail exponent of the i.i.d. variables (or  $\alpha = 2$  if they have finite variance), and it tends to the Lévy alpha-stable distribution (which is the Normal distribution if  $\alpha = 2$ ).

We take a more formal presentation from Nolan (2020), Theorem 3.12, and state it assuming  $1 < \alpha < 2$ , which is the relevant case here, and simplifies exposition<sup>10</sup>.

**Theorem 1** *Let  $X_1, X_2 \dots$  be i.i.d. copies of  $X$ , where  $X$  has characteristic function  $\phi_X(u)$  and satisfies the tail conditions*

$$x^\alpha F(-x) \rightarrow c^- \text{ and } x^\alpha (1 - F(x)) \rightarrow c^+ \text{ as } x \rightarrow \infty,$$

<sup>10</sup> For  $\alpha > 2$ , the rescaled sum converges to a normal distribution. For  $0 < \alpha \leq 1$ , the mean of  $X$  does not exist so the expressions for centering the sum are different, but the sum still converges to an alpha stable distribution. Note also that we state the Theorem in terms of convergence to Nolan's  $\mathbf{S}_1$  parametrization, but our empirical work is conducted with the  $\mathbf{S}_0$  parametrization.

where  $F(x) = P(X \leq x)$  is the cumulative distribution function of  $X$ ,  $c^- \geq 0$ ,  $c^+ \geq 0$ ,  $0 < c^- + c^+ < \infty$ , and  $1 < \alpha < 2$ . Then

$$a_N \left( X_1 + X_2 + \dots + X_N - NE[X] \right) \xrightarrow{d} Z \sim S_1(\alpha, \beta, 1, 0),$$

where

$$\beta = \frac{c^+ - c^-}{c^+ + c^-}$$

and

$$a_n = \left( \frac{2\Gamma(\alpha) \sin(\pi\alpha/2)}{\pi(c^+ + c^-)} \right)^{1/\alpha} N^{-1/\alpha}.$$

In plain English, the sum can be re-centered using the first moment (thanks to  $1 < \alpha < 2$ ), the rescaling factor is proportional to  $N^{1/\alpha}$  (and a function of  $\alpha$  and the tail balance parameters  $c^+$  and  $c^-$ ), and the asymmetry parameter  $\beta$  depends on the relative difference between tail balance parameters  $c^+$  and  $c^-$ .

This theorem is remarkably general. For instance, assume that employees draw their productivity from a heavy tail distribution<sup>11</sup>. Firms' productivity being an average of the productivity of their employees, sufficiently large firms will have Lévy distributed productivity.

A legitimate concern with this justification for the alpha stable distribution is that employees are unlikely to draw their productivity from the same distribution, and even if they did, these draws are unlikely to be independent: the two "i"s of the i.i.d. assumption are violated.

In the case of finite variance variables, the CLT has been extended to cover independent but heterogeneous variables, non-independent but identically distributed variables, and non-independent, heterogeneous variables, see [White \(2000\)](#). There are similar attempts for infinite variance variables, but less is known and the conditions are typically more difficult to state, less intuitive, and harder to check empirically.

Starting with dependent variables, there is growing literature but it is generally focused on time series, rather than cross-sectional correlations<sup>12</sup>. That said, it is known that under "mild" time series dependence (e.g.  $m$ -dependent sequences), the GCLT still applies (see [Bartkiewicz et al. \(2011\)](#) for precise conditions and a history of this literature).

Regarding heterogeneous variables, it is important to realize that thanks to the property of stability, a sum of independent but heterogeneous stable distributions (with the same  $\alpha$ , also called "index of stability") is still a stable distribution, with parameters being an explicit function of the weights and parameters of underlying distributions ([Nolan 2020](#), Proposition 1.3). As a result, if we have a fixed number of heterogeneous distributions (say, two types of employees), we can let the number of copies go to infinity for each kind of distributions, obtain a stable distribution for each of the sums, and then sum up all the stable variables, which will give another stable distribution. A more precise statement appears in [Shintani & Umeno \(2018\)](#).

While this discussion suggests that the basic GCLT has a broader scope than i.i.d. variables, there remain a number of limitations, such as a fixed  $\alpha$  for heterogeneous variables<sup>13</sup>, the relative lack of

<sup>11</sup> There exist many mechanisms generating power laws, see [Mitzenmacher \(2004\)](#) and [Gabaix \(2009\)](#). Although not focusing on tail dispersion, [Ilzetzki & Simonelli \(2017\)](#) provide evidence of substantial dispersion in vote counting productivity across polling stations.

<sup>12</sup> In spatial and network econometrics, it is well-known that spatial correlations are qualitatively different from time series correlations, as reflected for instance in the fact that one needs to decide whether to use "infill" or "increasing domain" asymptotics ([Kelejian & Piras 2017](#)).

<sup>13</sup> Generically, the smaller  $\alpha$  dominates (see e.g. [Gabaix \(2009\)](#), Eq. 3), but convergence might be slow so that these theoretical results do not always work well in samples of even moderate size. See also [Cohen et al. \(2020\)](#) for the case of heterogeneous  $\alpha$  (but focusing on  $0 < \alpha < 1$ ).

results on GCLT for cross-sectional dependence (but see [Cohen et al. \(2020\)](#) for progress on this and on heterogenous variables), and limited results on large deviations. As a result, we refrain from attempting to provide a specific data-generating process (DGP) here. It is likely that several plausible DGPs can deliver a stable distribution, but in our view, evaluating which one is more plausible would require fitting its parameters, and simulating it using *finite* samples of sizes similar to the empirical data, as well as checking additional predictions of each model. We thus leave this for further research, and in this paper, we focus on establishing that the distribution of productivity is alpha stable, and discuss the values of the parameters for subsamples, such as different points in time, countries, and industries.

### 3.2 Distribution fit and model comparison

We use two members of the Asymmetric Exponential Power (AEP) family of distributions as benchmark. Like the Lévy alpha-stable distribution, AEP distributions generalize the Gaussian distribution to allow for skewness and more probability density in the tails. In contrast to the Lévy distribution, however, the AEP does not feature power law tails, so all its moments are finite. Various formulations of AEP distributions have been introduced since the 1980s, notably the 4-parameter AEP by [Delicado & Goría \(2008\)](#) and the 5-parameter AEP by [Bottazzi & Secchi \(2011\)](#). They have been suggested as distributional models for quantities that are known to have heavier tails than Gaussians, such as changes in currency exchange rates ([Ayebo & Kozubowski 2003](#)), and (logarithmic) firm growth rates ([Bottazzi et al. 2007](#)).

The AEP distributions provide a good benchmark because (i) they have been applied to firm-level data related to productivity (namely, sales growth rates ([Bottazzi et al. 2007](#))), (ii) they are broadly able to fit the body of the empirical distribution, but (iii) they differ radically from the Lévy alpha-stable distribution in their tail behavior.

Among the approaches to fitting AEP distributions are applications of maximum likelihood with interval-constrained optimization ([Bottazzi 2014](#)) and the method of L-moments ([Asquith 2014](#)). While the former returns a fit that is faithful to the likelihood of the empirical density, the latter will match the overall shape of the distribution better at the expense of exactness around the modal value, given the role of the third and fourth moment in the fitting procedure<sup>14</sup>. Appendix C gives the functional forms of the 4-parameter and 5-parameter AEP models, the details of the fitting procedure and goodness comparisons to the Lévy alpha-stable model fit.

---

<sup>14</sup> See [Moran et al. \(2020\)](#) for a discussion of the inability of the Laplace distribution to fit growth rates due to the sharp change of behaviour around the mode.



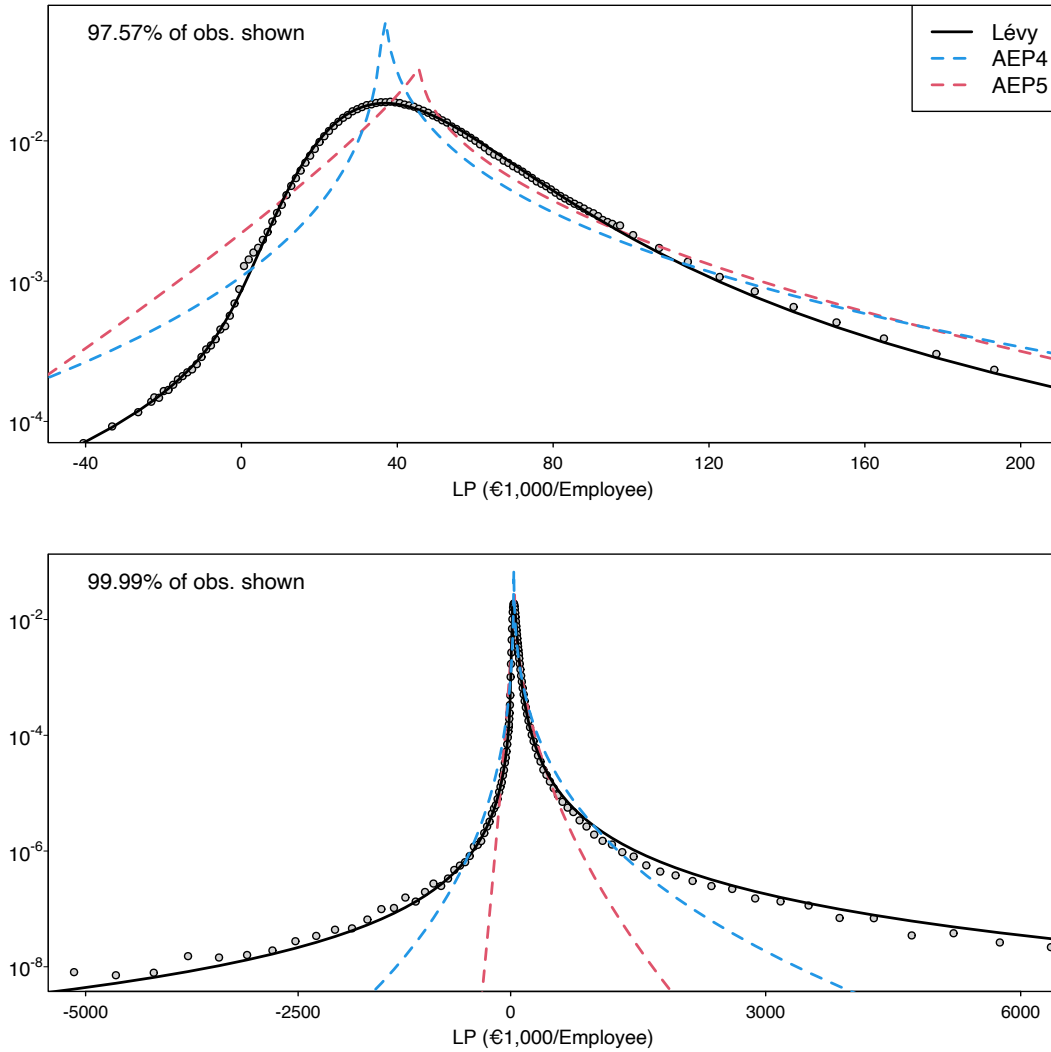


Figure 7: **Distribution of labour productivity with Lévy alpha-stable and AEP fits.** The solid black line indicates the Lévy alpha-stable, The solid red line indicates the 4-parameter AEP fit obtained with the method of L-moments and the dashed blue line indicates the 5-parameter AEP fit obtained with interval-constrained likelihood optimization.

Fig 7 shows the empirical densities and the fits for the Lévy alpha-stable, AEP distributions fitted with interval-constrained likelihood optimization and again with L-moments<sup>15</sup> for the France sample. The Lévy alpha-stable distribution is able to fit the entire domain of the empirical distribution extremely well, except for a small irregularity in the data at zero, and possibly a slightly imperfect ability to fit deal with asymmetric tail behavior (as noted in Section 2.4). In contrast, the AEP distributions fail to fit the tails and have an unsatisfactory fit for the body of the distribution. The same result holds for the other country samples, as shown in Table 9 where the log-likelihood of each model is compared.

As a further check, Fig 8 compares the scaling of the estimated standard deviation with sample size in the Lévy alpha-stable model and the 5-parameter AEP model (using a simulated sample of 2.4 million observations, which is the size of our empirical data for France). Both distributions are generated

<sup>15</sup> We employ Bottazzi (2014)'s Subbotools package for the interval-constrained likelihood optimization of the 5-parameter AEP and the R-package lmomco (Asquith 2014) for the L-moments fitting of the 4-parameter AEP. L-moments of the 5-parameter AEP distribution have to our knowledge not yet been derived.

