Flavio Calvino, Chiara Criscuolo, Carlo Menon and Angelo Secchi

Growth volatility and size: a firm-level study

Item type
Article (Accepted version) (Refereed)

Growth volatility and size: a firm-level study

Angelo Secchi, Flavio Calvino, Chiara Criscuolo, Carlo Menon

PII: S0165-1889(18)30117-9
DOI: 10.1016/j.jedc.2018.04.001
Reference: DYNCON 3585

To appear in: *Journal of Economic Dynamics & Control*

Received date: 2 November 2017
Revised date: 26 February 2018
Accepted date: 2 April 2018

Please cite this article as: Angelo Secchi, Flavio Calvino, Chiara Criscuolo, Carlo Menon, Growth volatility and size: a firm-level study, *Journal of Economic Dynamics & Control* (2018), doi: 10.1016/j.jedc.2018.04.001

Growth volatility and size: a firm-level study*

Flavio Calvino[1], Chiara Criscuolo[1,2], Carlo Menon[1], and Angelo Secchi[3]

[1]OECD Directorate for Science, Technology and Innovation

[2]CEP - London School of Economics and Political Science

[3]Paris School of Economics, Université Paris 1 Panthèon-Sorbonne

April 11, 2018

*The views expressed here are those of the authors and cannot be attributed to the OECD or its member countries. The DynEmp (Dynamics of Employment) project would not have progressed this far without the active participation of all members of the DynEmp network and, in particular, of national delegates (listed in Table A1) from the OECD Working Party on Industry Analysis (WPIA). The authors would like to thank the Editor and two anonymous referees for their valuable suggestions. The authors would also like to thank participants to seminars at the OECD, Scuola Superiore Sant'Anna and Paris School of Economics. Regarding the United Kingdom, this work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. Access to French data benefited from the use of Centre d'accès sécurisé aux données (CASD), which is part of the "Investissements d'Avenir" program (reference: ANR-10-EQPX-17) and supported by a public grant overseen by the French National Research Agency. Angelo Secchi acknowledges financial support from the European Union Horizon 2020 research and innovation programme under grant agreement No. 649186 (ISIGrowth).

Abstract

This paper provides a systematic cross-country investigation of the relation between a firm's growth volatility and its size. For the first time the analysis is carried out using comparable and representative sets of data sourced by official business registers of an important number of countries. We show that there exists a robust negative relation between growth volatility and size with an average elasticity equal to -0.18. We check the robustness of this result against a number of potential sources of bias and in particular with respect to sectoral disaggregation and against the inclusion of firm age. Our result is consistent with the idea that independently from specific country characteristics there exists a common underlying mechanism driving the elasticity between size and growth volatility. We then propose two mechanisms able to explain our result and we conclude discussing its relevance with respect to the recent literature on granularity.

Keywords: Firm size; Gibrat's law; Volatility of growth.

JEL Classification: D22, L25.

1 Introduction

Is the growth dynamics of business firms tied-up with their size? In trying to answer this question the literature has largely focused on the relation between a firm's size and its average growth rate,[1] while only a small number of studies have investigated if there exists a link between a firm's size and the volatility of its growth.[2] After the first evidence in Hymer and Pashigian (1962) recent estimates obtained on U.S. companies support the idea that larger firms tend to display a less volatile growth dynamics than smaller ones (Stanley et al., 1996). These pieces of evidence remain, however, to a large extent inconclusive. First, as suggested in Gabaix (2011), these estimates may well be biased since they are obtained focusing on large listed companies only.[3] Second, as recently discussed in Di Giovanni and Levchenko (2012), the extent to which results for the U.S. economy can be generalised to other countries is unclear.[4]

This paper overcomes both these limitations providing the first systematic cross-country investigation of the relation between a firm's growth volatility and its size, using an original data source containing comparable and representative data on business firms for 20 countries. We show that there exists a robust negative relation between the volatility of growth and size: averaging across countries, an increase by 10% of a firm's size is accompanied by a 1.8% decrease of its growth volatility. This relation appears quite homogeneous across countries, with 17 out of 20 countries characterized by an estimated elasticity lying in the interval $[-0.24, -0.16]$. We check the robustness of our result against a number of potential confounding factors and, in particular, we show that it is not an artifact due to the aggregation of firms belonging to different industrial sectors and that it is not entirely driven by firms' age. Our estimates suggest the striking result that economies that are very different in terms of size, industrial structure and institutional framework show very similar estimated elasticity. This is consistent with the idea that, independently from specific country characteristics, there exists a common underlying mechanism generating the relation between firm size and growth volatility.

Quantifying the elasticity between a firm's size and its growth volatility and assessing the extent to which this relation is common across diverse countries is important for a number of reasons. At the micro level, it can help discriminating among different theories of firm growth that are generally grounded on the assumption that a firm can be seen as an aggregation of several elementary units.

[1] See Lotti et al., 2003 for a review of the literature originating from the pioneering work by Gibrat (1931).

[2] Conceptually, cross-sectional variance (or standard deviation) is a measure of between-firm dispersion of growth rates at a given time while volatility is a measure of within-firm variation of growth rates over time (rolling window). The two concepts are, however, very related. Empirically Davis et al. (2007), using Compustat data show that, while capturing different aspects of business dynamics, the two measures track each other well. Using a different data source Calvino et al. (2016) provides further support to the existence of a positive correlation between volatility and dispersion.

[3] Similarly Capasso and Cefis (2012) discuss the effects of the existence of natural and/or exogenously imposed thresholds in firm size distributions on estimations of the relation between firm size and the variance of firm growth rates.

[4] The elasticity between growth volatility and size has been found close to -0.1 with a sample of French manufacturing firms (Coad, 2008) and practically zero with a sample of Italian manufacturing firms (Bottazzi et al., 2007).

Indeed, the fact that we observe an elasticity not far from -0.18 provides evidence against a simple model where these elementary units display similar size and their growth dynamics are independent. On the contrary, our result can be interpreted alternatively as supporting the existence of some correlation among sub-units (Mansfield, 1962 and Boeri, 1989), of a hierarchical structure among sub-units (Amaral et al., 1997b), or of a fat-tailed distribution of the size of sub-units (Sutton, 2002; Fu et al., 2005; Riccaboni et al., 2008). Bottazzi and Secchi (2006) show that if the probability that a firm diversifies into a new sub-market (i.e., generating a new sub-unit) increases with the number of existing sub-units, the negative relation between growth volatility and size can be traced back to a more fundamental positive correlation between a firm's size and the number of its sub-units.

At the macro level, assessing the existence of the scaling relation between growth volatility and size is important to determine the extent to which micro-level volatility is associated with aggregate fluctuations (see Comin and Mulani, 2006, Comin and Philippon, 2006 and Davis et al., 2007). In granular economies Gabaix (2011) shows that the mechanism which transmits microeconomic shocks into aggregate fluctuations is limited by the extent to which large firms present less volatile growth patterns than smaller ones. In the same vein, Di Giovanni and Levchenko (2012) show that the increase in aggregate volatility due to trade opening is magnified when a firm's volatility scales down with its size. In their model, a scaling elasticity of about -0.17 almost triplicates the contribution of trade to aggregate fluctuations.[5]

This paper is organized as follows. Section 2 describes the data and defines the variables used in the empirical investigation. Section 3 presents the main result together with an extensive set of robustness checks. Section 4 provides an economic interpretation of the coefficient of interest in terms of firm diversification and discusses its relevance for the transmission of micro-economic shocks into aggregate fluctuations. Section 5 concludes.

2 Data

The data used in this study come from a distributed data collection exercise aimed at creating a harmonized cross-country micro-aggregated database sourced from firm-level data collected in national business registers.[6] For example the data sources for France and the U.S. are "Fichier Complet Unifié de SUSE" (FICUS) and Census Bureau's Business Dynamics Statistics (BDS) and Longitudinal Business Database (LBD) respectively, which are both built on administrative data with a quasi-universal coverage. These are the typical data used for studies using firm size such as Garicano et al. (2016) and

[5] Another related stream of research analyses focuses on business cycles, with particular attention to the countercyclical nature of microeconomic volatility (see Decker et al., 2016 and Ilut et al., 2014).

[6] "Micro-aggregated" refers to the fact that the aggregation is much finer than what can be found in more common country-sector-year databases. Other data sources, beyond standard business registers, include social security records, tax records, censuses or other administrative sources. See Calvino et al. (2016) for further details.

Haltiwanger et al. (2013). The high representativeness of the underlying data sources and the large country coverage are two of the key features that make our dataset unique and particularly suitable for the present investigation.

These data are produced within the DynEmp project led by the OECD, with the support of national delegates and national experts of member and non-member economies. The DynEmp project builds upon the distributed micro-data methodology proposed by Bartelsman et al. (2004) for analysing and comparing harmonized firm demographics across countries.[7]

Data produced by the DynEmp routine include the "annual flow datasets" and the "transition matrices". The "flow datasets" contain annual statistics on gross job flows, such as gross job creation and gross job destruction and on several other statistical indicators of unit-level employment growth, such as mean, median, and standard deviation. "Transition matrices", which are used in this paper, summarize instead the growth trajectories of different cohorts of firms – defined according to their age, size, and macro sectors of activity – from year t to year $t+j$, where t takes the values 2001, 2004, and 2007 and j is equal to 3, 5, or 7.[8] The matrices contain a number of statistics, such as the number of units in the cell, median employment at t and at $t+j$, total employment at t and at $t+j$, mean growth rate, average size, and, most importantly, employment growth volatility. These statistics are computed on balanced cohorts of entering and incumbent firms that are observed in a time window of length j without considering exiting firms. Therefore the investigations in the following should be considered conditional on surviving.[9]

The DynEmp database currently includes 20 countries, namely Australia, Austria, Belgium, Brazil, Costa Rica, Denmark, Finland, France, Hungary, Italy, Japan, Luxembourg, the Netherlands, Norway, New Zealand, Portugal, Spain, Sweden, Turkey, the United Kingdom and the United States and covers firms in manufacturing and non-financial business services.[10] Data from most countries cover the 2001-2011 period. A detailed coverage table is provided in Appendix A (see Table A2).

Variables of interest. In this subsection we provide details on how the main variables used in this study are built. In particular, we focus on how the DynEmp routine creates the measure of employment growth volatility, σ, using confidential firm-level data for 20 countries.

Following Davis and Haltiwanger (1999), annual employment growth $R_{i,t}$ of firm i at time t is defined as

$$R_{i,t} = \frac{S_{i,t} - S_{i,t-1}}{0.5(S_{i,t} + S_{i,t-1})} \quad , \tag{1}$$

[7] Details on the data collection and harmonisation procedure are discussed extensively in Criscuolo et al. (2015).

[8] Therefore, if data are available, transition matrices are calculated for the periods 2001-2004, 2001-2006, 2001-2008; 2004-2007, 2004-2009, 2004-2011; 2007-2010, 2007-2012, 2007-2014.

[9] While we will not be able to directly control for a possible selection bias, some of our robustness checks in Section 3 provide evidences mitigating concerns with this respect.

[10] Data for Japan are limited to the manufacturing sector only and Costa Rica is excluded from the sample due to the limited time coverage and unavailability of the transition matrix database.

where $S_{i,t}$ indicates employment of firm i at time t. Next we define firm-level employment growth volatility of firm i as the standard deviation of its employment growth rates over a time window of length j

$$\sigma_{i,t}^{j} = \sqrt{\frac{1}{j-1}\sum_{h=1}^{j}(R_{i,t+h} - \overline{R}_{i,t+1}^{j})^2} \quad , \tag{2}$$

where $\overline{R}_{i,t+1}^{j}$ is the average employment growth rate of firm i over the period between $t+1$ and $t+j$. Firm-level data are then aggregated to avoid confidentiality issues. However, this aggregation is very detailed (we call this a micro-aggregation). Indeed, the DynEmp routine aggregates confidential micro-data in cells on the basis of five dimensions: i) the starting year t, with $t = 2001, 2004$ or 2007; ii) the length of the time window j over which firms are followed, with $j = 3, 5, 7$; iii) firms' age classes a, with a =[entrants, $1-2$ years old, $3-5$ years old, $6-10$ years old, 11 or more years old]; iv) firms' size classes s, with s =[less than 10 employees, $10-49$ employees, $50-99$ employees, $100-249$ employees, $250-499$ employees, 500 or more employees]; v) macro sectors of economic activity m, with m =[manufacturing, non-financial business services].[11] Further details on the methodology and cleaning procedure are presented in Criscuolo et al. (2015).