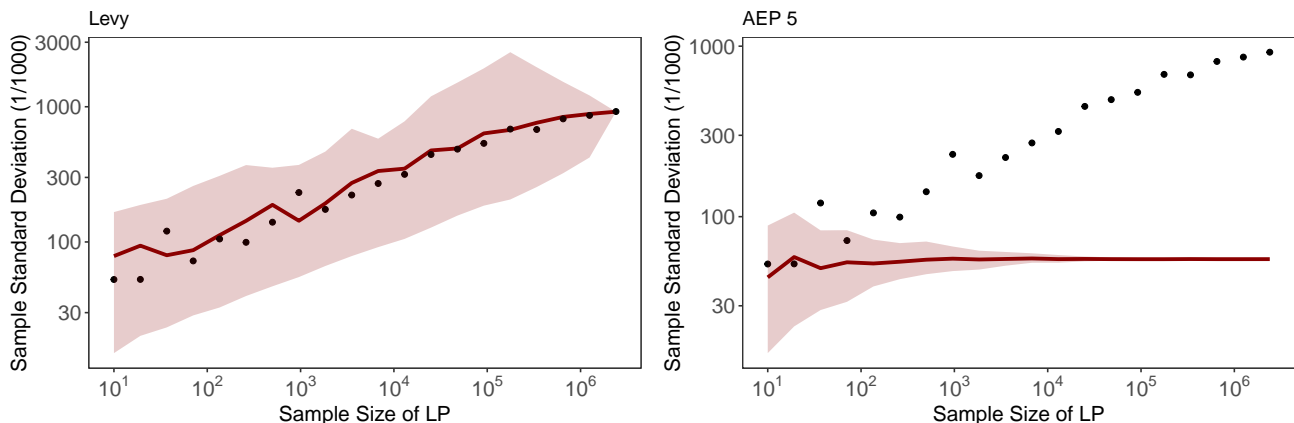


Figure 8: **Simulation of the scaling of the sample standard deviation with sample size for Lévy alpha-stable and AEP distributed data.** We generate surrogate datasets that follow either the Lévy alpha-stable (left) and for the 5-parameter AEP distribution (right), using the same sample size as the empirical data for France (2.4 million), and the parameters estimated on the empirical data. We then estimate the standard deviation in subsamples of the surrogate datasets, using many subsamples for each subsample size. The figures show mean and 5%-95% quantiles for estimates of the standard deviation. The black dots show the mean (across subsamples of given size) of the sample standard deviation using the empirical dataset.

with the parameters we estimated for the respective model with our empirical data for France. Clearly, an AEP random sample fails to reproduce the pattern which we have shown holds for both the Lévy distribution and the empirical data (Fig 1, and, for the case of France, Fig 4).

## 4 Patterns of productivity dispersion

In this section, we compare our indicators of productivity dispersion with other metrics, using two case studies: dispersion over time, and correlation of dispersion with intangible capital intensity. For the case studies, we use the most disaggregated samples at the country-year-industry level, which leads to nearly 3,500 sub-groups<sup>16</sup>. We start by explaining our choice of dispersion metrics.

### 4.1 Comparing measures of dispersion

To sum up from the earlier discussion, we cannot measure productivity dispersion using the empirical standard deviation because it is likely to be a measurement of a moment that is infinite, so its measured value depends on sample size and is not meaningful. We also cannot measure dispersion as the standard deviation of the log of VA/worker, because too many values are negative, and would need to be removed when taking the log. As a result, we are left with only one serious contender: Inter-quantile ratios.

<sup>16</sup> In principle, we would have preferred to estimate and test goodness of fit at the disaggregated level, and then discuss whether a mixture model is more appropriate when data is pooled. We acknowledge that the problem of estimating mixture models remains, and will be an important way forward to understand the extent to which country-level heterogeneity is due to within- vs between-sector heterogeneity. However, the sample sizes at a detailed level are fairly small, so here we can establish the quality of the Lévy hypothesis only at a pooled level. We use country-year-industry samples to provide an example of how our metrics could be used in practice when one has access to the larger samples typically available in administrative data.

**Computing Inter-Quantile Ratios** Inter-Quantile Ratios (IQRs) are very popular in part because they are easy to interpret. Let us denote the value of the  $p^{th}$  quantile for VA per worker by  $Q_p(V)$ . Considering for instance firms at the  $90^{th}$  and  $10^{th}$  quantiles, we define the IQR as the ratio

$$\text{IQR}_{90/10} = \frac{Q_{90}(V)}{Q_{10}(V)}, \quad (7)$$

which tells us how many times more productive the firm at the  $90^{th}$  percentile is compared to the firm at the  $10^{th}$  percentile. Usually, the variables are initially log transformed. If all values are positive, the firm sitting at the  $p^{th}$  quantile of VA per worker is also the firm sitting at the  $p^{th}$  quantile of the log(VA per worker), so we have  $\log(Q_p(V)) = Q_p(\log(V))$ , and therefore Eq. 7 can be rewritten

$$\text{IQR}_{90/10} = \exp\left(Q_{90}(\log V) - Q_{10}(\log V)\right).$$

In other words, the Inter-Quantile Ratio is the exponential of the Inter Quantile *range* of the log-transformed values. While one usually log transforms the data before computing an Inter-Quantile range, this is by no means necessary because one could just directly compute an Inter-Quantile Ratio.

This makes clear a key advantage of IQRs: even when there are negative values, IQRs can still be computed, as long as the bottom quantile is positive. For example, Spain has roughly 3-4% of negative labour productivity observations (Table 1), so we do not need to drop the negative values if we want to compare, say, the bottom 5% to the top 5%. Unfortunately, the practice of automatically taking log as a first step of the analysis has sometimes led to removing non-positive values before computing IQRs, leading to a bias in the measurement of dispersion.<sup>17</sup>

Fig 9 checks how substantial this bias can be in practice, by comparing the ratio of two quantiles before and after removing non-positive values,  $Q_{90}/Q_{10}$  and  $Q_{90_{\text{pos}}}/Q_{10_{\text{pos}}}$ . We use the country-year-industry samples at a 2-digit industry level. We exclude the samples where the bottom 10% quantile of the labour productivity is negative. By definition, the log quantile ratio based only on positive productivity values (y-axis) is always smaller than the actual log ratio, so all the points are below the 45 degree line. It appears that the larger the actual dispersion is, the higher is the degree of distortion. While in the vast majority of cases, the distortion is likely to be minor, in extreme cases the biased IQRs is two orders of magnitude smaller than the true IQRs. This is likely to happen when the lower quantile is close to zero.

---

<sup>17</sup> If one wants to take the log but keep all the observations (to keep the ranks, at least for the positive values (Campbell et al. 2019)), one can code the log of negative values as some value strictly lower than the lowest value of the log values of the positive values available in the sample. Note that coding log labour productivity as 0 would be incorrect as there could be values of VA per worker that are between 0 and 1 (particularly if expressed in thousands or millions of currency units), leading to log values that are negative.

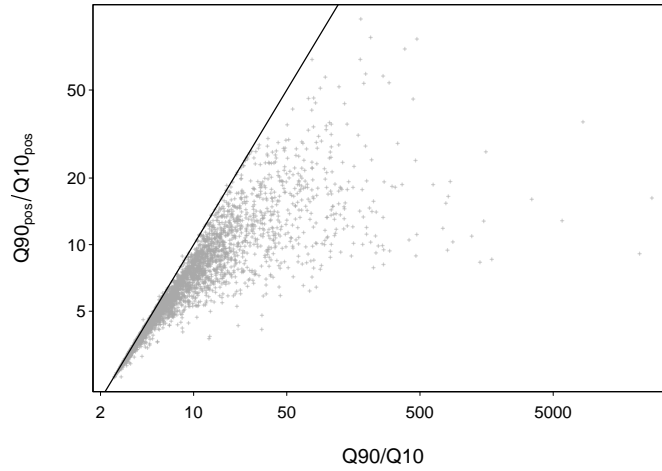


Figure 9: **Scatterplot of quantiles ratios, with and without removing non-positive values.** Both axes are on the log scale. The black line is a 45 degree line.

**Choosing IQRs** We will study two inter-quantile ratios as non-parametric measures of dispersion:  $\log(Q_{90}/Q_{10})$ , which measures dispersion in the body of the distribution, and  $\log(Q_{95}/Q_{50})$ , which measures the extent to which super-star firms deviate from a “typical” firm. Our choice is driven by previous work that has identified an increase in productivity dispersion mostly due to the divergence of superstar firms (Andrews et al. 2016, De Loecker et al. 2022). We compute the quantiles on the non-log data, but we remove observations of IQRs that are negative (i.e., when  $Q_{10}$  or  $Q_{50}$  is negative).

**Measuring dispersion using Lévy parameters.** Because the IQRs are relative, we do not want to compare them directly with  $\gamma$ , which is expressed in monetary units. To express dispersion in the body of the distribution in relative terms, we propose to use  $\gamma/\delta$ . Using a ratio of a scale over a location parameter is fairly common - for instance the standard deviation over the mean defines the well-known coefficient of variation, which is the inverse of the well-know Sharpe ratio in finance. Of course,  $\delta$  can be close to zero or even negative, but the issue also exists and is indeed worse with IQRs: the lower quantile of a (body dispersion) IQR is more likely than  $\delta$  to be very close to zero or negative, since  $\delta$  is usually close to the median.

Second, we consider  $\alpha$ , the tail parameter. We consider that a fatter tail ( $\alpha$  closer to zero) indicates more dispersion, since fatter tails indicate that there is a higher mass of super-productive firms (in almost all cases,  $\beta$  is substantially above 0, so we consider  $\alpha$  as an indicator of the fatness of the right tail.). We often use  $-\alpha$  instead of  $\alpha$ , so that this indicator, like all the others, indicates more dispersion when it increases.

We compute these four indicators at the level of industry (2 digit)-country-year. The samples are fairly small, so we remove samples with size  $N < 50$ , and we use maximum likelihood to estimate the Lévy parameters when  $50 \leq N < 200$ . We remove values of  $\delta$  that are negative and a few cases where the MLE fails or the quantile methods gives corner solutions. This leaves us with more than 3700 country-industry-year triples with estimated values of the parameters, out of an ideal maximum that would have been 19 countries  $\times$  20 industries  $\times$  12 years = 4560 triples. Appendix B.1 justifies these choices based on a Monte Carlo simulation, and Fig. 15 shows the histograms of the estimated parameters.

We also checked that the scaling of the sample standard deviation with sample size persists in some country-year and country-industry subsamples (Appendix E).

As representative examples, Fig 10 shows the distributions with the fitted lines for the French sample in 2015, looking at 4 industries. The Lévy alpha-stable model generally gives a very good fit. In particular, the Lévy alpha-stable model captures the power-law decay of the tails well. Further, the Lévy alpha-stable model outperforms the AEP 5 in terms of the log-likelihood for 88% of the subsamples.

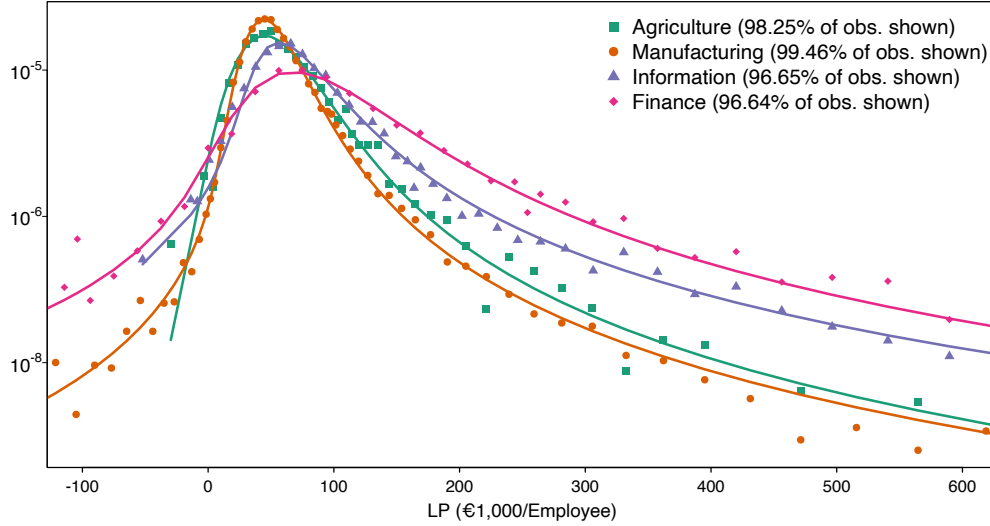


Figure 10: **Distribution of French labour productivity by industry for 2015 with Lévy alpha-stable fits.** Solid lines show the fitted Lévy alpha-stable distributions.

Table 2: **Correlations between metrics of productivity dispersion computed at the industry-country-year level.**

	$\log(\text{IQR}_{90/10})$	$\log(\text{IQR}_{95/50})$	$\log(\gamma/\delta)$	$-\alpha$
$\log(\text{IQR}_{90/10})$	1.00	0.72	0.69	0.46
$\log(\text{IQR}_{95/50})$	0.72	1.00	0.67	0.85
$\log(\gamma/\delta)$	0.69	0.67	1.00	0.37
$-\alpha$	0.46	0.85	0.37	1.00

**Correlations between metrics: Body vs Tail as two concepts of dispersion** Table 2 shows the correlation between these four metrics. It confirms that  $\text{IQR}_{90/10}$  and  $\gamma/\delta$  capture similar aspects of dispersion (dispersion in the body), and  $\text{IQR}_{95/50}$  and  $-\alpha$  capture similar aspect of dispersion (dispersion from a fatter tail). The correlation between the body- and tail- measures is fairly high for the quantile-based metrics (0.72), but much lower for the Lévy-based measures (0.37). This shows that the Lévy-based measures are potentially better able to capture distinct aspects of dispersion.

We are now ready to address two case studies: productivity dispersion over time, and the correlation between dispersion and intangible intensity.

## 4.2 Has productivity dispersion increased?

Time series patterns of productivity dispersion have been a central concern of the current literature on productivity, including reports by the OECD for advanced economies (Andrews et al. 2016, Berlingieri, Blanchenay & Criscuolo 2017), Haldane (2017) and De Loecker et al. (2022) for the UK, Gopinath et al. (2017) for European economies, and Cette et al. (2018) for France. Our goal here is to propose alternative measures of dispersion and compare them with the existing literature.

Fig 11 shows the time evolution of the dispersion metrics. For each country-year, we use unweighted averages of the industry-level estimates<sup>18</sup>

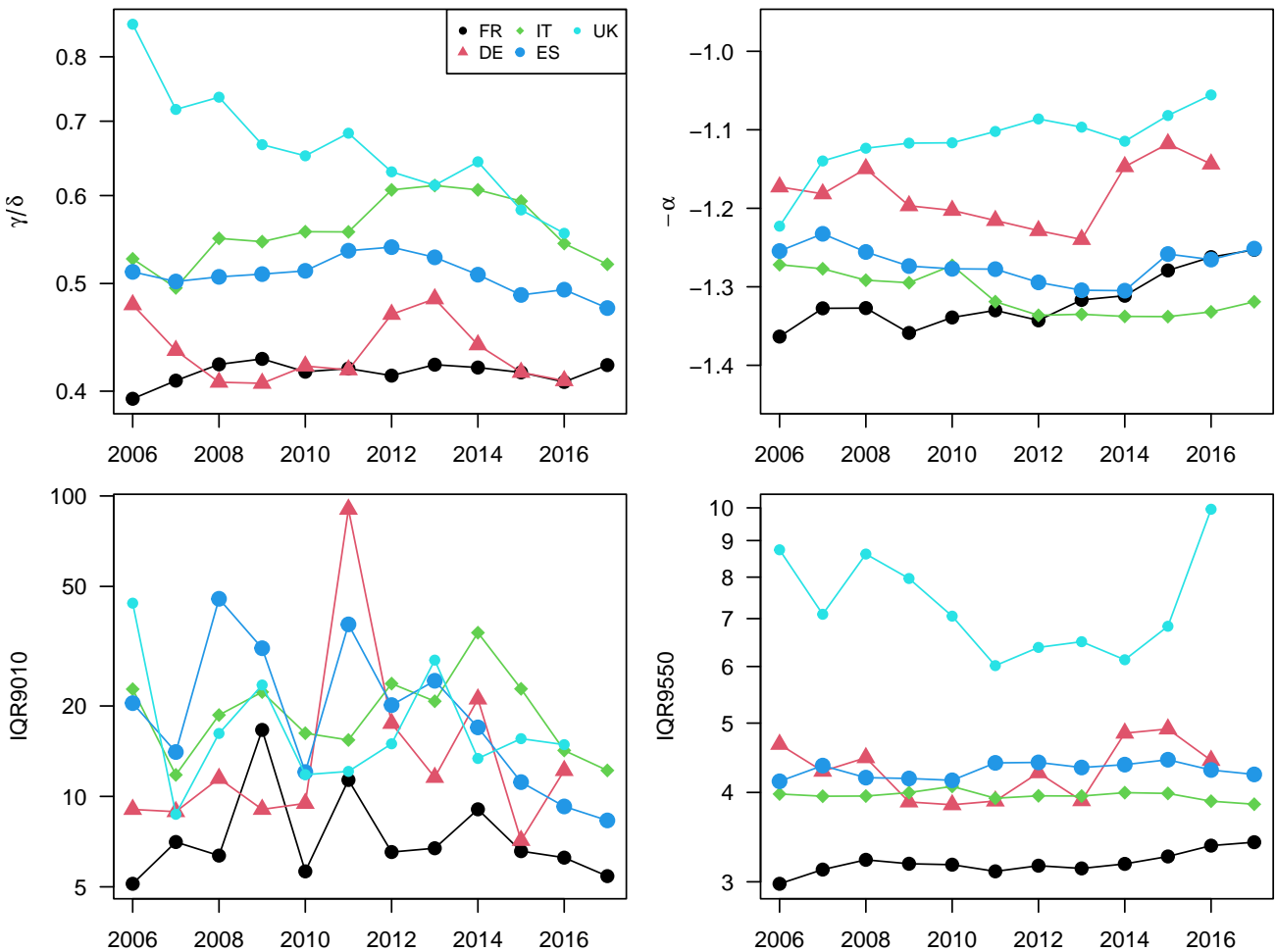


Figure 11: **Dispersion metrics parameters for selected countries over time.** Each dot is an (unweighted) average over all industries for which a value is available. We show  $-\alpha$  so that all 4 metrics should go up if dispersion increases.

There is no evidence of a generalized increase in productivity dispersion in our sample. The only country where both body and tail dispersion have increased is France. The time trends of estimated parameters are quite heterogeneous across countries (See also the country-year Lévy parameters in Fig 14 in Appendix B.2 for all 19 countries).

<sup>18</sup>This procedure implies that different years will not include the same set of industries, as some industries fail to have the required sample size only in some years, or have a negative  $\delta$  or a negative 10<sup>th</sup> percentile only in some years.

This is not necessarily at odd with previous work. Some have found that the increase in productivity dispersion took place mostly before 2006 (typically with data starting in 1996), with productivity dispersion only weakly increasing or possibly stagnating after 2006 (Andrews et al. (2016, Fig.2), De Loecker et al. (2022, Figs 3 and 4b) for the UK); De Loecker et al. (2022, Fig 4a) find increasing dispersion mostly after 2006 using Orbis, but note that using administrative data, the patterns are weaker. Oliveira-Cunha et al. (2021, Figs 8 and 10) find broadly stable dispersion for the UK and OECD countries. Various studies use different datasets, different subsets (i.e. industries), different concepts (employment weighted or not), cleaning strategies (in particular, include or exclude negative value added), periods and base years, and ways to present the data.

Putting aside the substantive results on the evolution of dispersion, however, we see that the year-on-year volatility of body dispersion is higher using quantiles rather than Lévy-based parameters, suggesting an advantage of Lévy-based metrics at capturing signal over noise. We think that part of the reason for this is the denominator, which is more stable for our Lévy-based metrics than for  $IQR_{90/10}$ . This can be seen by comparing the range of the y-axis in Fig. 11 (top-left vs bottom-left), where we see that IQRs have a numerator that is up to 100 times higher than the denominator, while  $\gamma/\delta$  is always less than 1. IQRs are also more likely to have a negative denominator, in which case these are simply dropped from our sample, which, aside from the loss of information, may increase volatility. Turning to the comparison of tail dispersion metrics, we find that Lévy-based and quantile-based metrics have relatively similar tendencies and overall volatility, with perhaps the exception of the UK, where tail behaviour clearly suggests increasing dispersion but  $IQR_{95/50}$  are volatile without clear tendency.

### 4.3 Productivity dispersion and intangible capital intensity

A prominent hypothesis in contemporary macroeconomics is that market concentration, markups and productivity dispersion have increased, and that this may be partly due to the rise of intangible capital (see e.g. Goldin et al. (2021) or De Loecker et al. (2022) for a review). Intangible capital - software and databases, knowledge capital developed from investment in R&D, branding, and employee skills - is typically a large fixed cost that also makes marginal cost lower. This results in an increase in economies of scale, which makes firms able to invest even more productive, diverging from the rest, and able to charge higher markups and capture more market shares. A recent paper (Corrado et al. 2021) found that industries with high intangible capital per worker also tend to have a higher dispersion of productivity. Here we revisit this finding to compare traditional parametric and non-parametric measures of dispersion.

We gather data on industry-level intangibles from the latest version of EU-KLEMS. Our measure of intangible intensity is a volume measure of capital services per hours worked, for each letter-level industry in 19 countries, between 2006 and 2017<sup>19</sup>.

Since we are interested only in correlations (rather than causation), we regress intangible intensity on measures of dispersion rather than the other way around as Corrado et al. (2021) do. This has several advantages: first, we can compare the goodness of fit obtained using different metrics: putting intangible intensity on the left-hand side ensures that the  $R^2$  and the standard error of the regression are comparable. Second, we can see whether including two separate metrics in the same regression

<sup>19</sup> See <https://euklems-intanprod-11ee.luiss.it>. We use the February 2022 release of EU KLEMS 2021, downloaded in March 2022 (Bontadini et al. 2021). We divide the variable  $Kq\_Intang$  (Total Capital stock intangibles, volume 2015 ref.prices, NAC mn) by the variable  $H\_EMP$  (from the “statistical” national accounts). We obtained similar results when using GFCF rather than capital stocks.

provides additional explanatory power. And third, because  $\alpha$  is bounded between 0 and 2, this avoids having to consider an appropriate estimator for bounded dependent variables.

Table 3: **Intangible intensity and quantile-based metrics of productivity dispersion.**

	Dep. variable: Log intangible intensity of a country-industry-year triples					
log(q90/q10)	0.123*		0.045	-0.029*		-0.030*
	(0.075)		(0.094)	(0.016)		(0.016)
log(q95/q50)		0.431***	0.221		-0.005	0.005
		(0.104)	(0.181)		(0.034)	(0.056)
log(q50)	1.060***	1.040***	1.059***	0.056	0.049	0.056
	(0.046)	(0.042)	(0.046)	(0.057)	(0.040)	(0.057)
Fixed effects	None	None	None	CxJ, TxJ	CxJ, TxJ	CxJ, TxJ
Degrees of freedom	2530	2835	2529	2064	2361	2063
Observations	2,533	2,838	2,533	2,533	2,838	2,533
R <sup>2</sup>	0.683	0.689	0.684	0.996	0.996	0.996
Adjusted R <sup>2</sup>	0.683	0.688	0.684	0.995	0.995	0.995
Residual Std. Error	1.114	1.109	1.112	0.135	0.141	0.135

Notes: This table regresses intangible capital intensity on various inter-quantile measures of labour productivity dispersion, by industry-country-year. The fourth, fifth and sixth columns include country-industry and year-industry fixed effects. Corresponding standard errors are in parentheses.

Significance level: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 4: **Intangible intensity and Lévy-based metrics of productivity dispersion.**

	Dep. variable: Log intangible intensity of a country-industry-year triples					
log( $\gamma/\delta$ )	0.754***		0.678***	-0.008		-0.007
	(0.115)		(0.128)	(0.038)		(0.038)
$-\alpha$		1.043***	0.506*		0.010	0.008
		(0.277)	(0.283)		(0.068)	(0.069)
log( $\delta$ )	1.048***	1.022***	1.046***	0.011	0.018	0.012
	(0.043)	(0.044)	(0.042)	(0.041)	(0.018)	(0.042)
Fixed effects	None	None	None	CxJ, TxJ	CxJ, TxJ	CxJ, TxJ
Degrees of freedom	2826	2826	2825	2352	2352	2351
Observations	2,829	2,829	2,829	2,829	2,829	2,829
R <sup>2</sup>	0.688	0.667	0.691	0.996	0.996	0.996
Adjusted R <sup>2</sup>	0.688	0.666	0.691	0.995	0.995	0.995
Residual Std. Error	1.110	1.148	1.106	0.140	0.140	0.140

Notes: This table regresses intangible capital intensity on Lévy-based measures of labour productivity dispersion, by industry-country-year. The fourth, fifth and sixth columns include country-industry and year-industry fixed effects. Corresponding standard errors are in parentheses.

Significance level: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

We broadly follow [Corrado et al. \(2021\)](#): we include a control for the central values of the distri-



bution, we provide a specification with country-by-industry and time-by-industry dummies, and we provide conservative standard errors, clustered at the industry-country level.

Tables 3 and 4 show the results. The first and perhaps most important remark is that the number of observations is generally higher when we use the Lévy-based metrics, particularly compared to regressions that use the 10<sup>th</sup> quantile, because the 10<sup>th</sup> quantile is negative in more than 10% of the cases<sup>20</sup>.

Considering regressions that do not include fixed effects, we find similar results using parametric and non parametric measures. Taken independently, both body and tail metrics have the expected sign and are significant, in both the parametric and the non parametric cases. The estimated parameters suggest an important economic effect. The elasticity of intangible capital to the 90/10 IQR is only slightly weaker than in Corrado et al. (2021, Table 3) (0.16-0.24). The elasticity for the tail dispersion IQR is more than three times higher. The elasticity for  $\gamma/\delta$ , close to 0.75, is substantial. For the tail parameter  $\alpha$ , an increase of  $-\alpha$  by 0.1 (e.g. a fattening of the tail from say 1.4 to 1.3, roughly like the evolution observed for France in Fig. 11) is associated with an increase in intangible intensity of around 10%.

The goodness of fit is similar across the IQR and Lévy tables, because they are driven by the large explanatory power of the statistics for central location (median, or  $\delta$ ). There is a sensibly higher  $R^2$  for the Lévy variables when body and tail dispersion are included together.

Table 5: LASSO estimates for intangible capital intensity and labour productivity dispersion.

	Quantiles-only	Lévy-only	Both
$\log(q90/q10)$	-0.00		-0.00
$\log(q95/q50)$	0.02		0.01
$\log(q50)$	0.45		0.43
$\log(\gamma/\delta)$		0.21	0.01
$-\alpha$		0.09	-0.02
$\log(\delta)$		0.27	0.02
Observations	2533	2829	2522
Non-zero coeffs	470	476	474

Notes: This table regresses intangible capital intensity on various measures of labour productivity dispersion, by industry-country-year, using a LASSO algorithm. Dispersion measures in the first column are based on quantiles of the relevant labour productivity distribution, and are from Lévy-based measures in the second column. The third column includes all dispersion measures available. Each regression includes country-industry and year-industry fixed effects, although the number set to zero by the algorithm is reported in the final row.

When including fixed effects (the last three columns of each Table), in all cases the coefficients drop to levels where they are no longer significant at the 5% level. Including all these dummies leads to a loss of statistical power (degrees of freedom fall from roughly 2800 to 2300), and clear overfitting ( $R^2 > 0.99$ ). As a final check, we repeat the regression of the last column of Tables 3 and 4, but using a standard penalized likelihood method. Specifically, we use a standard Lasso procedure with 10-fold cross validation as implemented in Friedman et al. (2010). Table 5 shows that once insignificant dummies are selected out, the coefficients are zero or very close to zero for the IQR metrics, but of the expected sign for the Lévy variables, although substantially smaller than when no fixed effects

<sup>20</sup> Campbell et al. (2019) note that in Australian administrative data, “At the 4-digit ANZSIC level, (...) only four industries (around 1.5 per cent of all 4-digit industries) [are] excluded for the [Inter-Quartile Range] analysis but 139 industries (around 51 per cent) [are] excluded for the 90-10 differential analysis. Nearly all 4-digit industries in Retail Trade and Wholesale Trade have a negative value added in the 10<sup>th</sup> percentile.”

are included at all. If we attempt to regress intangible intensity on all metrics of dispersion, we find inconclusive evidence, which we interpret as showing that the correlations among the regressors lead to unstable results.