Accordingly, and in line with the literature (Davis et al., 2007), we define cell-level employment growth volatility $\sigma_{c,t}^{j}$ as the weighted average of firm-level volatilities of firms i in cell c, computed over a time window of length j

$$\sigma_{c,t}^{j} = \sum_{i \in c,t} w_{i,t}^{j}\sigma_{i,t}^{j} \quad , \tag{3}$$

where weights $w_{i,t}^{j}$ are average employment shares of firms i over the period between t and $t+j$ and the cell c is micro-aggregated according to age classes, size classes and macro-sectors, as previously discussed ($c = a, s, m$). The focus of the analysis is on firms surviving until time $t+j$ as for these units a complete window to calculate growth volatility is available.

At the same level of aggregation, for each detailed cell, the DynEmp database provides information on average size $S_{c,t}$, as the average of initial size (measured in terms of employment at time t) of all firms in the cell, defined as[12]

$$S_{c,t} = \frac{\sum_{i \in c,t} S_i}{N_{c,t}} \quad , \tag{4}$$

where N_c is the number of firms in cell c.

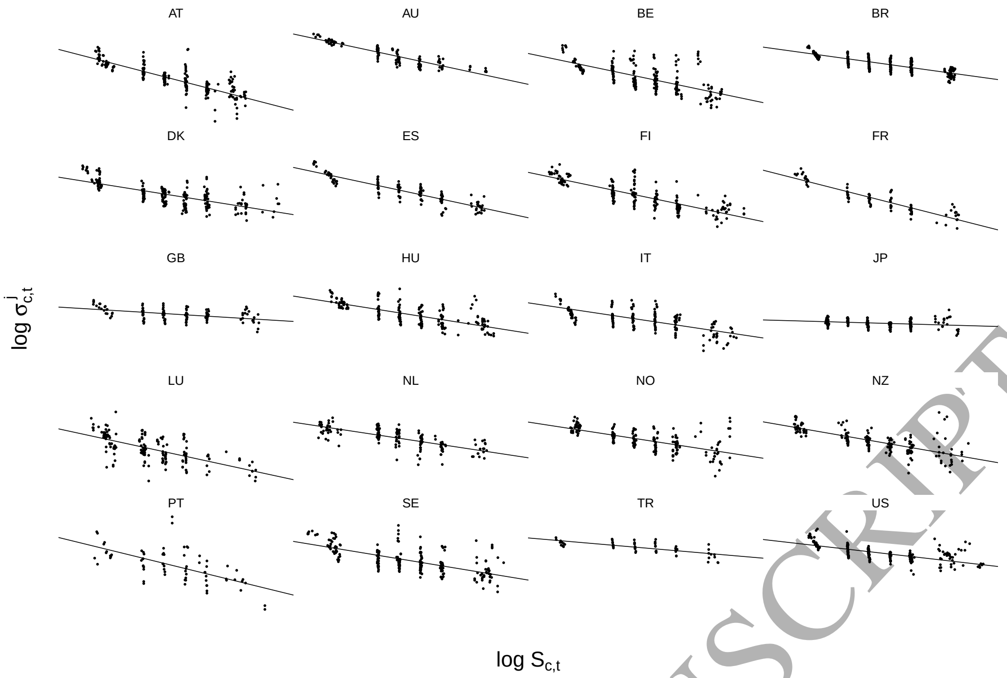

Figure 1: Growth volatility and size - Manufacturing sector

Notes: Scatter plot and linear regression line of log volatility (y axis) and log size (x axis) by country. Manufacturing firms, pooling data over 3, 5, and 7 years time windows and observations from 2001, 2004 and 2007.

3 Empirical results

This work focuses on the analysis of the relation between firms' growth volatility and size. With this aim we estimate the following regression model

$$\log \sigma^j_{c,t} = \alpha + \beta \ \log S_{c,t} + \epsilon_{c,t} \quad , \tag{5}$$

where $\sigma^j_{c,t}$ is the cell-level growth volatility between t and $t+j$, $S_{c,t}$ cell-level average size at initial time t of all firms in cell c and $\epsilon_{c,t}$ is an error term. We estimate equation (5) for each country in our data-set separately. The coefficient of interest, β, is identified mainly from the variation of (log) size across cells and we interpret its value as the value of a conditional correlation that does not reflect causality. The double log transformation implies that growth volatility scales with size according to a power law $\sigma_{c,t} \sim S^{\beta}$, with β measuring the correlation between size and growth volatility in terms of an elasticity.

Figure 1 provides the reader with a simple graphical representation of our data by reporting scatter

[11] The last size class includes firms in the right tail of the size distribution.
[12] Different countries record zero employment units in non-homogeneous ways. This caveat shall be taken into account when interpreting the results in the light of this definition of cell average size.

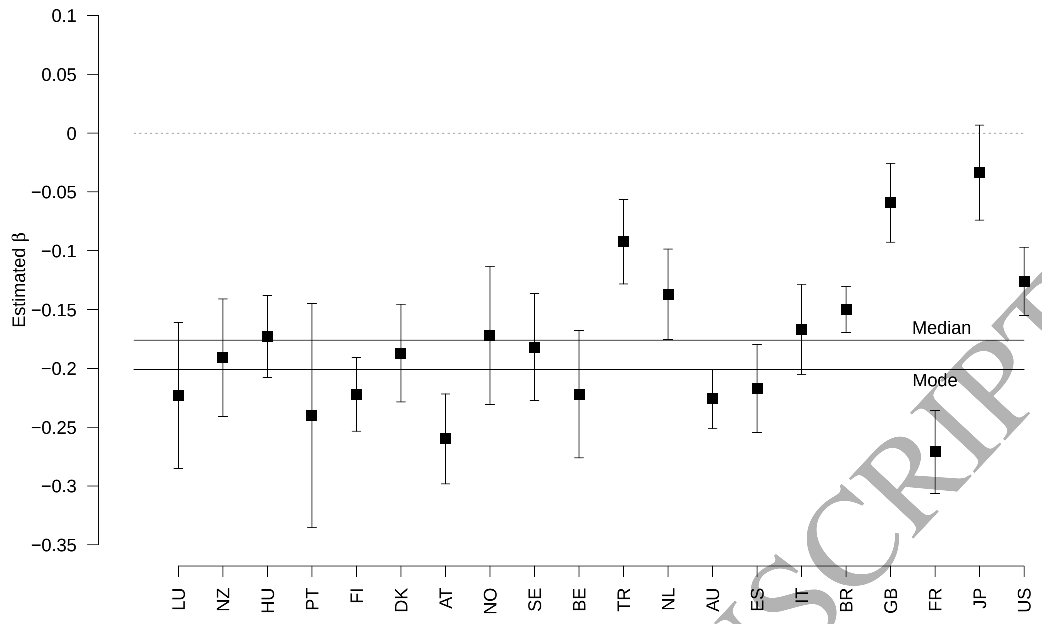

Figure 2: Estimated β - Baseline specification

Notes: Results of the regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Manufacturing firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. Standard errors used to compute the error bars are robust against heteroskedasticity. Countries are ranked based on their GDP in 2010.

plots of (log) firm size, $\log S_{c,t}$ against (log) volatility of growth rates, $\log \sigma^j_{c,t}$. Note that these scatter plots include all observations in different years t, over different time horizons j and with different age without distinguishing them. This is done to avoid the potential disclosure of any confidential information while in the regression analysis we control for all these factors. With this caveat in mind a simple visual inspection of Figure 1 suggests the existence of a negative, almost linear, relation between the two variables in every country, a bit flatter for Japan.

In order to provide a quantitative and statistically robust assessment of this relation we estimate equation (5) with OLS. In the baseline specification we focus on manufacturing firms and on a time window of length equal to three years ($j = 3$) and we estimate the model separately for each country pooling cells corresponding to all size classes, age classes and available years ($t = 2001, 2004, 2007$ conditional on availability).[13] Results from these estimations are displayed in Figure 2 and reported in Table B1 in Appendix B. They deserve a few comments.

First, they robustly confirm a negative and significant relation between volatility and average size in almost all the countries considered. Second, the estimated β appear quite similar across countries:

[13]Note that this excludes from the sample cells that have zero volatility. We will return to this issue in the following.

Table 1: Regression using firm level data from France

	(1)	(2)	(3)	(4)
Sample	All	10+	All	10+
Aggregation	Cell-level	Cell-level	Firm-level	Firm-level
log size	-0.271***	-0.232***	-0.260***	-0.295***
	(0.018)	(0.0249)	(0.012)	(0.004)
constant	-0.884***	-1.074***	-1.344***	-1.410***
	(0.0672)	(0.105)	(0.041)	(0.014)
Obs.	60	50	173,120	64,543
Adj. R^2	0.821	0.665		
R^2			0.205	0.095

Notes: Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. In column (1) we use all observations at the cell level, in column (2) we use observations at cell level focusing only on cells including firms with 10 or more employees, in column (3) we use observations at the firm level weighted by their relative size and in column (4) observations at the firm level focusing only on those with 10 or more employees. In all 4 columns we consider manufacturing firms over a 3 years time window and pooling together observations from 2001 and 2004. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.

most elasticities in the manufacturing sector (for 17 out of 20 countries) lie with a 95% significance level in the interval $[-0.24, -0.16]$, with their mean and median values equal to -0.18. Since β is an elasticity, this means that if a firm's size increases by 10% the volatility of its growth tends to decrease by 1.8%. A slightly different specification on a pooled sample including country dummies provides a very consistent scaling coefficient equal to -0.18, significant at 1% level when standard errors are clustered at country-year level.[14] Third, when looking at Figure 2, where countries are ranked by GDP, we do not observe any clear relation between the latter and the estimated β. The lack of this relation is confirmed by mean of a Least Absolute Deviation regression between β and the (log) GPD which returns an insignificant coefficient. Similarly we do not observe any relation between the estimated β and the (log) GDP per capita, the share of employment in services, or policy indicators capturing employment protection or the strength of legal rights of the countries in our sample. The striking implication is that economies that are very different in terms of size, industrial structure and institutional framework show very similar estimated β, suggesting the existence of an underlying mechanism possibly common across them and independent from specific country characteristics. Finally, for 4 countries (Japan, Great Britain, France and Turkey) and to a less extent for the US and Brasil we observe second order deviations from the benchmark of $\beta = -0.18$. This aspect and its implications certainly deserve to be further investigated but this is left for future research.

Validating the result using source data for France. As a first important exercise to validate our result, we compare the estimates presented in Figure 2 with those obtained using firm-level micro-data, in a country for which direct access to the underlying confidential firm-level data source is possible

[14]Results are available upon request.

for the authors, i.e. France. The data source for France is FICUS (Fichier Complet Unifié de SUSE), which is constructed from administrative (fiscal) data with almost universal coverage.[15] As confirmed in Garicano et al. (2016) this is the most appropriate database to study the firm size distribution in France.

In line with what we have done in the previous section, we define $S_{i,t}$ as firm i's size in term of employees at time t and $\sigma^j_{i,t}$ as firm-level growth volatility built over a j-years time window. We focus on manufacturing firms and we pool together observations for 2001 and 2004. Even with these precautions the two datasets are not directly comparable. Indeed while average cell volatility (as defined in Equation 3) includes into the computation firms with zero volatility, this is not the case when we estimate the relation on individual data where these zero volatility firms are dropped by the log transformation. Since firms with zero volatility tend to be micro firms, using individual data would then underestimate β. To deal with this source of bias in the comparison we adopt two strategies. First, we follow the procedure used in the DynEmp routine and we weight individual data using the employment weights previously described, calculated over the moving window on which volatility is computed. Second, we estimate the regression model using exclusively unweighted observations regarding firms that have 10 or more employees.