Overall, these results indicate that our metrics provide similar or slightly superior results than traditional non-parametric measures. To be sure, we are confident that our Lévy-based dispersion metrics are superior *in principle*: they make it possible to keep negative values, avoid having to select specific quantiles, and while they may be slightly more difficult to interpret than quantile-based measures, they provide a link to the large body of literature in economics and finance that has shown the relevance of heavy tails and the importance of estimating precisely the tail exponent. However, we also acknowledge that in specific applications such as the one explored here, a number of methodological choices<sup>21</sup> could affect the conclusions regarding which metric provides a better fit.

## 5 Discussion and conclusion

The distributions of firm-level VA per worker have extremely large support, are asymmetric, and are heavy-tailed. A major consequence of this is that measuring dispersion is not straightforward: standard deviations are poor metrics because second moments do not exist, and log transformations are not recommended due to the proportion of the negative values.

We propose the Lévy alpha-stable distribution as a sensible distributional model for labour productivity, motivated by empirical evidence of an infinite variance, and by the fact that the Lévy distribution should be fairly common, as it emerges from the generalized Central Limit Theorem. We show that the Lévy distribution provides a better fit to the data than its main competitor, the five-parameter AEP distribution.

Good distributional models make it possible to offer a richer picture of dispersion. While the scale parameter captures the overall width of the distribution, the tail parameter captures the occurrence of the extreme events. These are qualitatively distinct aspects of dispersion. While existing research does attempt to distinguish between body and tail dispersion using quantile ratios, we argue that parametric measures make this distinction clearer and more objective, as there is no need to log-transform the data or to choose specific quantiles.

We have provided two case studies, comparing our parametric measures with non parametric measures of dispersion. We find that they give similar results: a change in dispersion (2006-2017) that differs across countries, and a positive but noisy correlation between intangible intensity and dispersion. While the results are similar, our parametric measures limit the need to discard data when measuring body dispersion, and limit the need to make arbitrary choices of quantiles.

To go further, the research on productivity dispersion needs administrative data, which typically cannot be made available. A limitation of our results is that country-industry-year samples, even at the 2-digit level, remain fairly small so it is difficult not only to estimate the parameters of the Lévy distribution, but also to test the quality of the fit. Our findings have implications for statistical offices releasing moments of micro-data. Because the Lévy alpha-stable distribution can be estimated using 5 quantiles (more conveniently the 5<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 95<sup>th</sup>), if statistical offices release more than these 5 quantiles, then the Lévy alpha-stable can be estimated on 5 quantiles and the fit can be tested

---

<sup>21</sup> For instance, using investment instead of capital, changing the threshold for the minimum sample size, changing the estimators for the Lévy-based parameters, changing the quantiles of the quantile-based metrics, and adding or removing controls.

on the other quantiles. If the fit is good, researchers can then compute any other statistic about the distribution almost as if they had access to the raw data.

Yet, our paper establishes a number of facts that are useful for the burgeoning literature on misallocation and productivity dispersion, suggesting that models should be able to reproduce Lévy-distributed VA per worker, with tail parameters  $\alpha \in [1, 1.5]$  (although possibly with a broader range when considering country-industry-year samples), substantially positive skewness ( $\beta \in [0.5, 1]$ ), and a “body” dispersion  $\gamma/\delta \in [0.2, 1]$ .

## References

- Acemoglu, D., Ozdaglar, A. & Tahbaz-Salehi, A. (2017), ‘Microeconomic origins of macroeconomic tail risks’, *American Economic Review* **107**(1), 54–108.
- Andrews, D., Criscuolo, C. & Gal, P. N. (2016), ‘The best versus the rest’, *OECD Productivity Working Papers* **05**.
- Aoyama, H., Fujiwara, Y. & Gallegati, M. (2015), ‘Micro-macro relation of production: double scaling law for statistical physics of economy’, *Journal of Economic Interaction and Coordination* **10**(1), 67–78.
- Aoyama, H., Yoshikawa, H., Iyetomi, H. & Fujiwara, Y. (2009), ‘Labour productivity superstatistics’, *Progress of Theoretical Physics, Supplement* **179**, 80–92.
- Aoyama, H., Yoshikawa, H., Iyetomi, H. & Fujiwara, Y. (2010), ‘Productivity dispersion: facts, theory, and implications’, *Journal of Economic Interaction and Coordination* **5**(1), 27–54.
- Aradanaz-Badia, A., Awano, G. & Wales, P. (2017), ‘Understanding firms in the bottom 10% of the labour productivity distribution in great britain: “the laggards”, 2003 to 2015’, *Office for National Statistics* .
- Asquith, W. (2018), *lmomco - L-moments, censored L-moments, trimmed L-moments, L-comoments, and many distributions*. R package version 2.3.1.
- Asquith, W. H. (2014), ‘Parameter estimation for the 4-parameter Asymmetric Exponential Power distribution by the method of L-moments using R’, *Computational Statistics & Data Analysis* **71**, 955 – 970.
- Axtell, R. & Guerrero, O. (forthcoming), ‘Dynamics of firms from the bottom up: Data, theories and models’.
- Axtell, R. L. (2001), ‘Zipf distribution of U.S. firm sizes’, *Science* **293**(5536), 1818–1820.
- Ayebo, A. & Kozubowski, T. J. (2003), ‘An asymmetric generalization of gaussian and laplace laws’, *Journal of Probability and Statistical Science* **1**(2), 187–210.
- Baily, M. N., Bartelsman, E. J. & Haltiwanger, J. (1996), ‘Downsizing and productivity growth: myth or reality?’, *Small Business Economics* **8**(4), 259–278.
- Bartelsman, E., Haltiwanger, J. & Scarpetta, S. (2013), ‘Cross-country differences in productivity: The role of allocation and selection’, *American Economic Review* **103**(1), 305–34.
- Bartelsman, E. J. & Doms, M. (2000), ‘Understanding productivity: Lessons from longitudinal microdata’, *Journal of Economic Literature* **38**(3), 569–594.
- Bartelsman, E. J. & Wolf, Z. (2018), Measuring productivity dispersion, in E. Grifell-Tatjé, C. K. Lovell & R. C. Sickles, eds, ‘Oxford Handbook of Productivity Analysis’, Oxford University Press, United Kingdom.
- Bartkiewicz, K., Jakubowski, A., Mikosch, T. & Wintenberger, O. (2011), ‘Stable limits for sums of dependent infinite variance random variables’, *Probability Theory and Related Fields* **150**(3), 337–372.

- Berlingieri, G., Blanchenay, P., Calligaris, S. & Criscuolo, C. (2017), 'The multiprod project: A comprehensive overview', *OECD Science, Technology and Industry Working Papers* **2017(4)**, 1.
- Berlingieri, G., Blanchenay, P. & Criscuolo, C. (2017), 'The great divergence(s)', *OECD STI Policy Papers* **39**.
- Bontadini, F., Corrado, C., Haskel, J., Iommi, M. & Jona-Lasinio, C. (2021), 'Euklems & intanprod: methods and data descriptions', *EUKLEMS website*. [https://euklems-intanprod-1lee.luiiss.it/wp-content/uploads/2022/02/EUKLEMSINTANProd\\_2021\\_Methods-and-data-description-Rev1.pdf](https://euklems-intanprod-1lee.luiiss.it/wp-content/uploads/2022/02/EUKLEMSINTANProd_2021_Methods-and-data-description-Rev1.pdf).
- Bottazzi, G. (2014), Subbotools user's manual: For version 0.9. 7.1, 8 september 2004, Technical report, LEM Working Paper Series. Originally published 2004, updated 2014, <http://www.lem.sssup.it/WPLem/files/2004-14.pdf>.
- Bottazzi, G., Cefis, E., Dosi, G. & Secchi, A. (2007), 'Invariances and diversities in the patterns of industrial evolution: Some evidence from italian manufacturing industries', *Small Business Economics* **29**(1-2), 137–159.
- Bottazzi, G. & Secchi, A. (2003), 'Common properties and sectoral specificities in the dynamics of US manufacturing companies', *Review of Industrial Organization* **23**(3), 217–232.
- Bottazzi, G. & Secchi, A. (2011), 'A new class of asymmetric exponential power densities with applications to economics and finance', *Industrial and Corporate Change* **20**(4), 991–1030.
- Bouchaud, J.-P. & Potters, M. (2003), *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, 2 edn, Cambridge University Press.
- Campbell, S., Nguyen, T., Sibelle, A. & Soriano, F. (2019), 'Measuring productivity dispersion in selected australian industries', *Treasury-ABS Working Paper*.
- Canty, A. & Ripley, B. D. (2021), *boot: Bootstrap R (S-Plus) Functions*. R package version 1.3-28.
- Cette, G., Corde, S. & Lecat, R. (2018), 'Firm-level productivity dispersion and convergence', *Economics Letters* **166**, 76–78.
- Clauset, A., Shalizi, C. R. & Newman, M. E. (2009), 'Power-law distributions in empirical data', *SIAM review* **51**(4), 661–703.
- Cohen, J. E., Davis, R. A. & Samorodnitsky, G. (2020), 'Heavy-tailed distributions, correlations, kurtosis and taylor's law of fluctuation scaling', *Proceedings of the Royal Society A* **476**(2244), 20200610.
- Corrado, C., Criscuolo, C., Haskel, J., Himbert, A. & Jona-Lasinio, C. (2021), 'New evidence on intangibles, diffusion and productivity', *OECD STI Working Papers* **2021/10**.
- Cunningham, C., Foster, L., Grim, C., Haltiwanger, J., Pabilonia, S. W., Stewart, J. & Wolf, Z. (2021), 'Dispersion in dispersion: Measuring establishment-level differences in productivity', *IZA Discussion Paper* **14459**.
- De Loecker, J., Obermeier, T. & Van Reenen, J. (2022), Firms and inequalities, in IFS, ed., 'The IFS Deaton Review of Inequalities', IFS.
- Delicado, P. & Goría, M. (2008), 'A small sample comparison of maximum likelihood, moments and l-moments methods for the asymmetric exponential power distribution', *Computational Statistics & Data Analysis* **52**(3), 1661–1673.
- Embrechts, P., Kluppelberg, C. & Mikosch, T. (1997), Time series analysis for heavy-tailed processes, in 'Modelling Extremal Events', Springer, Berlin, Heidelberg, pp. 371–412.
- Fagiolo, G., Napoletano, M. & Roventini, A. (2008), 'Are output growth-rate distributions fat-tailed? some evidence from oecd countries', *Journal of Applied Econometrics* **23**(5), 639–669.

- Fama, E. F. (1965), 'Portfolio analysis in a stable Paretian market', *Management Science* **11**(3), 404–419.
- Foster, L., Grim, C., Haltiwanger, J. C. & Wolf, Z. (2021), Innovation, productivity dispersion, and productivity growth, in C. Corrado, J. Haskel, J. Miranda & D. Sichel, eds, 'Measuring and Accounting for Innovation in the Twenty-First Century', University of Chicago Press, p. 103.
- Friedman, J. H., Hastie, T. & Tibshirani, R. (2010), 'Regularization paths for generalized linear models via coordinate descent', *Journal of Statistical Software* **33**(1), 1–22.
- Gabaix, X. (2009), 'Power laws in economics and finance', *Annual Review of Economics* **1**(1), 255–294.
- Gabaix, X. (2011), 'The granular origins of aggregate fluctuations', *Econometrica* **79**(3), 733–772.
- Gaffeo, E. (2008), 'Lévy-stable productivity shocks', *Macroeconomic Dynamics* **12**(3), 425–443.
- Gaffeo, E. (2011), 'The distribution of sectoral TFP growth rates: International evidence', *Economics Letters* **113**(3), 252–255.
- Ghiglino, C. (2012), 'Random walk to innovation: Why productivity follows a power law', *Journal of Economic Theory* **147**(2), 713–737.
- Goldin, I., Koutroumpis, P., Lafond, F. & Winkler, J. (2021), 'Why is productivity slowing down?', *INET Oxford Working Paper* .
- Gopinath, G., Kalemli-Özcan, Ş., Karabarbounis, L. & Villegas-Sanchez, C. (2017), 'Capital allocation and productivity in south europe', *The Quarterly Journal of Economics* **132**(4), 1915–1967.
- Gouin-Bonenfant, E. (2020), 'Productivity dispersion, between-firm competition and the labor share', *Draft* .
- Gu, W. (2019), 'Frontier firms, productivity dispersion and aggregate productivity growth in canada', *International Productivity Monitor* (37), 96–119.
- Haldane, A. (2017), 'Productivity puzzles', *Speech given by Andy Haldane, Chief Economist, Bank of England at the London School of Economics* .
- Haltiwanger, J., Kulick, R. & Syverson, C. (2018), 'Misallocation measures: The distortion that ate the residual', *NBER Working Paper* **24199**.
- Heinrich, T. & Winkler, J. (2020), *finity: Test for Finiteness of Moments in a Distribution*. R package version 0.1.4.1. **URL:** <https://CRAN.R-project.org/package=finity>
- Holly, S., Petrella, I. & Santoro, E. (2013), 'Aggregate fluctuations and the cross-sectional dynamics of firm growth', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **176**(2), 459–479.
- Hsieh, C.-T. & Klenow, P. J. (2009), 'Misallocation and manufacturing TFP in China and India', *The Quarterly Journal of Economics* **124**(4), 1403–1448.
- Ijiri, Y. & Simon, H. A. (1977), *Skew distributions and the sizes of business firms*, Vol. 24, North Holland, Amsterdam.
- Ilzetzki, E. & Simonelli, S. (2017), 'Measuring productivity dispersion: Lessons from counting one-hundred million ballots', *CEPR Discussion Paper* **DP12273**.
- Jäger, K. & The Conference Board (2018), EU KLEMS growth and productivity accounts 2017 release, statistical module 1, Technical report, EU KLEMS Project. Revised 2018, Available online [http://www.euklems.net/TCB/2018/Methology\\_EUKLEMS\\_2017\\_revised.pdf](http://www.euklems.net/TCB/2018/Methology_EUKLEMS_2017_revised.pdf), [http://www.euklems.net/TCB/2018/EUKLEMS\\_2018\\_revision.pdf](http://www.euklems.net/TCB/2018/EUKLEMS_2018_revision.pdf).

- Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V. & Yesiltas, S. (2015), 'How to construct nationally representative firm level data from the orbis global database', *NBER Working Paper* 21558.
- Kelejian, H. & Piras, G. (2017), *Spatial Econometrics*, Academic Press.
- König, M. D., Lorenz, J. & Zilibotti, F. (2016), 'Innovation vs. imitation and the evolution of productivity distributions', *Theoretical Economics* 11(3), 1053–1102.
- Lucas, R. E. & Moll, B. (2014), 'Knowledge growth and the allocation of time', *Journal of Political Economy* 122(1), 1–51.
- Mandelbrot, B. (1960), 'The Pareto-Lévy law and the distribution of income', *International Economic Review* 1(2), 79–106.
- Mandelbrot, B. (1963), 'New methods in statistical economics', *Journal of Political Economy* 71(5), 421–440.
- McCulloch, J. H. (1986), 'Simple consistent estimators of stable distribution parameters', *Communications in Statistics-Simulation and Computation* 15(4), 1109–1136.
- Mitchell, W. C. (1915), 'The making and using of index numbers', *US BLS Bulletin* 284. Reprinted 1938, accessible at [https://fraser.stlouisfed.org/files/docs/publications/bls/bls\\_0656\\_1938.pdf](https://fraser.stlouisfed.org/files/docs/publications/bls/bls_0656_1938.pdf).
- Mitzenmacher, M. (2004), 'A brief history of generative models for power law and lognormal distributions', *Internet Mathematics* 1(2), 226–251.
- Mizuno, T., Ishikawa, A., Fujimoto, S. & Watanabe, T. (2012), 'Power laws in firm productivity', *Progress of Theoretical Physics Supplement* 194, 122–134.
- Moran, J., Secchi, A. & Bouchaud, J.-P. (2020), Statistical facts on growth rates, in J. Moran, ed., 'Statistical physics and anomalous macroeconomic fluctuations, PhD Thesis', EHESS press.
- Newman, M. E. (2005), 'Power laws, Pareto distributions and Zipf's law', *Contemporary Physics* 46(5), 323–351.
- Nolan, J. P. (1998), 'Parameterizations and modes of stable distributions', *Statistics & Probability Letters* 38(2), 187–195.
- Nolan, J. P. (2020), *Stable Distributions - Models for Heavy Tailed Data*, Springer.
- Office for National Statistics (2022), 'Firm-level labour productivity estimates from the annual business survey (abs): summary statistics', available online at <https://www.ons.gov.uk/economy/economicoutputandproductivity/productivitymeasures/datasets/firmlevellabourproductivityestimatesfromtheannualbusinesssurveyabssummarystatistics>.
- Oliveira-Cunha, J., Kozler, J., Shah, P., Thwaites, G. & Valero, A. (2021), Business time: How ready are uk firms for the decisive decade?, Technical report, The Resolution Foundation.
- Oulton, N. (2000), 'A tale of two cycles: Closure, downsizing and productivity growth in uk manufacturing, 1973-89', *National Institute Economic Review* 173, 66–79.
- Perla, J. & Tonetti, C. (2014), 'Equilibrium imitation and growth', *Journal of Political Economy* 122(1), 52–76.
- Resnick, S. I. (2007), *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer Series in Operations Research and Financial Engineering, Springer.
- Rotemberg, M. & White, T. K. (2021), 'Plant-to-table (s and figures): Processed manufacturing data and measured misallocation', *Working Paper*. [https://wp.nyu.edu/mrotemberg/wp-content/uploads/sites/8049/2021/07/RW\\_plant\\_to\\_table\\_July\\_2021.pdf](https://wp.nyu.edu/mrotemberg/wp-content/uploads/sites/8049/2021/07/RW_plant_to_table_July_2021.pdf).

- Samuelson, P. A. (1967), 'Efficient portfolio selection for Pareto-Lévy investments', *Journal of Financial and Quantitative Analysis* 2(2), 107–122.
- Schwarzkopf, Y., Axtell, R. & Farmer, J. D. (2010), 'An explanation of universality in growth fluctuations', *SSRN* 1597504.
- Shintani, M. & Umeno, K. (2018), 'Super generalized central limit theorem—limit distributions for sums of non-identical random variables with power laws—', *Journal of the Physical Society of Japan* 87(4), 043003.
- Sornette, D. (2006), *Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: concepts and tools*, Springer Science & Business Media.
- Souma, W., Ikeda, Y., Iyetomi, H. & Fujiwara, Y. (2009a), 'Distribution of labour productivity in japan over the period 1996–2006', *Economics: E-journal* 3(1).
- Souma, W., Ikeda, Y., Iyetomi, H. & Fujiwara, Y. (2009b), 'Distribution of labour productivity in japan over the period 1996-2006', *Economics: The Open-Access, Open-Assessment E-Journal* 3.
- Syverson, C. (2011), 'What determines productivity?', *Journal of Economic Literature* 49(2), 326–65.
- Trapani, L. (2016), 'Testing for (in)finite moments', *Journal of Econometrics* 191(1), 57 – 68.
- Voitalov, I., van der Hoorn, P., van der Hofstad, R. & Krioukov, D. (2019), 'Scale-free networks well done', *Physical Review Research* 1(3), 033034.
- White, H. (2000), *Asymptotic theory for econometricians, revised edition*, Academic press.

## A Data appendix

### A.1 Raw data

We use the Orbis Europe firm level database, which is part of the Orbis data provided by Bureau van Dijk. The database encompasses more than 21 million firms of all sizes from 44 different European countries. Around 23 million firm-year observations for 7 million unique firms in 19 countries are used in the present analysis; see Table 7. The variables used are listed in Table 6.

Orbis Code	Notation	Description
IDNR		Firm's identification number
EBITDA	$\pi$	Nominal Earnings Before Interest, Taxes, Depreciation & Amortization
STAF	$\omega$	Nominal Wages (staff costs)
EMPL	$L$	Employment
NACE_PRIM_CODE		Industrial classification code (NACE Rev. 2)
CLOSDATE_year		Year

Table 6: Variables from Orbis Europe used for the analysis. See Eq. 1 for the construction of VA and Labour productivity.

### A.2 Data cleaning

The first step in processing the data is to ensure that no observations have missing values for their reporting IDNR or year. We only use unconsolidated in order to avoid double counting. Since some accounts provide a default year for various reasons, observations where the reported year is not between 2006 and 2017 are removed. Negative total asset, fixed asset, sales, wages, and employment observations are regarded as missing.

Furthermore, we remove firm-year observations where employment or value added are missing. In addition, we discard country-year samples without at least 10,000 observations for five years or more. This reduces the list of countries considered from 44 to 19 countries (See Table 7).<sup>22</sup>

### A.3 Deflation

The KLEMS database<sup>23</sup> provides the most comprehensive data on deflation for the countries and industries covered by the Orbis Europe sample. In particular, tables `Statistical_National-Accounts.rds` and `Statistical_Capital.rds` provide value added and gross output for two-digit NACE Rev.2 industries for the countries in the Orbis Europe sample. Depending on its location, industry, and year of reporting, we deflate firms' value added using the value added deflator

<sup>22</sup> For 25 countries, there was insufficient data: Albania, Austria, Belarus, Bosnia and Herzegovina, Cyprus, Greece, Iceland, Ireland, Kosovo, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, Norway, North Macedonia, Moldova, Russian Federation, Serbia, Switzerland, Turkey, Ukraine.

<sup>23</sup> <http://www.euklems.eu>. This is version that was available when we started the project, so it differs from the EU-KLEMS data used in Section 4.3.



	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Belgium	24,823	88,877	94,373	94,899	95,631	96,810	97,357	97,093	97,324	97,629	96,909	60,423
Bulgaria	0	12,231	38,205	38,136	39,190	81,215	93,304	94,817	95,577	99,594	99,453	102,942
Croatia	16,451	38,875	39,957	42,387	40,536	39,360	40,340	40,472	41,208	42,434	44,062	0
Czech Republic	24,328	52,697	46,049	70,294	73,284	77,457	76,924	78,795	75,475	77,139	60,192	40,554
Denmark	0	0	0	0	0	0	10,153	13,116	12,649	14,232	41,054	45,520
Estonia	0	14,505	16,095	13,939	15,932	16,909	18,088	19,017	20,185	20,511	20,926	19,876
Finland	14,585	31,271	30,235	32,541	32,344	38,403	38,749	39,827	40,825	37,529	35,797	36,760
France	188,682	248,892	235,860	249,724	259,352	233,105	199,456	207,906	207,671	155,439	128,023	82,804
Germany	19,537	36,256	37,810	38,440	39,498	42,164	50,222	73,045	40,261	37,871	33,155	0
Hungary	0	65,589	16,220	109,059	91,033	88,773	120,705	121,833	129,449	132,779	133,472	129,107
Italy	80,270	205,580	300,766	255,141	201,252	406,372	425,845	418,763	422,876	437,585	435,280	380,695
Poland	13,548	25,968	31,798	54,341	18,129	14,024	15,311	13,514	0	0	0	25,350
Portugal	97,015	202,418	208,151	201,864	197,862	192,167	187,614	184,607	187,208	191,268	193,563	193,080
Romania	0	167,761	175,666	161,536	162,342	173,394	179,953	182,637	178,535	183,662	191,385	0
Slovakia	0	16,158	22,700	33,137	32,910	35,012	33,821	39,242	43,622	34,241	34,956	0
Slovenia	0	0	0	9,659	29,250	29,453	27,599	27,442	28,932	29,746	30,959	31,813
Spain	270,505	424,877	468,998	461,257	436,631	424,314	404,536	390,900	392,478	403,412	405,003	345,054
Sweden	18,283	83,593	100,106	102,341	104,274	106,361	107,975	108,888	110,179	112,690	114,515	100,681
United Kingdom	16,128	49,873	53,686	51,906	50,259	49,457	49,832	51,090	52,126	52,793	52,818	0

Table 7: Number of observations per country-year (after cleaning)

#### A.4 Industry classification

Each firm reported in Orbis Europe includes its four-digit NACE Rev. 2 industrial classification. We further aggregate the industry classes using the highest aggregation level in the original NACE Rev. 2 industrial classification. This yields 19 industries. Table 8 summarizes the industry category and descriptions.

Broad Category	Original NACE Rev. 2 Category
Agriculture	Agriculture, forestry and fishing (Agr)
	Mining and quarrying (Mine)
Manufacturing	Manufacturing (Manu)
Energy	Electricity, gas, steam and air conditioning supply (Elec)
	Water supply; sewerage, waste management and remediation activities (Water)
Construction	Construction (Cons)
Non-Financial Service	Wholesale and retail trade; repair of motor vehicles and motorcycles (Whole)
	Transportation and storage (Trans)
	Accommodation and food service activities (Accom)
	Administrative and support service activities (Adm-S)
	Arts, entertainment and recreation (Art)
	Other service activities (O-Serv)
Information	Information and communication (Info)
Finance, Insurance, Real Estate	Financial and insurance activities (F&I)
	Real estate activities (Real)
	Professional, scientific and technical activities (Prof-S)
Non-Market Service	Public administration and defence; compulsory social security (Pub-S)
	Education (Edu)
	Human health and social work activities (Health)

Table 8: Industry Category Codes and Descriptions

## B Fitting procedure

### B.1 Fitting methods

Nolan (2020, Chap. 4) provides a recent review and evaluation of the various methods for estimating the parameters of stable distributions. For asymmetric distributions, there are three main contenders: maximum likelihood (MLE), the quantile-based estimator of McCulloch (1986), and a method based on empirical characteristic functions. Nolan evaluates these methods for 10 different points in the parameter space. The most relevant to us is likely to be his case 8 ( $\alpha = 1.5, \beta = 0.5$ ). In this case and for sample sizes of 100, the MLE is the best method, followed by the empirical characteristic function, and the quantile method. As can be seen in his Fig. 4.22 p.196, while the differences in performance are not very high in general, they are fairly substantial for  $\alpha$ , with a RMSE of 0.156 for the MLE and 0.227 for the quantile estimator. This means that a roughly constructed 95% interval ( $\alpha \pm 2$  RMSE), which assumes unbiasedness as suggested by Nolan's Fig.4.22, would be [1.19-1.81] for the MLE, and [1.05-1.95] for the quantile estimator, so the difference is noticeable.

Unfortunately, the MLE is computationally highly prohibitive even for moderate sample sizes, so we want to use it only when absolutely necessary. To decide on the minimal sample size needed for using the quantile estimator, we run a Monte Carlo simulation as follows. We set  $\alpha = 1.3, \beta = 0.75, \gamma = 10, \delta = 20$ , and we investigate sample sizes  $N = 20, 50, 100, 200, 500$ . For each sample size we draw 200 samples and estimate the parameters using the MLE and the quantiles method. We then compute the mean and the 10<sup>th</sup> and 90<sup>th</sup> quantile of these 200 observations. We discard non convergent MLE results, but we keep "corner" solutions from the quantile estimator ( $\alpha = 0.5$ , which is the lower bound for this estimator (Nolan 2020, p.168)).