Results are reported in Table 1. For the sake of comparison, column (1) reports the estimated coefficient for France obtained with micro-aggregated data as reported in Table B1 and column (2) reports the same coefficient focusing on cells that include firms with 10 or more employees. Columns (3) and (4) report results for the regressions on weighted observation and on firms with 10 or more employees, respectively. Estimates are very similar across the 4 different specifications confirming that our micro-aggregated setting is well suited for investigating the volatility-size relation. Moreover the procedure based on micro-aggregated data allows also to preserve information on zero volatility firms that in a simple individual data setting would be lost. Availability of micro-data allows us to further test whether the results are driven by the growth rate definition (see Equation 1), by the measure of volatility, or by the size proxy (employment versus other size measures) chosen in the DynEmp routine. Unreported estimates on the French manufacturing sector suggest that similar results hold when using a definition of employment growth based on log-differences, even with coefficients slightly lower in absolute value both on micro-data and on micro-aggregated data. Using the standard deviation of employment growth instead of volatility[16] and using turnover as a size proxy also result in negative statistically significant coefficients. Additional checks also corroborate these findings when estimating the scaling relationship in a panel framework with firm fixed effects and when using as dependent

[15]FICUS is based on the mandatory reporting of firms' income to the tax authority. It excludes micro-enterprises and enterprises that are subject to *bénéfices agricoles* (tax regime dedicated to the agricultural sector).

[16]This volatility is computed on a pooled dataset with 25 bins with the same number of observations.

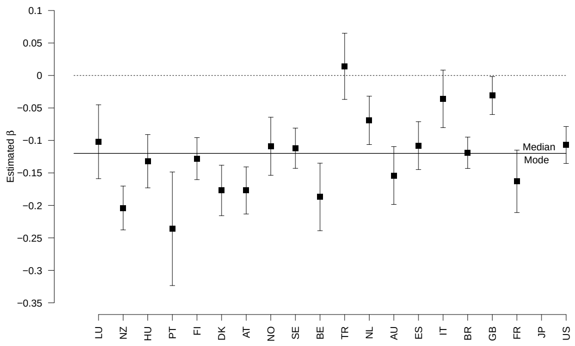

Figure 3: Estimated β - Services sector

Notes: Result of the regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Non financial business services firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. Standard errors used to compute the error bars are robust against heteroskedasticity. Countries ranked based on their GDP in 2010.

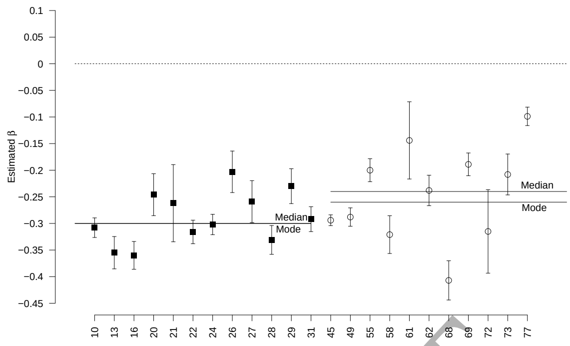

Figure 4: Estimated β - 2-digit sectors (France)

Notes: Result of the regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Firms in manufacturing and services in France only over a 3 years time window and pooling together observations from 2001 and 2004. Standard errors used to compute the error bars are robust against heteroskedasticity. See Table A3 for sector codes legend.

variable a volatility measure computed on growth rates demeaned by common shocks.[17]

Controlling for sectoral composition. So far, the focus has been on the manufacturing sector. However, one might suspect that the observed result is a statistical artifact due to the aggregation of firms operating in different sectors where, in turn, volatility scales down with size following different patterns.

We tackle the issue presenting the estimation of the baseline model for firms operating in non-financial business services. Results are displayed in Figure 3 and reported in Table B5 in Appendix. Two main messages emerge. First, once again the estimated β for almost all countries is negative and statistically significant, the two exceptions being Italy and Turkey.[18] Second the mean and median estimated values are -0.12 and the standard deviation 0.04. With respect to firms in the manufacturing sector, the scaling relation in services tends therefore to be flatter and less dispersed across countries. This is an interesting result since it will be consistent with the economic interpretation of the β coefficient we will discuss in Section 4.

Since considering only two macro sectors (manufacturing and non-financial business services) significantly limits the possibility of observing sectoral specificities we further investigate this issue reverting to the French micro-data. These data allow us to estimate the scaling relationship at a finer level of sec-

[17]These estimates are available upon request.

[18]Notably Turkey is the country for which we have the lowest number of observations due to the limited time period available. Its number of observations is lower than Portugal because no firm reports missing age, and therefore the "missing" age class is not defined in the micro-aggregated data.

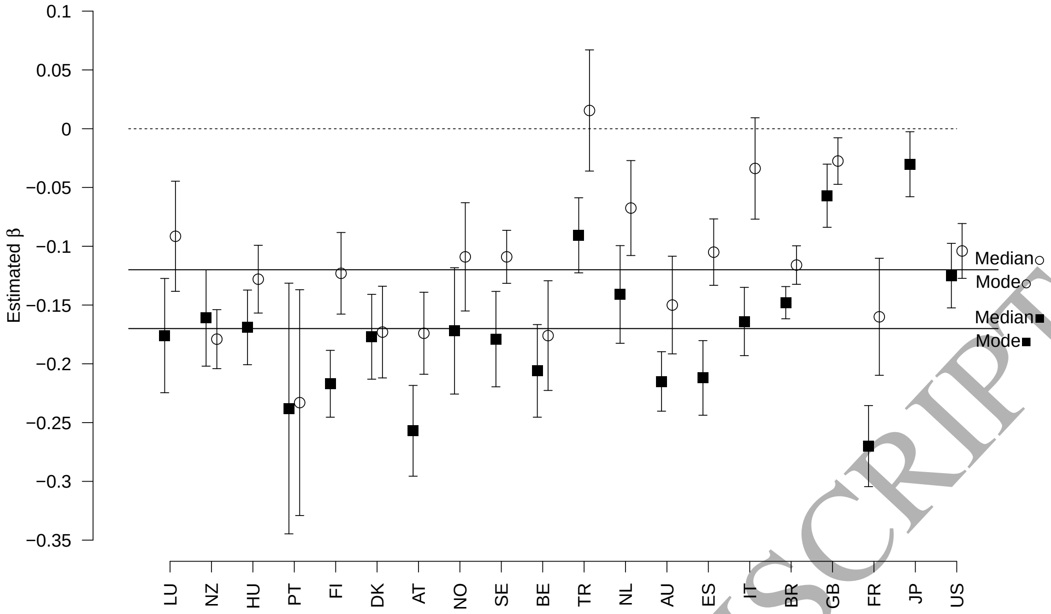

Figure 5: Estimated β - Controlling for age

Notes: Result of the regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$ with a full set of age dummies. Manufacturing (black squares) and Service (green circles) firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. Standard errors used to compute the error bars are robust against heteroskedasticity. Countries ranked based on their GDP in 2010.

toral aggregation.[19] Results, visually displayed in Figure 4, are overall consistent with those reported in Table 1 even if, as expected, they show a certain degree of heterogeneity both in the manufacturing and services sectors.[20] As in the cross-country setting, the scaling relationship tends to be flatter in services. Again, this is going to be consistent with the interpretations we propose below.

Controlling for age. An important improvement with respect to the existing literature (see for example Stanley et al., 1996) is that, in our investigations, we can exclude that the result we obtain is entirely driven by an age effect. With this aim we enrich the baseline model with a set of age dummies. These dummies are based on the age class aggregation described in Section 2.[21]

Results for both manufacturing and service firms are displayed together in Figure 5 and reported in Tables B6 and B7 in Appendix. When including age dummies, while older firms tend to be less

[19]Namely using the OECD STAN A38 classification in 38 sectors, focusing on manufacturing and non-financial market services, excluding the Coke and refined petroleum industry.

[20]The estimated β ranges between -0.36 and -0.20 in manufacturing and -0.41 and -0.10 in services. The full set of these results are available upon request.

[21]In these estimates the baseline age category is set to entering firms and $j = 3$. Qualitatively similar results also hold when changing the baseline of age (from the first to the last category) and when changing the length of the time window j.

volatile than younger ones in most cases, estimates of the β coefficients in the manufacturing and non-financial business services sectors remain consistent with the baseline specification. The mean and median values of the coefficient estimates are both equal to -0.17 in the manufacturing sector and to -0.12 in non-financial business services. This confirms that the scaling relation robustly holds also when controlling for age and that the estimated β tend to be flatter in services.[22]

As a further check we also estimated a more flexible specification that includes age class dummies and interactions of age class dummies with average size (see Table C4 in the Appendix). In this case, for firms in the Manufacturing sector, the estimated mean and median of the β coefficients is equal to -0.20 and -0.21, respectively, even though again older firms tend to be less volatile than younger ones in their growth dynamics. The same robust patterns emerge for firms in Services.

Other robustness checks. We further test the robustness of our main finding along a number of dimensions. First we run a set of basic checks by extending j (the length of the time window over which volatility is computed) from 3 to 5 and 7 years, conditional on availability, and by including in the baseline regression a set of year dummies to control for common macroeconomic factors. Results for these regressions are reported in Appendix (Table B2, B3 and B4) and they all show that our result emerges as very stable with only a minor reduction (in absolute value) of the estimated coefficient. The median estimated β is -0.17 and -0.16 when j is equal to 5 and j is equal to 7 respectively and -0.18 when we include year fixed effects.[23]

Second, we examine the robustness of our finding by estimating Equation (5) using a technique more robust than OLS to the presence of extreme observations to be sure that they are not driving our result. The first column of Table 2 reports the results when Equation 5 is estimated using a Least Absolute Deviations approach (see also Table C1 in Appendix for further details). Again findings are in line with the baseline result with only minor changes in the coefficients. Estimates report a cross-country mean and median value of -0.18 and -0.20.

Third, we adopt a fully non-parametric approach to test whether the estimates are somehow driven by the particular functional form estimated. We follow the approach proposed by Li and Racine (2004) and report the results in the second column of Table 2 (see also Table C2 and Figure C1 in Appendix).[24] Estimates are qualitatively similar to the main result, with some coefficients (including Belgium, Spain and Sweden) that have a tendency to decrease in absolute value. The cross-country mean is -0.16 and the median -0.17.

Then, we adopt a grouped data approach to regression to further test whether the estimates are

[22]This is also consistent with the results found by Garda and Ziemann (2014) based on Orbis data.

[23]If selection bias was very severe in our data-set we should have observed apparent changes in the estimated β when we extend, from 3 to 5 and 7 years, the time horizon over which we compute size and volatility. This is not the case as most estimated coefficients do not seem statistically different in case of different j.

[24]For the cross-validate bandwidth selection we used in most cases the method described in Hurvich et al. (1998).

Table 2: Robustness: LAD - Non-parametric - Weighted

	LAD				NP					WEIGHTED			
	log size	s.e.	Obs.	R^2	Mean	Mode	av. s.e.	Obs.	R^2	log size	s.e.	Obs.	Adj. R^2
AT	-0.241***	(0.0164)	89	0.487	-0.260***	-0.260	(0.0000)	89	0.6662	-0.269***	(0.0140)	89	0.872
AU	-0.234***	(0.0199)	49	0.612	-0.240***	-0.228	(0.0366)	49	0.8693	-0.258***	(0.0160)	49	0.915
BE	-0.236***	(0.0307)	89	0.334	-0.182***	-0.170	(0.0662)	89	0.4069	-0.275***	(0.0243)	89	0.736
BR	-0.153***	(0.0169)	102	0.392	-0.112***	-0.136	(0.0268)	102	0.6684	-0.190***	(0.0127)	102	0.727
DK	-0.189***	(0.0287)	84	0.279	-0.199*	-0.124	(0.0510)	84	0.6206	-0.234***	(0.0262)	84	0.616
ES	-0.229***	(0.0237)	59	0.474	-0.161***	-0.192	(0.0493)	59	0.7261	-0.271***	(0.0281)	59	0.708
FI	-0.208***	(0.0224)	85	0.442	-0.222***	-0.217	(0.0113)	85	0.6438	-0.261***	(0.0277)	85	0.635
FR	-0.301***	(0.0201)	60	0.592	-0.269***	-0.268	(0.0252)	60	0.8419	-0.331***	(0.0255)	60	0.832
GB	-0.0609***	(0.0212)	60	0.105	-0.056***	-0.0396	(0.0100)	60	0.2231	-0.138***	(0.0288)	60	0.454
HU	-0.167***	(0.0202)	90	0.342	-0.137***	-0.172	(0.0437)	90	0.5243	-0.225***	(0.0154)	90	0.749
IT	-0.163***	(0.0229)	90	0.229	-0.078***	-0.0952	(0.0154)	90	0.4754	-0.175***	(0.0287)	90	0.329
JP	-0.0443**	(0.0190)	58	0.0819	-0.0261***	-0.046	(0.0259)	58	0.1692	-0.0549***	(0.0164)	58	0.270
LU	-0.263***	(0.0476)	62	0.242	-0.223***	-0.223	(0.0000)	62	0.395	-0.260***	(0.0272)	62	0.635
NL	-0.129***	(0.0279)	63	0.220	-0.148***	-0.184	(0.0298)	63	0.4036	-0.110***	(0.0317)	63	0.167
NO	-0.211***	(0.0274)	72	0.352	-0.175***	-0.180	(0.0114)	72	0.4492	-0.207***	(0.0254)	72	0.670
NZ	-0.205***	(0.0226)	80	0.366	-0.191***	-0.191	(0.0000)	80	0.4983	-0.228***	(0.0193)	80	0.586
PT	-0.203***	(0.0637)	36	0.227	-0.238***	-0.237	(0.0094)	36	0.3615	-0.321***	(0.0462)	36	0.687
SE	-0.186***	(0.0211)	90	0.294	-0.0884***	-0.0809	(0.0179)	90	0.4594	-0.235***	(0.0360)	90	0.440
TR	-0.0801***	(0.0260)	30	0.238	-0.102***	-0.103	(0.0173)	30	0.5919	-0.0763***	(0.0185)	30	0.367
US	-0.140***	(0.0141)	98	0.311	-0.107***	-0.147	(0.0332)	98	0.5404	-0.177***	(0.0175)	98	0.697
Pooled	-0.180***	(0.00963)	1,446	0.411	-	-	-	-	-	-0.190***	(0.00912)	1,446	0.789
Mean	-0.1822	-	-	-	-0.1608	-	-	-	-	-0.2148	-	-	-
Median	-0.1960	-	-	-	-0.1683	-	-	-	-	-0.2310	-	-	-

Notes: i) Least Absolute Deviations regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Manufacturing sector.; ii) Non-parametric local linear regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Average gradient, mode of the gradient estimates and average standard errors of the gradient estimates are reported. Significance of the average gradient is assessed via kernel regression significance test; iii) Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$ with observation weighted by the number of firms in each cell. Manufacturing sector. All regressions include a constant. Volatility is calculated over a 3 years time window and pooling together observations from 2001, 2004 and 2007. Robust standard error in parenthesis with *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

somehow driven by the micro-aggregated setting. In particular, Angrist (1998) and Angrist and Pischke (2008) suggest that a regression where individual data are averaged by group and weighted by the number of individuals in each group produces coefficients identical to those generated using original individual level observations (Angrist, 1998; Angrist and Pischke, 2008). In our case, however, cell averages are themselves employment weighted averages and the regression is estimated on a logarithmic transformation of the variables, therefore the coefficients will not be perfectly equal to those we would obtain with individual data. Still, since the number of firms in each cell is an information available in our data set, we re-estimate the main model weighting the observations using this number.[25] The estimates, reported in the third column of Table 2, show that there are no radical changes with respect to the baseline (see also Table C3 in Appendix for further details). Indeed, even if there is a tendency for the coefficients to increase in absolute value, the mean and the median estimated β remains -0.21 and -0.23, respectively which are very close to the original values obtained above. The highest changes occur for Portugal, where the coefficient becomes -0.32, and for the United Kingdom, where it becomes equal to about -0.14.

An extensive number of additional robustness checks on the micro-aggregated data have been also carried out. These include estimating the main equation in a quantile regression framework at different points of the conditional volatility distribution (p25 and p75), excluding the smallest size class from the estimation, and including cell-level average employment growth as an additional control, to make sure that the relationship between growth volatility and size is not mechanical. These additional checks corroborate the main results, confirming the stability of the scaling between growth volatility and size.[26]

4 Discussion

In this section we present a simple statistical framework in which the scaling relation between a firm's size and the volatility of its growth, $\sigma(R) \sim S^{\beta}$, emerges naturally and we discuss possible mechanisms able to justify a value of β close to -0.18. Then we further motivate the relevance of the findings presented in this paper, particularly focusing on the relation between the scaling and the magnitude of aggregate fluctuations.

Why is $\boldsymbol{\beta = -0.18}$? To anwer this question we consider a stylized framework in which a firm of initial size S_0 is composed by N different sub-units. In this context, each sub-unit represents a specific market in which the firm operates. S_0 can be written as $\sum_{u=1}^{N} k_u$ where k_u represent the

[25]Note however that the standard errors from this regression do not measure the asymptotic sampling variance of the slope estimate in the micro-data (see Angrist and Pischke, 2008 for further discussion).

[26]Not observing a significant change in the estimated β when we remove the cell corresponding to the smallest firms provides further evidence on the limited impact of selection bias on our result. Indeed statistics computed in this cell are those more likely to be affected by exiting firms. All these results are available upon request from the authors.

initial size of the sub-unit u. Let us assume that the rate of change of the size of each sub-unit is given by $\frac{\Delta k_u}{k_u} = \sigma_k \epsilon_u$, where ϵ_u are random variables with 0 mean and unit variance, and σ_k^2 represents the variance of growth assumed common to all sub-units. Let instead $R = \frac{S_1 - S_0}{S_0}$ be the firm growth rate.

Within this framework, we propose two explanations of why β is higher than -0.5. The first one is based on the idea that there might exist a positive correlation among growth shocks ϵ_u, while the second one is grounded in the possibility that the size of the sub-unit k_u is correlated with firm size.[27]

In developing our reasoning let us consider as a benchmark the case in which $k_u = k \ \forall u$ and where the growth shocks ϵ_u are uncorrelated. Under these assumptions, a firm's size is proportional to the number of its sub-units N, $S_0 = Nk$, and the standard deviation of R for a firm with initial size S_0 reads

$$\sigma(R|S_0) = \sqrt{\operatorname{var}\left(\frac{1}{S_0}\sum_{u=1}^{N}\Delta k_u\right)} = \sqrt{\operatorname{var}\left(\sum_{u=1}^{N}\frac{1}{N}\sigma\epsilon_u\right)} = \frac{\sigma}{\sqrt{N}} \sim S_0^{-0.5} \quad . \tag{6}$$

Equation (6) implies that β is equal to -0.5 and that as S grows the associated $\sigma(R|S_0)$ scales down proportionally with the inverse of $\sqrt{N}$. This means that, while growing, the firm benefits from a perfect diversification effect and that ultimately it can be seen as a simple agglomeration of small sub-units subject to independent shocks. Unfortunately, as discussed extensively throughout the paper, the observed β is in absolute value lower than 0.5, so this benchmark case is inconsistent with the available empirical evidence. In the following we propose two different modifications of the benchmark that are able to explain why β is approximately equal to -0.18.

In the first extension, we keep the proportionality between S_0 and N but we allow the shocks ϵ_u to be correlated. We assume that for $u \neq v$, $\rho_{uv} = N^{-\rho}$, where ρ_{uv} represents the correlation coefficient between ϵ_u and ϵ_v and ρ a positive parameter in the interval $[0, 1]$. This functional form implies that when $\rho = 0$ the shocks are perfectly correlated while when $\rho = 1$ the correlation between any two shocks scales down proportionally with N. With $0 < \rho < 1$ the correlation ρ_{uv} features a slower than proportional decay to 0. Under these assumptions, $\sigma(R|S_0)$ in (6) becomes

$$\sigma(R|S_0) = \frac{\sigma}{\sqrt{N}}\sqrt{1 + (N-1)N^{-\rho}} \sim S_0^{-\frac{\rho}{2}} \quad , \tag{7}$$

implying that $\beta = -\frac{\rho}{2}$. In this case, the value of β is not constant anymore but rather depends on the strength of the correlation between shocks. When $\rho = 1$, β is equal to -0.5. In this case the decay of ρ_{uv} is so fast that we recover the situation in which $\sigma(R|S_0)$ scales with the inverse of the $\sqrt{N}$ (perfect

[27] These are not the only two possibilities to explain why β is higher than -0.5. Amaral et al. (1997a) present, for example, a tree-like hierarchical model of the internal organization of the firm in which lower layers implement only imperfectly decisions made higher up in the hierarchy. They show that in this setting the value of β depends on the number of layers and on the degree of imperfection in executing orders. Given the aim of the paper, we focus on simpler explanations that do not require assumptions on the internal organization of firms.

diversification effect). When $\rho = 0$, and hence also $\beta = 0$, the growth shocks are perfectly correlated and $\sigma(R|S_0)$ is independent of S_0. In this case, when S increases there are no benefits associated with diversification and the firm can be seen as one single big entity.[28] When $0 < \rho < 1$ we have an intermediate situation in which when S grows ρ_{uv} scales down, but not sufficiently fast to fully take advantage of diversification effects. Equation (7) allows us to compute values of the correlation coefficient ρ_{uv} consistent with $\beta = -0.18$ for different N. For example for $N = 2$, $N = 100$ and $N = 1000$ ρ_{uv} must be approximately 0.78, 0.19 and 0.08.[29]

A second possibility to explain why β is greater than -0.5 without imposing any correlation among growth shocks can be obtained by relaxing the assumption that the size of sub-units k_u is independent of firm size, and assuming instead that it increases with S. To operationalize this intuition we assume that $S \sim N^{\frac{1}{\lambda}}$ with $0 < \lambda < 1$, i.e., larger firms tend to have larger sub-units.[30] With $S \sim N^{\frac{1}{\lambda}}$ and $\rho_{uv} = 0$ we obtain $\sigma(R|S_0) \sim S_0^{-\frac{\lambda}{2}}$ and so $\beta = -\frac{\lambda}{2}$. While this mathematical expression looks almost identical to equation (7), the economic interpretation of $\beta = -0.18$ (or equivalently $\lambda = 0.36$) is different: in this case it reflects a situation in which as S increases firms face limits to their diversification capabilities. Conditional on their size if they followed a pure risk-minimization strategy, they would have been composed by a higher number of sub-units. In this case, indeed, it is not the correlation among growth shocks that reduces the speed of the decay of $\sigma(R|S_0)$ but rather the inability of firms to expand their scope of operations beyond a certain limit.[31]

Assessing the extent to which these explanations are alternative or coexist is an empirical question and would require data on firms' sales disaggregated at the product level. This is beyond the scope of the present paper since this kind of data are not available to us. For the interested readers Sutton (2002), Bottazzi and Secchi (2006) and Riccaboni et al. (2008) provide investigations in this direction.[32]

Volatility scaling in a granular economy. The statistical framework we presented above can be reinterpreted, without major changes, to discuss why knowing that $\beta = -0.18$ is relevant in the context of a granular economy.[33] Consider an economy with M firms, each producing a quantity S_i, and where there are no linkages among firms. Similarly to what done above, let us assume that the rate of change of a firm's size is given by $\frac{\Delta S_i}{S_i} = \sigma_{\text{firm}}\epsilon_i$, where ϵ_i are uncorrelated random shocks with 0 mean and a unit variance, and σ_{firm} is the variance of growth common to all firms. The GDP of this economy can be proxied using $GDP = \sum_{i=1}^{M} S_i$ and the volatility of its growth, σ_{GDP}, can be

[28]Notice that ρ_{uv} needs not to be equal to 1 for this argument to remain valid. Any constant value for ρ_{uv} would work.