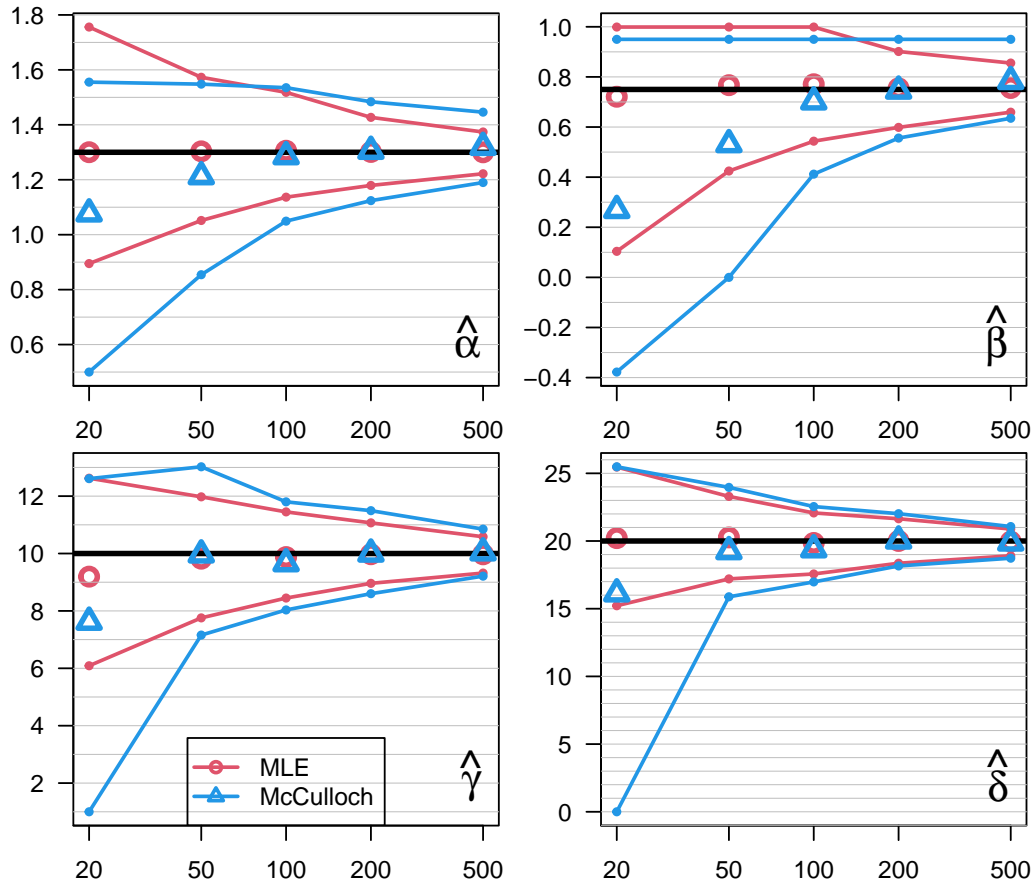


Figure 12: **Performance of the MLE and Quantile estimators in small samples.** For each sample size, we sample 200 sets of i.i.d. values. The black lines show the true values of the parameters. The dots show the average of the estimated parameters over the 200 repetitions, and the lines show the 10<sup>th</sup> and 90<sup>th</sup> percentiles.

Fig. 12 shows the results. As expected, the MLE generally outperforms the quantile estimator. A important point is that the quantile estimator gives *biased* estimates of all parameters for small sample sizes. However, the MLE remains approximately unbiased even for small samples. While the estimates are quite imprecise, it is still clear that valuable information can be extracted from a small sample, particularly if using the MLE on samples of size greater than 50. We choose  $N > 50$  as our criteria for the minimal sample size in the paper. When  $N$  reaches 200, both estimators appear unbiased, and the precision advantage of the MLE is less obvious (at least in absolute value, note that the y-axes are not in log scale). Given our computational constraints, we set  $N = 200$  as the minimum sample size for using the quantile estimator rather than the MLE.

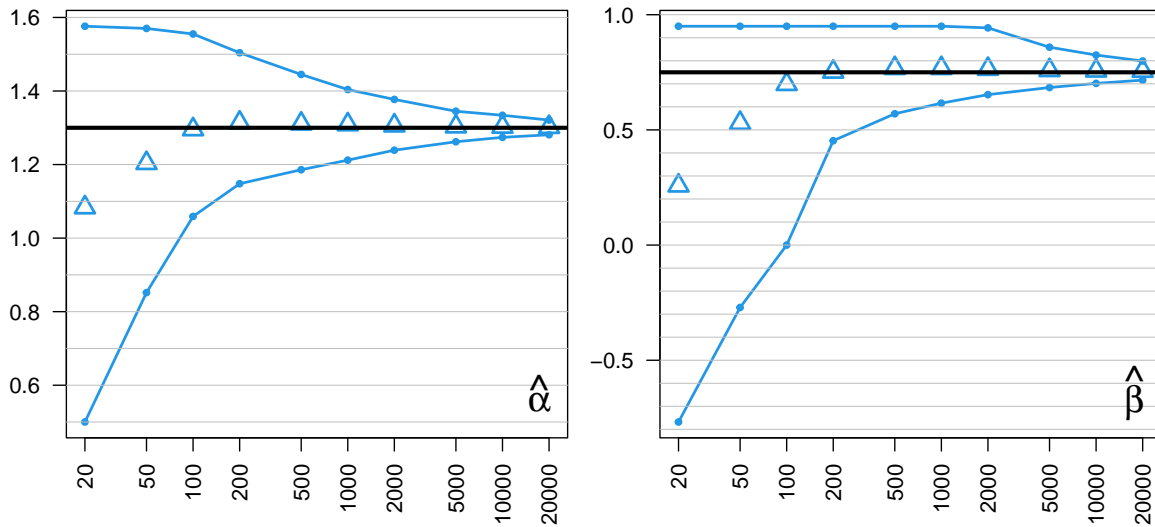


Figure 13: **Performance of the Quantile estimators in moderate sample sizes.** For each sample size, we sample 10,000 sets of i.i.d. values. The black lines shows the true values of the parameters. The dots show the average of the estimated parameters over the 200 repetitions, and the lines show the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles.

One remaining concern with the quantile estimator is the estimate of  $\beta$ . The implementation we use reports a maximum  $\hat{\beta} = 0.95$  rather than  $\hat{\beta} = 1$ , and this maximum is still reached quite often for moderate sample sizes  $N = 200, 500$  when  $\beta = 0.75$ . Similarly, the minimum value that the estimator of  $\alpha$  can return is 0.5, so the value 0.5 for the 10<sup>th</sup> percentile at  $N = 20$  is not very informative. To investigate this issue we run another set of simulations, using the same values of the parameters but 10,000 replications, and sample sizes up to 20,000. We also change the reported percentiles from  $[0.1, 0.9]$  to  $[0.025, 0.975]$ . Fig. 13 reports the results for  $\hat{\beta}$ , which is our main interest, and  $\hat{\alpha}$ , to provide an idea of the precision of the quantile estimator for this key parameter for large sample size. The 95% intervals for  $\hat{\beta}$  do still include 0.95 when  $N = 2,000$ , but not when  $N = 5,000$ . The Figure also complements Fig. 12 to make clear the bias and lack of precision for  $N < 200$ .

## B.2 Estimated parameters

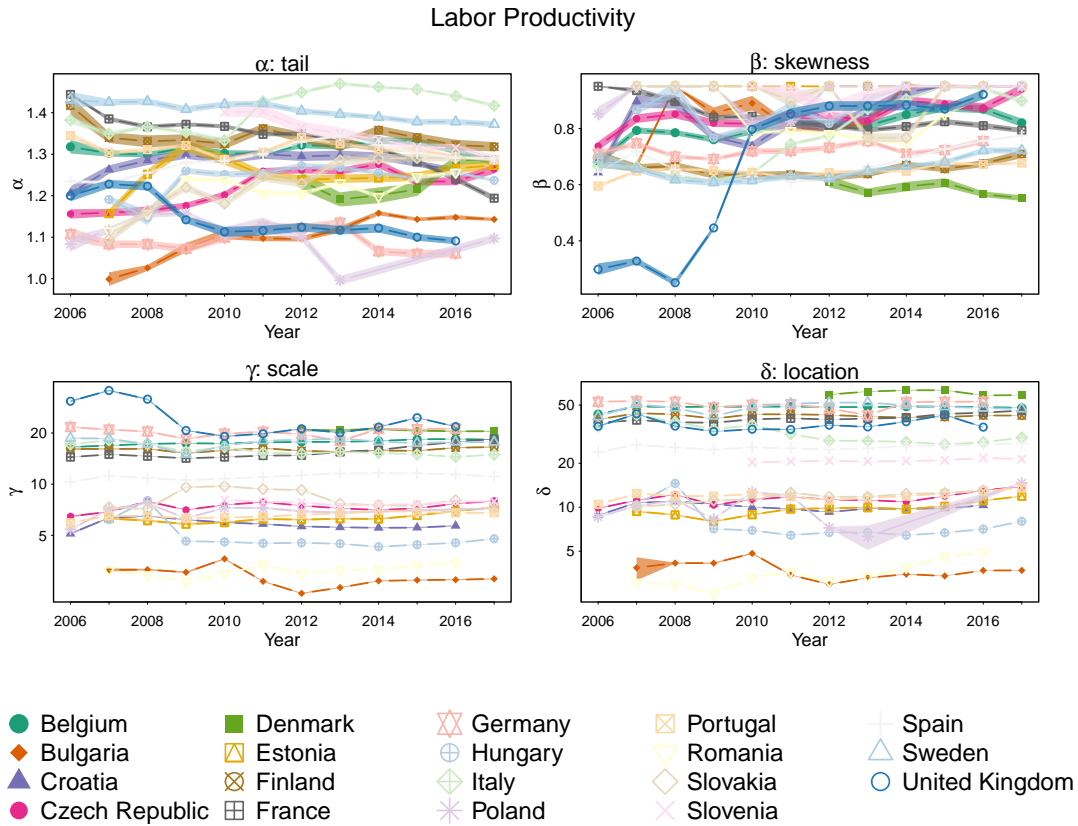


Figure 14: **Estimated parameters for country-year samples.** The four parameters of the fitted Lévy alpha-stable distributions for labour productivity levels (top two rows) and change (bottom two rows) are plotted by year (2006-2017 for levels, 2007-2017 for change).  $\gamma$  and  $\delta$  are denominated in €1,000/employee and in log scale,  $\alpha$  and  $\beta$  are dimensionless. Note that  $\gamma$  and  $\delta$  parameters in non-euro countries are converted to Euro using the average exchange rates for the country's sample time period. The exchange rate data is obtained from Eurostat<sup>24</sup>. The shaded area represents the range of  $\pm 1$  standard errors. The standard errors are calculated by bootstrapping with 1000 replications. Each replication is obtained by sampling from the data *with replacement*. We use R package boot (Canty & Ripley 2021).

<sup>24</sup>[https://ec.europa.eu/eurostat/databrowser/view/ert\\_bil\\_eur\\_a/default/table?lang=en](https://ec.europa.eu/eurostat/databrowser/view/ert_bil_eur_a/default/table?lang=en)

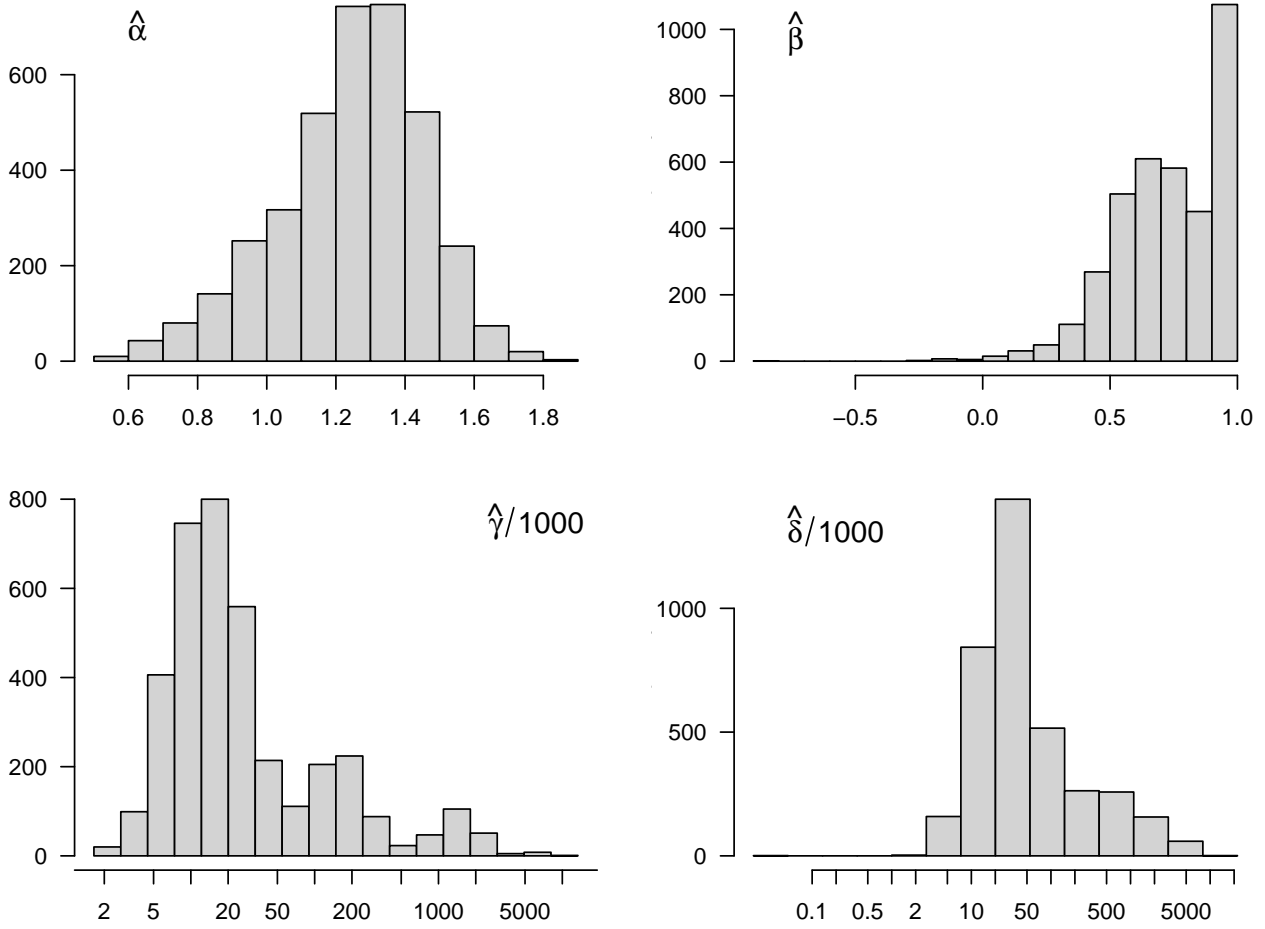


Figure 15: Distribution of the estimated parameters for the country-industry-year samples. The parameters are not estimated for sample sizes  $N < 50$ , are estimated by MLE for sample sizes  $50 \leq N < 200$  and are estimated by the quantile method for samples  $N \geq 200$ . We removed a few cases where  $\hat{\delta} < 0$  or where we had both  $\hat{\alpha} = 0.5$  and  $\hat{\gamma} = 1$  (a sign of failure of the quantile method).