[29]These figures for ρ_{uv} seem significantly higher than those found in Bottazzi and Secchi (2006) using pharmaceutical data at the product level. However, they are not directly comparable since pharmaceutical products, often based on different chemical entities, are likely to be by construction less correlated than the average product.

[30]In this case indeed the size of sub-units is proportional to $S^{1-\lambda}$.

[31]This interpretation is consistent with the resource-based view of the firm where diversification strategies are limited by what a firm is able to do (Penrose, 1959; Teece et al., 1997).

[32]The two mechanisms described can be used as the core of more structural model like the one of endogenous firm-level risk over the business cycle based on market exposure presented in Decker et al. (2016).

[33]In what follows we draw from Gabaix (2011) which originally developed the granularity argument.

computed as

$$\sigma_{\text{GDP}} = \sqrt{\text{var}\left(\frac{1}{\text{GDP}}\sum_{i=1}^{M}\Delta S_i\right)} = \sqrt{\text{var}\left(\sum_{i=1}^{M}\frac{S_i}{GDP}\sigma_{\text{firm}}\epsilon_i\right)} = \sigma_{\text{firm}}\sqrt{\text{HHI}} \quad , \tag{8}$$

where HHI is the Herfindal index of the economy. Equation (8) provides a link between the idiosyncratic volatility at the firm level and the volatility of GDP growth that measures the magnitude of aggregate fluctuations. The main contribution of Gabaix (2011) is to show that when M increases the behaviour of σ_{GDP} depends on the shape of the firm size distribution of the economy and in particular on the behaviour of its tails. To briefly illustrate this argument let us assume that firm size is distributed according to a power law, that is $\text{P}(S > s) = (S_{min}/s)^{\gamma}$ with $\gamma \geq 1$. This distribution is said to feature *fat tails* and *thin tails* when $1 \leq \gamma < 2$ and $\gamma \geq 2$ respectively.[34]

It can be shown that when the firm size distribution is thin-tailed and M grows large then σ_{GDP} scales down with the inverse of the $\sqrt{M}$. This implies that idiosyncratic shocks hitting individual firms "average out" in the aggregate. On the contrary, when the tails are fat, the economy is said to be granular and σ_{GDP} decays at a slower pace. Idiosyncratic shocks hitting large firms, in this latter case, generate part of the aggregate movements of GDP.[35] What is more relevant for the present paper is that it can be shown that this relation between microeconomic shocks and macroeconomic fluctuations is mediated by the extent to which micro-level volatility scales with size. Indeed, if one assumes that $\sigma_{\text{firm}} \sim S^{\beta}$ with $-0.5 \leq \beta \leq 0$, then

$$\sigma_{\text{GDP}} \propto GDP^{\beta'} \quad , \tag{9}$$

where $\propto$ means that σ_{GDP} and $GDP^{\beta'}$ are proportional and $\beta' = \max\{(1+\beta)/\gamma - 1, -0.5\}$.[36] Equation (9) states that the volatility of GDP growth depends on the size of the economy (GDP), on the extent to which the firm size distribution is fat-tailed (γ) but also on the extent to which micro-level volatility scales with size (β). This result has important consequences.

First, conditional on γ and β, an economy whose GDP is half of that of a second economy is characterized by a GDP growth volatility that is $(0.5)^{\beta'}$ times larger. Moreover, conditional on GDP and on the shape of the firm size distribution γ an economy tends to be more risky, generating more dramatic fluctuations, the higher its β, that is the less "diversified" are its firms. In presence of full diversification ($\beta = -0.5$) even when the firm size distribution is very fat tailed ($\gamma = 1$) any granular effects would disappear. The intuition behind this result is simple. Aggregate volatility in an economy where the firm size distribution is fat tailed is influenced by the shocks hitting the

[34] More precisely when $1 \leq \gamma < 2$ the variance of the distribution does not exist while with $\gamma \geq 2$ it does. In the latter case one can invoke the Central Limit Theorem in the former one cannot.

[35] Gabaix (2011) suggests that the largest 100 firms in the US account for about one third of the variation in output growth.

[36] For a formal exposition of these results see Proposition 2 and Proposition 3 in Gabaix (2011).

largest firms. In a bigger economy the largest firms tend to be larger[37] so that when a firm's growth volatility does not depend on its size we observe an amplification of aggregate fluctuations. This transmission mechanism of microeconomic shocks to aggregate fluctuations is instead limited when large firms present less volatile growth patterns, that is when large firms show at least a certain degree of diversification. When $\beta = -0.5$ large firms are simply collections of independent smaller firms annihilating any granular fluctuation.

Second, equation (9) is also useful to provide a-back-of-the-envelope quantification of the effect of β in shaping the behaviour of σ_{GDP}. Consider an economy with $\gamma = 1$ so that $\beta' = \beta$. If $\beta = -0.5$ (perfect diversification effect) halving the size of an economy will result in an increase of aggregate volatility σ^{GDP} proportional to $\sqrt{2}$, that is around 41%. With $\beta = -0.4$ and $\beta = -0.3$ the increase in the volatility of GDP growth will become 32% and 23% respectively. Even in an economy characterized by a seemingly mild scaling ($\beta = -0.18$) the same halving of the size of the GDP will be accompanied by an increase in the magnitude of aggregate fluctuations of about 13%.

Third, the same kind of mechanisms we discuss here in terms of closed economies have been shown to influence the impact of a trade opening event on aggregate fluctuations. Di Giovanni and Levchenko (2012) show that allowing volatility to decrease in firm size with a slope of about -0.17, very much in line with our estimates, implies an increment of about 30% of macroeconomic volatility when countries experience a trade opening event. This effect is 3 times larger than in their baseline model where there is no scaling.

5 Conclusions

This paper is the first study that, using comparable and representative data from a significant number of diverse countries, provides robust evidence on the existence of a negative relation between a firm's growth volatility and its size. We estimate an average elasticity of -0.18 with a remarkable homogeneity across countries. We check that this result is robust to a number of potential confounding factors showing, in particular, that it holds true when one performs the investigations at the sectoral level and controls for firm age. Our result suggests that economies that are very different in terms of size, industrial structure and institutional framework show very similar estimated β. This is consistent with the idea that independently from specific country characteristics there exists a common underlying mechanism driving the elasticity between size and growth volatility.

We discuss two mechanisms able to explain why β is higher than -0.5 interpreting our result in terms of a correlation among the growth shocks hitting sub-units within a firm and of the relation between the size of sub-units and of the firm itself. In both cases the emergence of a scaling relation

[37]This is a property of the power laws. See Newman (2005) for details.

between size and growth volatility with $\beta = -0.18$ is associated with the inability of businesses to fully "diversify" their structure to protect themselves against unexpected shocks. The importance of our result is that providing a precise estimate of β helps to better characterize the mechanism that in a granular economy translates idiosyncratic shocks at the firm level into aggregate fluctuations and to quantify more precisely how much a trade opening event would increase GDP fluctuations.

Our finding raises questions for future research. First, the result of a flatter elasticity estimated in the services sector seems interesting and stimulates additional investigations on the drivers of the observed relation. Second, estimating the Pareto coefficients of countries' firm-size distributions and appropriately taking into account the scaling relation of growth volatility with firm size would allow to directly link a country's productive structure and its resilience to economic shocks.

References

Amaral, L. A. N., S. V. Buldyrev, S. Havlin, , P. Maass, M. A. Salinger, H. E. Stanley, and M. H. Stanley (1997a): "Scaling behavior in economics: the problem of quantifying company growth," *Physica A*, 1–24.

Amaral, L. A. N., S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, H. E. Stanley, and M. H. Stanley (1997b): "Scaling behavior in economics: Empirical results for company growth," *Journal de Physique I*, 7, 621–633.

Angrist, J. D. (1998): "Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants," *Econometrica*, 66, 249–288.

Angrist, J. D. and J.-S. Pischke (2008): *Mostly harmless econometrics: An empiricist's companion*, Princeton and Oxford: Princeton University Press.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2004): "Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries," World Bank Policy Research Working Paper Series No. 3464, World Bank, Washington, DC.

Boeri, T. (1989): "Does firm size matter?" *Giornale degli Economisti e Annali di Economia*, 48, 477–495.

Bottazzi, G., E. Cefis, G. Dosi, and A. Secchi (2007): "Invariances and diversities in the patterns of industrial evolution: Some evidence from Italian manufacturing industries," *Small Business Economics*, 29, 137–159.

Bottazzi, G. and A. Secchi (2006): "Gibrat's law and diversification," *Industrial and Corporate Change*, 15, 847–875.

Calvino, F., C. Criscuolo, and C. Menon (2016): "No country for young firms? Start-up dynamics and national policies." OECD Science, Technology and Industry Policy Papers No. 29, OECD Publishing, Paris.

Capasso, M. and E. Cefis (2012): "Firm Size and Growth Rate Variance: The Effects of Data Truncation," *Review of Industrial Organization*, 41, 193–205.

Coad, A. (2008): "Firm growth and scaling of growth rate variance in multiplant firms," *Economics Bulletin*, 12, 1–15.

Comin, D. and S. Mulani (2006): "Diverging trends in aggregate and firm volatility," *The Review of Economics and Statistics*, 88, 374–383.

Comin, D. A. and T. Philippon (2006): "The rise in firm-level volatility: Causes and consequences," in *NBER Macroeconomics Annual 2005*, ed. by M. Gertler and K. Rogoff, MIT Press, vol. 20, 167–228.

Criscuolo, C., P. N. Gal, and C. Menon (2015): "dynemp: A routine for distributed microdata analysis of business dynamics," *Stata Journal*, 15, 247–274.

Davis, S. J. and J. Haltiwanger (1999): "Gross job flows," in *Handbook of Labor Economics*, Elsevier, vol. 3, Part B, chap. 41, 2711 – 2805.

Davis, S. J., J. Haltiwanger, R. Jarmin, and J. Miranda (2007): "Volatility and dispersion in business growth rates: Publicly traded versus privately held firms," in *NBER Macroeconomics Annual 2006*, MIT Press, vol. 21, 107–180.

Decker, R. A., P. N. D'Erasmo, and H. Moscoso Boedo (2016): "Market exposure and endogenous firm volatility over the business cycle," *American Economic Journal: Macroeconomics*, 8, 148–198.

Di Giovanni, J. and A. A. Levchenko (2012): "Country size, international trade, and aggregate fluctuations in granular economies," *Journal of Political Economy*, 120, 1083–1132.

Fu, D., F. Pammolli, S. V. Buldyrev, M. Riccaboni, K. Matia, K. Yamasaki, and H. E. Stanley (2005): "The growth of business firms: Theoretical framework and empirical evidence," *Proceedings of the National Academy of Sciences of the United States of America*, 102, 18801–18806.

Gabaix, X. (2011): "The granular origins of aggregate fluctuations," *Econometrica*, 79, 733–772.

Garda, P. and V. Ziemann (2014): "Economic Policies and Microeconomic Stability: A Literature Review and Some Empirics," OECD Economics Department Working Paper No. 1115, OECD Publishing, Paris, France.

Garicano, L., C. Lelarge, and J. Van Reenen (2016): "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, 106, 3439–79.