## C Reference model and comparison of the fit

As a reference model, we use two variants of the Asymmetric Exponential Power (AEP) distribution - the 4-parameter AEP by [Delicado & Goría \(2008\)](#) and the 5-parameter AEP by [Bottazzi & Secchi \(2011\)](#) - that are estimated with two different fitting procedures with radically different characteristics. In the following, we give the functional forms, details of the fitting procedures, and a comparison of the goodness of the results.

## C.1 Functional forms of the AEP models and fitting procedure

### C.1.1 4-parameter AEP

The probability density function of the distribution is

$$f(x) = \frac{\alpha \kappa}{\sigma(1 + \kappa^2)\Gamma\left(\frac{1}{\alpha}\right)} \exp\left(-\left(\kappa^{\text{sign}(x-\theta)} \left|\frac{x-\theta}{\sigma}\right|\right)^\alpha\right),$$

where the parameters are location ( $\theta$ ), scale ( $\sigma$ ), skewness ( $\alpha$ ) and tails ( $\kappa$ ). We use the method from (Asquith 2014) to estimate these parameters. It used uses L-moments, the 3rd and 4th moments specifically, and will therefore achieve a better representation of the shape of the empirical data in the tails at the expense of a more exact fit in the body. For the computation, we use the R-package **Imomco** (Asquith 2018).

### C.1.2 5-parameter AEP

The probability density function of the 5-parameter AEP distribution is (Bottazzi & Secchi 2011)

$$f(x) = \begin{cases} \frac{1}{A} \exp\left(-\frac{1}{b_l} \left|\frac{x-\theta}{a_l}\right|^{b_l}\right) & \text{for } x < \theta \\ \frac{1}{A} \exp\left(-\frac{1}{b_r} \left|\frac{x-\theta}{a_r}\right|^{b_r}\right) & \text{for } x > \theta \end{cases}$$

with

$$A = a_l b_l^{\frac{1}{b_l}} \Gamma\left(1 + \frac{1}{b_l}\right) + a_r b_r^{\frac{1}{b_r}} \Gamma\left(1 + \frac{1}{b_r}\right).$$

$\theta$  remains the location parameter; the scale parameter ( $\sigma$  in the 4-parameter version) is now separate for the left and right side of the distribution  $x < \theta$  and  $x > \theta$  as scale parameters  $a_l$  and  $a_r$ . The skewness and tail parameters of the 4-parameter version are rearranged into a left-shape and a right-shape parameter  $b_l$  and  $b_r$ . The distribution recovers the normal distribution  $\mathcal{N}(\mu, \sigma)$  for parameter values  $\theta = \mu, a_l = a_r = \sigma, b_l = b_r = 2$ .

The distribution can be fitted with interval-constrained likelihood optimization, which is, however, computationally expensive. We used the Subbotools package (Bottazzi 2014). To speed up the computation, we first perform a 3-parameter exponential power<sup>25</sup> fit and then a 5-parameter AEP fit that is initialized with the estimates from the 3-parameter exponential power fit.

The L-moments have to our knowledge not been derived, therefore the alternative fitting method of L-moments (Asquith 2014) is not available in this case.

## C.2 Comparison of the goodness of fit

Table 9 compares the average log-likelihoods of the Lévy alpha-stable and the AEP distributions for all country samples. The log-likelihood of Lévy alpha-stable model is always higher.

<sup>25</sup>This distribution is the standard Subbotin distribution and is obtained for  $a_l = a_r$  and  $b_l = b_r$  for the 5-parameter AEP.

Country	Obs	AEP4 Log Lik	AEP5 Log Lik	Lévy Log Lik
Belgium	1,042,148	-5.20	-5.11	-5.04
Bulgaria	794,664	-4.18	-4.31	-4.03
Croatia	426,082	-6.09	-6.05	-5.97
Czech Republic	753188	-7.70	-7.69	-7.51
Denmark	136,724	-7.65	-7.34	-7.25
Estonia	195,983	-4.17	-4.05	-4.02
Finland	408,866	-5.10	-5.02	-4.91
France	2,396,914	-5.04	-4.95	-4.86
Germany	448,259	-5.81	-5.46	-5.36
Hungary	1,138,019	-9.67	-9.50	-9.45
Italy	3,970,425	-5.00	-4.95	-4.87
Poland	211,983	-5.88	-5.75	-5.67
Portugal	2,236,817	-4.36	-4.16	-4.07
Romania	1,756,871	-4.96	-4.87	-4.84
Slovakia	325,799	-4.51	-4.41	-4.35
Slovenia	244,853	-4.26	-4.21	-4.19
Spain	4,827,965	-4.86		-4.61
Sweden	1,169,886	-7.30	-7.29	-7.20
United Kingdom	529,968	-5.72	-5.41	-5.32

Table 9: **Model Comparison: Lévy alpha-stable vs. AEP distributions.** The table shows the average log-likelihood of the Lévy alpha-stable model and the AEP models. The log-likelihood of Lévy alpha-stable model is generally higher than that of the AEP4 and AEP5 models. Note that some countries have a few observations that have a zero probability for the AEP5 density function. The probability is not actually zero but is so close to it and is recorded as zero in the numerical software. We exclude those observations when calculating the log likelihood for the AEP5. The number of observations that are excluded from the country sample is the following: Bulgaria (3), Czech Republic (1), Finland (2), France (8), and Sweden (2). The AEP5 result for Spain is omitted because its estimation fails to be completed within a reasonable amount of time.

## D Derivation of the scaling of the sample standard deviation with sample size

Here we provide a highly stylized, heuristic derivation for the scaling of the sample standard deviation with sample size in the case of a Lévy alpha-stable distributed random variable. The key to this phenomenon is that because the theoretical moment is infinite, the larger the sample size, the higher is the chance that an extreme event is drawn. These extreme events are so extreme that they dominate the sum of squares from which the variance is computed. Thus, the larger the sample size, the larger the sample variance.

Sornette (2006) and Bouchaud & Potters (2003), for instance, provide more precise statements. Here we expose the argument in the simplest, albeit non rigorous way. We discuss the maximum, but symmetric arguments apply to the minimum.

First, we note that for  $N$  large enough, the sample maximum will be dictated by the tail. A key characteristic of the Lévy alpha-stable distribution is that it has power law tails, that is, for large  $x$  (Nolan 2020),

$$P(X > x) \sim x^{-\alpha}. \quad (8)$$



Now, in a sample of size  $N$ , we would hardly expect to see an extreme value that has chances of occurring less than  $1/N$ . Thus, we may define the “typical” value of the maximum as the value  $X_{\max}$  such that  $1/N = P(X > X_{\max})$ . Using Eq. 8, we have<sup>26</sup>  $1/N \sim X_{\max}^{-\alpha}$ , and solving for  $X_{\max}$  gives

$$X_{\max} \sim N^{\frac{1}{\alpha}}, \quad (9)$$

For simplicity, let us assume a mean of zero<sup>27</sup>, so that the sample variance is just the average squared value,

$$\text{Var}(X) \sim \frac{1}{N} \sum_{i=1}^N X_i^2. \quad (10)$$

In a Lévy alpha-stable distribution, the square of the maximum (or minimum, if larger in absolute value) is so large that it dominates the entire sum of squares, such that we may approximate

$$\sum_{i=1}^N X_i^2 \approx X_{\max}^2 \sim N^{\frac{2}{\alpha}}, \quad (11)$$

where the last step uses Eq. 9. Now, inserting Eq. 11 into 10, and taking square root, we find that the standard deviation depends on the sample size as

$$\sqrt{\text{Var}(X)} \sim N^{\frac{1}{\alpha} - \frac{1}{2}}.$$

Note that when  $\alpha = 2$ , so that the distribution is Gaussian, the sample standard deviation does *not* increase with sample size, as one expects.

See also the derivations leading to (4.52) in [Sornette \(2006\)](#) and the heuristic proof based on the rank-size rule in [Gabaix \(2009, Proof of Prop. 2\)](#). A rigorous proof ([Gabaix 2011, Cohen et al. 2020](#)) simply uses the fact that since  $X_i$  is (by assumption) regularly varying with index  $\alpha$ , it follows that  $X_i^2$  is regularly varying with index  $\alpha/2$  ([Gabaix 2009, e.g. Eq.8](#)), and we can apply the GCLT to the sum of  $X_i^2$  to obtain the right scaling. The rigorous proof makes it clear that we expect the sample standard deviation to exhibit substantial variations, since the convergence is in distribution to a Lévy stable variable.

---

<sup>26</sup> [Sornette \(2006\)](#) shows that this is the value of the maximum that is not exceeded with probability  $1/e \approx 37\%$ . Generalizing to the value of the maximum that is not exceeded with probability  $p$  implies the condition  $\frac{\ln(1/p)}{N} = P(X > X_{\max})$ , which does not change the scaling of  $X_{\max}$  with  $N$ . [Newman \(2005\)](#) shows that the expected value of the maximum also scales as  $N^{\frac{1}{\alpha}}$ .

<sup>27</sup> This is without loss of generality, see (2.8) in [Cohen et al. \(2020\)](#).

## E The scaling of the estimated sample standard deviation with sample size

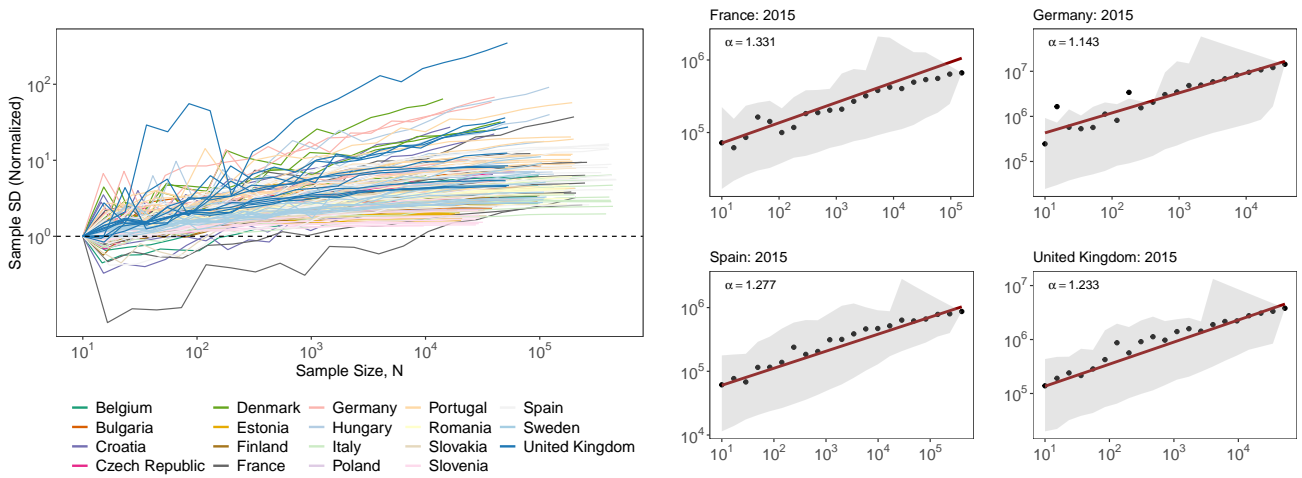


Figure 16: **Measured standard deviation in sub-samples of firm labour productivity, country-year sample.** We construct a distribution by pooling together the productivity levels of firms in each of 19 countries for each year. Then, for each subsample size  $N$ , we compute the standard deviation of each of 1,000 subsamples. The plot on the left shows the average of standard deviations (dots) for each of all country-year samples, while the plot on the right shows the same scaling for four selective countries, France, Germany, Spain and the UK, for Year 2015 with the linear scaling line calculated by  $N^{\frac{1}{\alpha} - \frac{1}{2}}$  from Appendix D (we use our estimate of the tail parameter  $\alpha$ ). The shaded area is the 5<sup>th</sup> and 95<sup>th</sup> percentiles of the sample standard deviation.

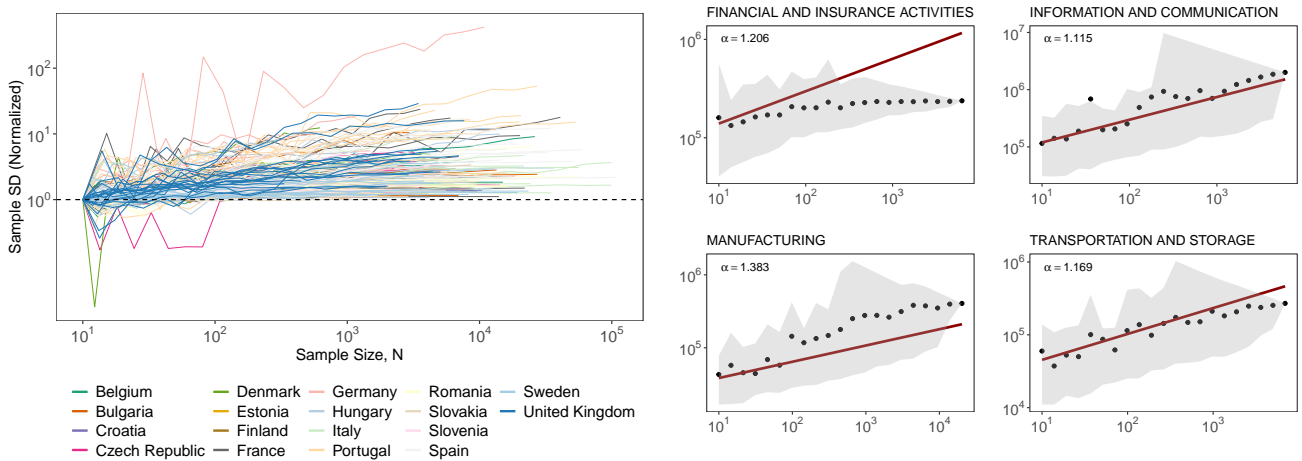


Figure 17: **Measured standard deviation in sub-samples of firm labour productivity, country-industry sample for year 2015.** We construct a distribution by pooling together the productivity levels of firms in each of 19 countries for each industry. Then, for each subsample size  $N$ , we compute the standard deviation of each of 1,000 subsamples. The plot on the left shows the average of standard deviations (dots) for each of all country-industry samples at a 2-digit industry level for year 2015, while the plot on the right shows the same scaling for four selective industries for France, Year 2015. Similar pattern holds for all other years.

## F Trapani's (2016) procedure for testing for infinite moments

Trapani (2016) suggests a test for the divergence of arbitrary moments of order  $p$ , including fractional (non-integer) moments. As the moment of order  $p$  may be infinite, it is unknown whether a limiting distribution of the moments exists. The test can therefore not be applied directly. Following a randomized testing approach, the test therefore adds artificial randomness to manipulate the quantity in question, the sample moment of order  $p$ , to yield a known distribution if the moment is infinite. The resulting test statistic can then be compared to this distribution to obtain a p-value for whether or not the null hypothesis, that the moment of order  $p$  is infinite, is correct.

### F.1 Test procedure

The test statistic is derived such that under the null hypothesis  $H_0$ , the  $p$ th moment of the distribution is infinite. For this, the approach starts with the absolute sample moment of order  $p$ ,

$$A_p = \frac{1}{n} \sum_{k=1}^n |X_k|^p.$$

To make the resulting test statistic scale-invariant and comparable, the absolute moment must be rescaled,

$$A_p^* = \frac{A_p}{(A_\psi)^{p/\psi}} \times \frac{(A_\psi^N)^{p/\psi}}{A_p^N},$$

with  $\psi \in (0, p)$ .  $A_p^N$  denotes the  $p$ th absolute fractional lower order moment of the standard normal distribution,  $A_\psi^N$  the  $\psi$ th absolute fractional lower order moment of the standard normal, etc. Next, an artificial random sample  $\xi$  of size  $r$ <sup>28</sup> is generated from a standard normal distribution and rescaled,

$$\varphi_r = \sqrt{e^{A_p^*}} \times \xi_r.$$

The intuition here is that  $\varphi_r$  follows a normal distribution with mean zero and a finite variance, as  $n \rightarrow \infty$ , if  $A_p^*$  is finite itself. Thus the problem has been reduced from testing for any moment  $p$ , to testing the existence of the variance of the transformed random variable  $\varphi_r$ .

The next step is to generate a sequence  $\zeta_r$ , given by

$$\zeta_r(u) \equiv I[\varphi_r \leq u],$$

where  $I[\cdot]$  is an indicator function, and  $u \neq 0$  is any real number. Under  $H_0$ ,  $\zeta_r(u)$  will have a Bernoulli distribution with mean  $\frac{1}{2}$  and variance  $\frac{1}{4}$ . This is not the case under the alternative, where  $A_p^* < \infty$ , as  $e^{A_p^*}$  converges to a finite value.

Values for  $u$  are picked from some density, but for simplicity it can be taken from a uniform distri-

---

<sup>28</sup>Trapani suggests  $r = n^{0.8}$  where  $n$  is the number of observations used to compute  $A_p$ .

bution  $U(-1, 1)$ . As a result, the test statistic of interest is obtained as

$$\Theta_r \equiv \int_{-1}^1 \frac{1}{2} \vartheta_r^2(u) du,$$

$$\vartheta_r(u) \equiv \frac{2}{\sqrt{r}} \sum_{j=1}^r \left[ \zeta_j(u) - \frac{1}{2} \right].$$

Under the null hypothesis that moment  $p$  is infinite,  $\vartheta_r(u)$  should reduce to zero given that  $\zeta_j(u)$  has a mean of two.  $\Theta_r$  is thus shown to follow a  $\chi^2$  distribution with  $df = 1$  if moment  $p$  is infinite.

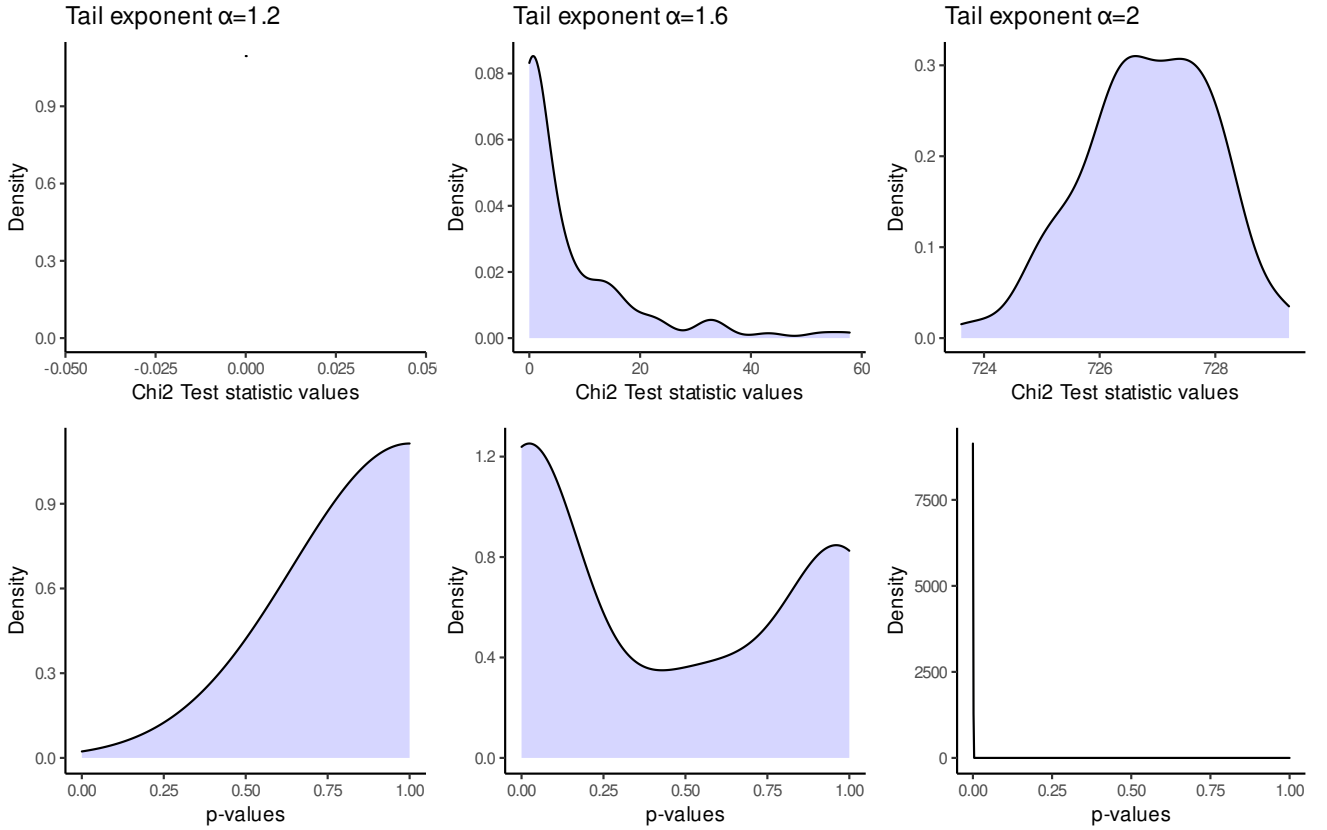


Figure 18: Performance test for Trapani's finite moment test conducted using the R package finity. Distribution of  $\chi^2$  values and p-values given by the test for the divergence of the second moment ( $k = 2$ ) for samples of size  $N = 100,000$  drawn from a Lévy alpha-stable distribution  $S(\alpha, 0.5, 1.0, 0.0; 0)$  for three different values of  $\alpha \in \{1.2, 1.6, 2.0\}$ .

## F.2 Performance of the test

Trapani's test depends on a number of parameters; the choice of  $\psi$ ,  $r$ , and  $u$  can influence the test results. We largely follow Trapani (2016) in our choices of the parameter values and choose

$$u = 1, r = n^{0.8}, \psi = \begin{cases} k - 1 & \text{if } k > 1 \\ k/2 & \text{if } k \leq 1 \end{cases}.$$

The test results should be expected to be somewhat noisy. Especially for samples drawn from a Lévy

alpha-stable distribution  $S(\alpha, \dots)$  in the vicinity of the true tail parameter  $k \approx \alpha$ , the test loses accuracy.

For the tests performed here, we use the R package `finty` (Heinrich & Winkler 2020). Fig. 18 shows the distribution of  $\chi^2$  statistics and p-values for a test of divergence of the second moment ( $k = 2$ ), for samples of size  $N = 100,000$  drawn from a Lévy alpha-stable distribution  $S(\alpha, 0.5, 1.0, 0.0; 0)$  for three different values of  $\alpha$ . The results in Fig. 18 reflect a typical pattern of performance of our implementation of the test at relatively large sample sizes ( $N > 1000$ ), where the test tends to seriously over-reject the null even for values of  $2 > \alpha > 1.6$ , reinforcing the robustness of our conclusion that productivity has an infinite variance.

### F.3 Finite moment test results for country-year sample

Country	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Belgium	1.00	0.30	0.72	0.00	0.00	0.00	0.27	0.08	0.95	0.98	0.99	0.98
Bulgaria		0.98	0.86	0.91	1.00	1.00	1.00	0.99	0.93	0.99	0.92	1.00
Croatia	1.00	1.00	1.00	1.00	0.99	0.99	0.98	0.99	0.92	0.00	0.32	
Czech Republic	0.99	0.97	0.99	1.00	1.00	0.99	0.83	1.00	0.65	0.52	1.00	0.99
Denmark							1.00	1.00	1.00	–	0.99	0.99
Estonia		1.00	0.03	0.01	1.00	0.42	0.00	0.09	0.00	0.00	0.00	0.00
Finland	0.10	0.08	0.63	0.08	0.28	0.59	0.63	0.82	0.92	0.62	0.76	1.00
France	0.08	0.32	1.00	0.99	1.00	1.00	0.99	0.99	0.24	0.99	0.99	1.00
Germany	1.00	0.99	1.00	1.00	1.00	0.99	0.99	1.00	–	–	1.00	
Hungary		1.00	0.98	1.00	0.99	0.99	–	–	0.99	0.99	1.00	0.99
Italy	0.99	0.44	0.93	0.99	0.67	0.51	1.00	0.00	0.00	0.00	0.87	0.00
Poland	1.00	1.00	0.59	0.99	0.99	0.98	0.94	1.00				0.96
Portugal	0.95	0.95	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00	0.99	0.99
Romania		1.00	0.99	0.99	0.90	0.35	1.00	1.00	0.74	0.99	1.00	
Slovakia		0.98	0.99	1.00	0.16	0.37	1.00	0.43	0.78	0.00	0.86	
Slovenia					0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
Spain	0.28	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Sweden	0.00	0.95	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.93	0.00	0.00
United Kingdom	0.98	0.99	0.99	0.99	0.99	0.99	–	0.99	–	1.00	0.99	

Table 10: **Testing for infinite moments of firm-level labour productivity.** P-values of Trapani’s finite moment test for the nonexistence of the second moment of the distribution of labour productivity, for country-year samples. In the vast majority of cases, the p.value is above any sensible threshold (e.g., 1%, 5%, or 10%), suggesting that the nonexistence of the second moment can not be rejected. For example, 84% of the P-values are greater than 0.05, meaning that, for 84% of country-year samples, the finite moment test failed to reject the null hypothesis of the infinite second moment at the 5% confidence level. See Appendix F.

### F.4 Finite moment test results for country-year-industry sample

Table 11 shows the proportion of the 2-digit industries in each country for which we cannot reject that the second moment diverges at the 5% level.

Country	# of Sectors	Proportion	Country	# of Sectors	Proportion
Belgium	228	0.58	Italy	216	0.56
Bulgaria	198	0.72	Poland	162	0.72
Croatia	198	0.33	Portugal	216	0.78
Czech Republic	216	0.67	Romania	190	0.53
Denmark	108	0.44	Slovakia	180	0.67
Estonia	198	0.28	Slovenia	162	0.17
Finland	216	0.28	Spain	228	0.68
France	216	0.50	Sweden	216	0.22
Germany	209	0.83	United Kingdom	209	0.94
Hungary	209	0.61			

Table 11: **Proportion of the country-year-industry samples for which the hypothesis of infinite second moment of the distribution of LP is not rejected at the 5% level.** Across all country-year-industry samples, the proportion is 55%.

Even at a highly disaggregated country-year- industry level, 55% samples do not have the second moment in the distribution of labour productivity. For the largest 5 European countries (France, Germany, Italy, Spain, and the UK), the same result holds for nearly 70% of samples.