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

Haltiwanger, J., R. S. Jarmin, and J. Miranda (2013): "Who creates jobs? Small versus large versus young," *Review of Economics and Statistics*, 95, 347–361.

Hurvich, C., J. Simonoff, and C. Tsai (1998): "Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion," *Journal of the Royal Statistical Society B*, 271–293.

Hymer, S. and P. Pashigian (1962): "Firm size and rate of growth," *The Journal of Political Economy*, 70, 556–569.

Ilut, C., M. Kehrig, and M. Schneider (2014): "Slow to hire, quick to fire: Employment dynamics with asymmetric responses to news," NBER Working Paper No. 20473, National Bureau of Economic Research.

Li, Q. and J. S. Racine (2004): "Cross-validated local linear nonparametric regression," *Statistica Sinica*, 485–512.

Lotti, F., E. Santarelli, and M. Vivarelli (2003): "Does Gibrat's Law hold among young, small firms?" *Journal of Evolutionary Economics*, 13, 213–235.

Mansfield, E. (1962): "Entry, Gibrat's law, innovation, and the growth of firms," *The American Economic Review*, 52, 1023–1051.

Newman, M. E. J. (2005): "Power laws, Pareto distributions and Zipf's law," *Contemporary Physics*, 46, 323–351.

Penrose, E. T. (1959): *The Theory of the Growth of the Firm*, Oxford: Basil Blackwell.

Riccaboni, M., F. Pammolli, S. V. Buldyrev, L. Ponta, and H. E. Stanley (2008): "The size variance relationship of business firm growth rates," *Proceedings of the National Academy of Sciences*, 105, 19595–19600.

Stanley, M. H., L. A. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, and H. E. Stanley (1996): "Scaling behaviour in the growth of companies," *Nature*, 379, 804–806.

Sutton, J. (2002): "The variance of firm growth rates: the 'scaling' puzzle," *Physica A: Statistical Mechanics and its Applications*, 312, 577–590.

Teece, D. J., G. Pisano, and A. Shuen (1997): "Dynamic capabilities and strategic management," *Strategic Management Journal*, 18, 509–533.

Appendix A

In this Appendix the contributors to the DynEmp v.2 data collection are reported (Table A1) and the coverage table of the DynEmp v.2 database is presented (Table A2).[38] A correspondence between the industry codes reported and industry names is also provided (Table A3).

Table A1: Contributors to the DynEmp v.2 data collection

Country	National representative	Institution
Australia	Antonio Balaguer, Diane Braskic David Hansell	Department of Industry, Innovation and Science and Australian Bureau of Statistics
Austria	Werner Hoelzl	WIFO Institute (Austrian Institute of Economic Research)
Belgium	Michel Dumont, Chantal Kegels, Hilde Spinnewyn	Federal Planning Bureau
Brazil	Carlos Henrique Leite Corseuil, Gabriel Lopes de Ulyssea	IPEA - Instituto de Pesquisa Econômica Aplicada
Costa Rica	David Bullon Patton and Tayutic Mena	Ministry for Foreign Trade
Denmark	Dorte Hoeg Koch, Morten Skov Poulsen	Ministry for Business and Growth
Finland	Mika Maliranta	The Research Institute of the Finnish Economy (ETLA) and Statistics Finland
France	DynEmp and MultiProd teams	OECD
Hungary	Adrienn Szep Szollosine, Erzsebet Eperjesi Lindnerne, Gabor Katay, Peter Harasztosi	Central Bank of Hungary, Hungarian Central Statistical Office
Italy	Stefano Costa	Italian National Institute of Statistics (ISTAT)
Japan	Kyoji Fukao and Kenta Ikeuchi	Hitotsubashi University and National Institute of Science and Technology Policy
Luxembourg	Leila Peltier - Ben Aoun, Chiara Peroni, Umut Kilinc	STATEC
The Netherlands	Michael Polder	Statistics Netherlands (Centraal Bureau voor de Statistiek)
New Zealand	Lynda Sanderson, Richard Fabling	New Zealand Treasury, Motu Economic and Public Policy Research and Statistics New Zealand
Norway	Arvid Raknerud, Diana-Cristina Iancu	Statistics Norway and Ministry of Trade and Industry
Portugal	Jorge Portugal	Presidencia da Republica
Spain	Valentin Llorente Garcia	Spanish Statistical Office
Sweden	Eva Hagsten	Statistics Sweden
Turkey	Faik Yücel Günaydin	Ministry of Science, Industry, and Technology
United Kingdom	Michael Anyadike-Danes	Aston Business School
United States	Javier Miranda	Center for Economic Studies, US Census Bureau

Notes: Countries included in the dataset used for this paper.

[38]Costa Rica was excluded due to the limited time coverage and unavailability of the transition matrix database. Data for Japan are limited to the manufacturing sector only. Data for Norway are restricted up to 2001 and 2004+3 due to unusual patterns in the data from 2009. Data related to 2004 in the Netherlands are excluded from the sample due to a redesign of the Dutch business register. Data for the United Kingdom in 2001 are excluded from the sample due to censoring issues related to the age class calculation. Data for France are restricted up to 2007 due to a redesign of the French statistical systems of data collection on firms (from FICUS to FARE).

Table A2: Temporal coverage of the DynEmp v.2 database over time

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
AUSTRALIA						■	■	■	■	■	■	■	■	■	■	■		
AUSTRIA					■	■	■	■	■	■	■	■	■	■	■	■	■	
BELGIUM						■	■	■	■	■	■	■	■	■	■	■		
BRASIL	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	
COSTA RICA													▒	▒	■	■	■	
DENMARK					■	■	■	■	■	■	■	■	■	■	■	■	■	
FINLAND						■	■	■	■	■	■	■	■	■	■	■		
FRANCE				■	■	■	■	■	■	■	■	■	▒	▒	▒	▒		
HUNGARY						■	■	■	■	■	■	■	■	■	■	■		
ITALY				■	■	■	■	■	■	■	■	■	■	■	■			
JAPAN			■	■	■	■	■	■	■	■	■	■	■	■				
LUXEMBOURG					■	■	■	■	■	■	■	■	■	■	■	■	■	
THE NETHERLANDS					■	■	■	■	■	■	▒	■	■	■	■	■	■	
NEW ZEALAND						■	■	■	■	■	■	■	■	■	■	■		
NORWAY			■	■	■	■	■	■	■	■	■	■	■	■	▒	▒		
PORTUGAL											■	■	■	■	■	■	■	
SPAIN								■	■	■	■	■	■	■	■	■	■	■
SWEDEN						■	■	■	■	■	■	■	■	■	■	■	■	
TURKEY											■	■	■	■	■	■	■	■
UNITED KINGDOM						▒	▒	■	■	■	■	■	■	■	■	▒		
UNITED STATES			■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

Notes: temporal coverage by country of the database used for the analysis. Years for which annual flow data are available are colored. Analysis based on flow data excludes the first available year, since most job flows statistics require two consecutive periods to be computed. For Costa Rica no transition matrix is available due to the limited time extension of the source data. For Japan data refer to the manufacturing sector only. Gray boxes correspond to years that have been excluded from the analysis due to data issues. Data for some countries are still preliminary.

Source: OECD DynEmp v.2 database.

Table A3: STAN A38 Sector codes

STAN A38 code	Sector name
10	Food and beverages [CA]
13	Textiles and apparel [CB]
16	Wood and paper products [CC]
20	Chemicals [CE]
21	Pharmaceuticals [CF]
22	Rubber and plastics [CG]
24	Metal products [CH]
26	Computer and electronics [CI]
27	Electrical equipment [CJ]
28	Machinery and equipment [CK]
29	Transport equipment [CL]
31	Furniture and other [CM]
45	Wholesale and retail [G]
49	Transportation and storage [H]
55	Hotels and restaurants [I]
58	Media [JA]
61	Telecommunications [JB]
62	IT [JC]
68	Real estate [L]
69	Legal and accounting [MA]
72	Scientific R&D [MB]
73	Marketing and other [MC]
77	Administrative services [N]

Notes: This classification follows the OECD STAN A38 industry classification focusing on manufacturing and non-financial market services, excluding the Coke and refined petroleum sector.

Appendix B

Additional Tables and Figures

In the following we report the full set of estimation results together with a number of robustness checks.

First, in table B1 we report estimates of the main equation focusing on the manufacturing sector with $j = 3$. Second, we estimate the main equation extending the length of the time window over which volatility is computed from 3 to 5 and 7 years and including in the baseline regression a set of year dummies to control for common macroeconomic factors. Results of these regressions are reported in Table B2, B3 and B4, respectively. Third, we report the estimates of the baseline model focusing on firms operating in non-financial business services. Results are reported in Table B5. Fourth, we report the estimates that control for firm age focusing on the manufacturing sector in Table B6 and on non-financial business services in Table B7.

Table B1: Baseline regression model

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
log size	-0.260***	-0.226***	-0.222***	-0.150***	-0.187***	-0.217***	-0.222***	-0.271***	-0.0594***	-0.173***
	(0.0195)	(0.0127)	(0.0276)	(0.00989)	(0.0212)	(0.0191)	(0.0160)	(0.0180)	(0.0170)	(0.0178)
constant	-1.038***	-0.307***	-1.060***	-0.804***	-1.087***	-0.796***	-0.965***	-0.884***	-1.387***	-0.979***
	(0.0647)	(0.0386)	(0.0997)	(0.0406)	(0.0826)	(0.0766)	(0.0643)	(0.0672)	(0.0702)	(0.0691)
Obs	89	49	89	102	84	59	85	60	60	90
Adj. R^2	0.662	0.837	0.391	0.650	0.538	0.724	0.632	0.821	0.177	0.499

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
log size	-0.167***	-0.0336	-0.223***	-0.137***	-0.172***	-0.191***	-0.240***	-0.182***	-0.0924***	-0.126***
	(0.0194)	(0.0206)	(0.0317)	(0.0196)	(0.0300)	(0.0255)	(0.0485)	(0.0232)	(0.0183)	(0.0148)
constant	-1.225***	-1.858***	-1.501***	-1.216***	-1.116***	-1.105***	-0.919***	-1.030***	-0.807***	-0.860***
	(0.0807)	(0.0686)	(0.115)	(0.0789)	(0.0933)	(0.0780)	(0.201)	(0.0920)	(0.0688)	(0.0577)
Obs	90	58	62	63	72	80	36	90	30	98
Adj. R^2	0.449	0.064	0.385	0.349	0.430	0.492	0.339	0.441	0.454	0.464

Notes: Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Manufacturing firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.

Table B2: Main regression - $j = 5$

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
log size	-0.266***	-0.219***	-0.217***	-0.136***	-0.153***	-0.228***	-0.224***	-0.244***	-0.0720***	-0.152***
	(0.0200)	(0.0133)	(0.0331)	(0.00888)	(0.0203)	(0.0139)	(0.0151)	(0.0183)	(0.0243)	(0.0211)
constant	-0.910***	-0.337***	-1.058***	-0.773***	-1.079***	-0.701***	-0.888***	-0.875***	-1.266***	-0.968***
	(0.0673)	(0.0451)	(0.110)	(0.0374)	(0.0754)	(0.0597)	(0.0669)	(0.0766)	(0.0879)	(0.0765)
Obs.	90	47	59	102	82	59	56	30	30	60
Adj. R^2	0.673	0.789	0.438	0.635	0.461	0.797	0.678	0.867	0.289	0.473

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
log size	-0.146***	-0.0328**	-0.214***	-0.169***	-0.172***	-0.170***	-0.265***	-0.162***	-0.0850***	-0.116***
	(0.0197)	(0.0148)	(0.0293)	(0.0136)	(0.0263)	(0.0312)	(0.0413)	(0.0204)	(0.0163)	(0.0155)
constant	-1.209***	-1.772***	-1.411***	-1.075***	-1.051***	-1.091***	-0.858***	-0.983***	-0.795***	-0.887***
	(0.0780)	(0.0515)	(0.105)	(0.0488)	(0.0835)	(0.0894)	(0.170)	(0.0799)	(0.0524)	(0.0596)
Obs.	60	58	60	62	72	52	35	90	30	94
Adj. R^2	0.491	0.112	0.402	0.681	0.479	0.490	0.425	0.461	0.535	0.478

Notes: Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Manufacturing firms only over a 5 years time window and pooling together observations from 2001, 2004 and 2007. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.

Table B3: Main regression - $j = 7$

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
log size	-0.278***	-0.211***	-0.203***	-0.139***	-0.149***	-0.212***	-0.198***	-	-0.0602***	-0.159***
	(0.0181)	(0.0152)	(0.0279)	(0.00982)	(0.0289)	(0.0161)	(0.0155)	-	(0.0177)	(0.0205)
constant	-0.865***	-0.275***	-1.119***	-0.748***	-1.126***	-0.750***	-0.953***	-	-1.274***	-0.925***
	(0.0652)	(0.0444)	(0.0968)	(0.0422)	(0.103)	(0.0696)	(0.0615)	-	(0.0721)	(0.0754)
Obs.	59	23	59	66	55	29	56	-	28	59
Adj. R^2	0.736	0.838	0.488	0.683	0.452	0.830	0.689	-	0.283	0.471

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
log size	-0.132***	-0.0168	-0.233***	-0.165***	-0.105**	-0.168***	-	-0.161***	-	-0.109***
	(0.0239)	(0.0200)	(0.0307)	(0.0219)	(0.0488)	(0.0291)	-	(0.0281)	-	(0.0215)
constant	-1.218***	-1.789***	-1.275***	-0.946***	-1.164***	-1.121***	-	-0.942***	-	-0.908***
	(0.0971)	(0.0708)	(0.120)	(0.0638)	(0.160)	(0.0827)	-	(0.107)	-	(0.0817)
Obs	30	30	39	29	36	51	-	60	-	61
Adj. R^2	0.489	0.002	0.513	0.641	0.214	0.563	-	0.377	-	0.447

Notes: Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Manufacturing firms only over a 7 years time window and pooling together observations from 2001, 2004 and 2007. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.

Table B4: Main regression - Year dummies

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
log size	-0.260***	-0.226***	-0.222***	-0.150***	-0.183***	-0.217***	-0.223***	-0.271***	-0.0594***	-0.172***
	(0.0189)	(0.0129)	(0.0278)	(0.0100)	(0.0218)	(0.0194)	(0.0158)	(0.0182)	(0.0171)	(0.0172)
year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
constant	-1.059***	-0.306***	-1.017***	-0.801***	-1.175***	-0.820***	-1.055***	-0.865***	-1.400***	-1.079***
	(0.0645)	(0.0479)	(0.120)	(0.0498)	(0.0912)	(0.0846)	(0.0883)	(0.0730)	(0.0688)	(0.0897)
Obs.	89	49	89	102	84	59	85	60	60	90
Adj. R^2	0.678	0.834	0.379	0.643	0.637	0.723	0.656	0.819	0.165	0.541

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
log size	-0.167***	-0.0352*	-0.223***	-0.140***	-0.172***	-0.191***	-0.240***	-0.182***	-0.0924***	-0.125***
	(0.0197)	(0.0196)	(0.0316)	(0.0174)	(0.0300)	(0.0257)	(0.0485)	(0.0236)	(0.0183)	(0.0147)
year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
constant	-1.238***	-1.811***	-1.495***	-1.411***	-1.050***	-1.052***	-0.919***	-1.048***	-0.807***	-0.793***
	(0.0940)	(0.0627)	(0.138)	(0.0722)	(0.101)	(0.0904)	(0.201)	(0.123)	(0.0688)	(0.0666)
Obs	90	58	62	63	72	80	36	90	30	98
Adj. R^2	0.439	0.108	0.371	0.597	0.444	0.485	0.339	0.434	0.454	0.487

Notes: Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$ including year fixed effects. Manufacturing firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.

Table B5: Main regression - Services sector

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
log size	-0.177***	-0.154***	-0.187***	-0.119***	-0.177***	-0.108***	-0.128***	-0.163***	-0.0309**	-0.132***
	(0.0185)	(0.0227)	(0.0265)	(0.0123)	(0.0198)	(0.0188)	(0.0165)	(0.0245)	(0.0149)	(0.0209)
constant	-1.187***	-0.499***	-1.055***	-1.061***	-1.076***	-0.930***	-1.135***	-0.913***	-1.263***	-1.049***
	(0.0565)	(0.0634)	(0.0903)	(0.0456)	(0.0710)	(0.0732)	(0.0585)	(0.0722)	(0.0602)	(0.0692)
Obs.	90	55	90	102	86	60	85	60	60	90
Adj. R^2	0.547	0.600	0.380	0.551	0.611	0.396	0.384	0.513	0.048	0.364

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
log size	-0.0360	-	-0.102***	-0.0691***	-0.109***	-0.204***	-0.236***	-0.112***	0.0141	-0.107***
	(0.0226)	-	(0.0290)	(0.0190)	(0.0228)	(0.0172)	(0.0446)	(0.0158)	(0.0260)	(0.0145)
constant	-1.494***	-	-1.619***	-1.088***	-1.129***	-1.013***	-0.815***	-1.053***	-1.117***	-0.901***
	(0.0901)	-	(0.0809)	(0.0759)	(0.0708)	(0.0639)	(0.138)	(0.0622)	(0.0918)	(0.0612)
Obs.	90	-	84	69	72	85	36	90	30	99
Adj. R^2	0.022	-	0.104	0.156	0.359	0.539	0.451	0.428	-0.022	0.370

Notes: Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Services firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.

Table B6: Age dummies - Manufacturing sector

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
log size	-0.257***	-0.215***	-0.206***	-0.148***	-0.177***	-0.212***	-0.217***	-0.270***	-0.0570***	-0.169***
	(0.0197)	(0.0129)	(0.0201)	(0.00704)	(0.0184)	(0.0162)	(0.0145)	(0.0176)	(0.0137)	(0.0162)
ageclass 1-2	0.123	-0.160**	-0.717***	-0.0965**	-0.202*	-0.00114	-0.107	0.112	-0.102	-0.0497
	(0.146)	(0.0651)	(0.132)	(0.0378)	(0.114)	(0.118)	(0.133)	(0.104)	(0.0709)	(0.103)
ageclass 3-5	0.0530	-0.130*	-0.895***	-0.202***	-0.257**	-0.0824	-0.337***	0.0866	-0.226***	-0.159
	(0.153)	(0.0679)	(0.125)	(0.0372)	(0.121)	(0.119)	(0.103)	(0.0984)	(0.0706)	(0.108)
ageclass 6-10	-0.0575	-0.185**	-0.928***	-0.289***	-0.391***	-0.192	-0.224*	-0.0443	-0.429***	-0.343***
	(0.143)	(0.0710)	(0.124)	(0.0394)	(0.120)	(0.117)	(0.117)	(0.0886)	(0.0630)	(0.0980)
ageclass 11+	-0.198	-0.224***	-1.085***	-0.396***	-0.380***	-0.385***	-0.286**	-0.179*	-	-0.451***
	(0.146)	(0.0621)	(0.124)	(0.0374)	(0.107)	(0.116)	(0.128)	(0.0895)	-	(0.0939)
ageclass missing	-	-	-	-0.451***	-	-	-	-	-0.240***	-
	-	-	-	(0.0368)	-	-	-	-	(0.0844)	-
constant	-1.030***	-0.193***	-0.386***	-0.580***	-0.861***	-0.676***	-0.783***	-0.882***	-1.197***	-0.791***
	(0.121)	(0.0519)	(0.0785)	(0.0262)	(0.108)	(0.103)	(0.0929)	(0.0877)	(0.0761)	(0.0969)
Obs.	89	49	89	102	84	59	85	60	60	90
Adj. R^2	0.692	0.856	0.792	0.907	0.600	0.811	0.662	0.848	0.500	0.653

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
log size	-0.164***	-0.0302**	-0.176***	-0.141***	-0.172***	-0.161***	-0.238***	-0.179***	-0.0907***	-0.125***
	(0.0148)	(0.0141)	(0.0248)	(0.0212)	(0.0274)	(0.0209)	(0.0544)	(0.0207)	(0.0163)	(0.0140)
ageclass 1-2	-0.343***	-0.0971	-0.136	0.191	-0.0478	-0.487***	0.126	-0.356***	-0.0321	-0.0826
	(0.0909)	(0.0710)	(0.144)	(0.243)	(0.185)	(0.147)	(0.297)	(0.125)	(0.0960)	(0.0660)
ageclass 3-5	-0.410***	-0.108	-0.351***	0.111	-0.307**	-0.739***	-0.168	-0.334***	-0.195*	-0.112*
	(0.0972)	(0.0664)	(0.110)	(0.252)	(0.142)	(0.149)	(0.247)	(0.120)	(0.109)	(0.0633)
ageclass 6-10	-0.528***	-0.195***	-0.576***	0.159	-0.252*	-0.739***	-0.276	-0.401***	-0.237**	-0.185***
	(0.0944)	(0.0673)	(0.0866)	(0.241)	(0.141)	(0.147)	(0.247)	(0.136)	(0.0873)	(0.0600)
ageclass 11+	-0.787***	-0.342***	-0.773***	0.00783	-0.250	-0.850***	-0.425	-0.580***	-0.345***	-0.364***
	(0.0878)	(0.0689)	(0.0846)	(0.245)	(0.150)	(0.142)	(0.254)	(0.108)	(0.0881)	(0.0564)
ageclass missing	-	-	-1.138***	0.282	-0.351**	-	-0.0301	-	-	-0.0390
	-	-	(0.192)	(0.237)	(0.159)	-	(0.527)	-	-	(0.170)
constant	-0.822***	-1.721***	-1.095***	-1.332***	-0.916***	-0.586***	-0.798**	-0.709***	-0.652***	-0.725***
	(0.0714)	(0.0489)	(0.0538)	(0.207)	(0.135)	(0.0973)	(0.293)	(0.104)	(0.0828)	(0.0654)
Obs.	90	58	62	63	72	80	36	90	30	98
Adj. R^2	0.773	0.480	0.676	0.357	0.481	0.739	0.314	0.555	0.696	0.588

Notes: Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Manufacturing firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. The baseline age category is entering firms. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.

Table B7: Age dummies - Services sector

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
log size	-0.174***	-0.150***	-0.176***	-0.116***	-0.173***	-0.105***	-0.123***	-0.160***	-0.0275***	-0.128***
	(0.0178)	(0.0212)	(0.0238)	(0.00835)	(0.0199)	(0.0144)	(0.0177)	(0.0254)	(0.0101)	(0.0147)
ageclass 1-2	0.0216	0.118	-0.441***	-0.102**	-0.0972	0.0837	0.0582	-0.186	-0.182**	-0.219**
	(0.114)	(0.112)	(0.147)	(0.0470)	(0.146)	(0.105)	(0.160)	(0.161)	(0.0716)	(0.0863)
ageclass 3-5	-0.0767	-0.00148	-0.567***	-0.180***	-0.162	-0.0280	-0.125	-0.186	-0.252***	-0.384***
	(0.103)	(0.118)	(0.145)	(0.0273)	(0.144)	(0.112)	(0.151)	(0.165)	(0.0572)	(0.0762)
ageclass 6-10	-0.0804	-0.131	-0.643***	-0.264***	-0.277**	-0.150	-0.258	-0.383**	-0.591***	-0.566***
	(0.105)	(0.106)	(0.154)	(0.0260)	(0.137)	(0.110)	(0.157)	(0.167)	(0.0552)	(0.0730)
ageclass 11+	-0.375***	-0.201*	-0.886***	-0.323***	-0.323**	-0.489***	-0.422***	-0.499***	-	-0.696***
	(0.0970)	(0.102)	(0.145)	(0.0324)	(0.138)	(0.106)	(0.151)	(0.161)	-	(0.0702)
ageclass missing	-	-	-	-0.471***	-	-	-	-	-0.363***	-
	-	-	-	(0.0385)	-	-	-	-	(0.0583)	-
constant	-1.095***	-0.467***	-0.586***	-0.853***	-0.911***	-0.823***	-0.996***	-0.674***	-0.998***	-0.692***
	(0.0765)	(0.102)	(0.0821)	(0.0340)	(0.117)	(0.0923)	(0.112)	(0.107)	(0.0527)	(0.0561)
Obs.	90	55	90	102	86	60	85	60	60	90
Adj. R^2	0.650	0.664	0.660	0.847	0.654	0.705	0.551	0.628	0.710	0.730

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
log size	-0.0338	-	-0.0915***	-0.0675***	-0.109***	-0.179***	-0.233***	-0.109***	0.0155	-0.104***
	(0.0220)	-	(0.0239)	(0.0206)	(0.0235)	(0.0128)	(0.0490)	(0.0115)	(0.0263)	(0.0119)
ageclass 1-2	-0.257*	-	0.0827	0.00509	-0.0377	-0.461***	-0.111	-0.167**	0.0265	-0.154***
	(0.137)	-	(0.146)	(0.151)	(0.129)	(0.115)	(0.420)	(0.0745)	(0.163)	(0.0460)
ageclass 3-5	-0.405***	-	-0.0114	0.0884	0.0821	-0.567***	-0.0335	-0.289***	-0.0506	-0.215***
	(0.136)	-	(0.153)	(0.150)	(0.0991)	(0.118)	(0.430)	(0.0704)	(0.151)	(0.0491)
ageclass 6-10	-0.509***	-	-0.154	-0.0289	-0.139	-0.737***	-0.329	-0.415***	-0.147	-0.303***
	(0.134)	-	(0.150)	(0.157)	(0.0860)	(0.108)	(0.435)	(0.0694)	(0.144)	(0.0460)
ageclass 11+	-0.799***	-	-0.569***	-0.234	-0.272***	-0.849***	-0.281	-0.589***	-0.342**	-0.484***
	(0.132)	-	(0.144)	(0.141)	(0.0970)	(0.111)	(0.432)	(0.0683)	(0.159)	(0.0437)
ageclass missing	-	-	-0.805***	0.0412	-0.0751	-	-0.193	-	-	-0.113
	-	-	(0.220)	(0.135)	(0.0847)	-	(0.453)	-	-	(0.181)
constant	-1.108***	-	-1.406***	-1.071***	-1.056***	-0.549***	-0.670**	-0.772***	-1.019***	-0.691***
	(0.0974)	-	(0.125)	(0.129)	(0.0999)	(0.0910)	(0.313)	(0.0517)	(0.137)	(0.0519)
Obs.	90	-	84	69	72	85	36	90	30	99
Adj. R^2	0.442	-	0.498	0.196	0.432	0.819	0.402	0.766	0.189	0.602

Notes: Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Services firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. The baseline age category is entering firms. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.

Appendix C

Online Version

In the following we first provide additional methodological and cleaning details. Then, we report the detailed estimates of a number of additional robustness checks, as discussed hereafter.

First, we estimate the main equation using a Least Absolute Deviation estimator, more robust to extreme observations. Results are reported in Table C1.

Second, we adopt a fully non-parametric approach (Li and Racine, 2004). Results are available in Table C2. We also plot micro-aggregated OLS and non-parametric regression lines, and kernel density of the betas of the non-parametric regression with bootstrapped error bands for France (FigureC1).

Third, we adopt a grouped data approach to regression (discussed in Angrist, 1998 and Angrist and Pischke, 2008). Results are reported in Table C3.

Fourth, we enrich the baseline model including age class dummies and interactions of age class dummies with average size in Table C4.

Methodological notes and cleaning

In this subsection we provide further details on the methodology that we apply to create the database used for estimation, including cleaning details. Starting from the aggregate Transition Matrix provided by the DynEmp database, we exclude the cells corresponding to all macro sectors, to avoid double counting. We further exclude cells where employment growth volatility is missing. The strictness of the blanking procedures – aimed at avoiding primary or secondary disclosure – applied on the micro-aggregated data differs from country to country. These differences (that involve a limited number of cells including few units, generally between 5 and 10) may influence cross-country comparability of the estimates. Countries that implement blanking are Australia, Denmark, Spain, Finland, France, Japan, Luxembourg, the Netherlands, New Zealand, Turkey, the United Kingdom and the United States.

Since the level of disaggregation for entering units (in terms of their size class at time $t + j$) produced by the DynEmp routine is more detailed with respect to other cells in the transition matrix, we re-aggregate them. We therefore proceed collapsing the dimension related to size class at time $t + j$, in order to obtain a cell that includes all surviving entrants together, as for the other age classes. Weights used for the aggregation of employment growth volatility in this case are cell average employment shares, calculated using cell-level employment at time t and $t + j$. Note that, in order to implement this aggregation, cells for which average employment for entrants is missing, are dropped (this involves a limited number of cells, mostly in the United Kingdom and New Zealand). Before collapsing, we also drop cells for which the number of entering units is not available due to blanking (United Kingdom only), otherwise the subsequent average size calculation would be influenced by these cells, as well. This re-aggregation does not seem to qualitatively affect the nature of findings proposed.

As highlighted in the paper, the focus is restricted to surviving units for which a window to calculate volatility is available.

Since the relation between volatility and average size is assessed in log terms, cells for which volatility and average size equal zero are dropped. Further details on and robustness linked to this issue have been provided in Section 3 of the paper.

Additional tables and Figures

Table C1: LAD regression

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
log size	-0.241***	-0.234***	-0.236***	-0.153***	-0.189***	-0.229***	-0.208***	-0.301***	-0.0609***	-0.167***
	(0.0164)	(0.0199)	(0.0307)	(0.0169)	(0.0287)	(0.0237)	(0.0224)	(0.0201)	(0.0212)	(0.0202)
constant	-1.121***	-0.305***	-1.140***	-0.805***	-1.142***	-0.738***	-1.058***	-0.807***	-1.394***	-1.073***
	(0.0551)	(0.0506)	(0.119)	(0.0634)	(0.110)	(0.0946)	(0.0840)	(0.0692)	(0.0930)	(0.0681)
Obs.	89	49	89	102	84	59	85	60	60	90
Pseudo R^2	0.487	0.612	0.334	0.392	0.279	0.474	0.442	0.592	0.105	0.342

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
log size	-0.163***	-0.0443**	-0.263***	-0.129***	-0.211***	-0.205***	-0.203***	-0.186***	-0.0801***	-0.140***
	(0.0229)	(0.0190)	(0.0476)	(0.0279)	(0.0274)	(0.0226)	(0.0637)	(0.0211)	(0.0260)	(0.0141)
constant	-1.269***	-1.831***	-1.352***	-1.232***	-1.003***	-1.095***	-1.094***	-1.078***	-0.874***	-0.814***
	(0.0869)	(0.0620)	(0.155)	(0.105)	(0.0950)	(0.0699)	(0.238)	(0.0912)	(0.0997)	(0.0618)
Obs.	90	58	62	63	72	80	36	90	30	98
Pseudo R^2	0.229	0.0819	0.242	0.220	0.352	0.366	0.227	0.294	0.238	0.311

Notes: Least Absolute Deviations regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Manufacturing firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.

Table C2: Non-parametric regression

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
Mean	-0.260	-0.240	-0.182	-0.112	-0.199	-0.161	-0.222	-0.269	-0.056	-0.137
Mode	-0.260	-0.228	-0.170	-0.136	-0.124	-0.192	-0.217	-0.268	-0.0396	-0.172
av. s.e.	(0.0000)	(0.0366)	(0.0662)	(0.0268)	(0.0510)	(0.0493)	(0.0113)	(0.0252)	(0.0100)	(0.0437)
Obs.	89	49	89	102	84	59	85	60	60	90
R^2	0.6662	0.8693	0.4069	0.6684	0.6206	0.7261	0.6438	0.8419	0.2231	0.5243

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
Mean	-0.078	-0.0261	-0.223	-0.148	-0.175	-0.191	-0.238	-0.0884	-0.102	-0.107
Mode	-0.0952	-0.046	-0.223	-0.184	-0.180	-0.191	-0.237	-0.0809	-0.103	-0.147
av. s.e.	(0.0154)	(0.0259)	(0.0000)	(0.0298)	(0.0114)	(0.0000)	(0.0094)	(0.0179)	(0.0173)	(0.0332)
Obs.	90	58	62	63	72	80	36	90	30	98
R^2	0.4754	0.1692	0.395	0.4036	0.4492	0.4983	0.3615	0.4594	0.5919	0.5404

Notes: Non-parametric local linear regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$. Average gradient, mode of the gradient estimates and average standard errors of the gradient estimates are reported. Volatility is calculated over a 3 years time.

Figure C1: Non-parametric estimates - France

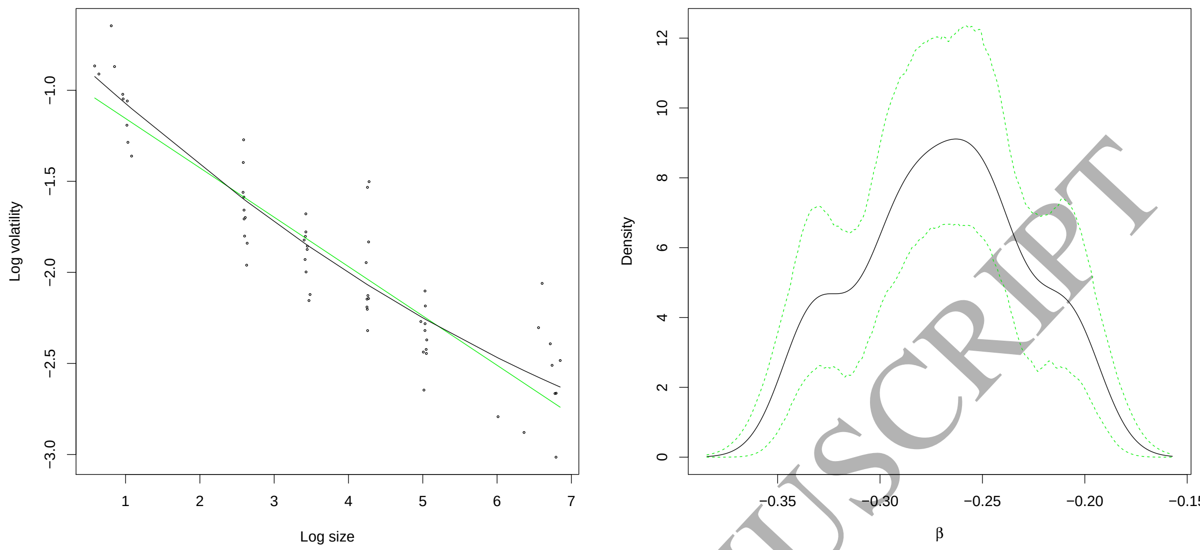

Table C3: Weighted regression

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
log size	-0.269***	-0.258***	-0.275***	-0.190***	-0.234***	-0.271***	-0.261***	-0.331***	-0.138***	-0.225***
	(0.0140)	(0.0160)	(0.0243)	(0.0127)	(0.0262)	(0.0281)	(0.0277)	(0.0255)	(0.0288)	(0.0154)
constant	-1.155***	-0.290***	-1.205***	-0.762***	-1.051***	-0.869***	-0.972***	-0.838***	-1.350***	-0.954***
	(0.0420)	(0.0296)	(0.0711)	(0.0399)	(0.0806)	(0.0784)	(0.0883)	(0.0766)	(0.0896)	(0.0480)
Obs.	89	49	89	102	84	59	85	60	60	90
Adj. R^2	0.872	0.915	0.736	0.727	0.616	0.708	0.635	0.832	0.454	0.749

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
log size	-0.175***	-0.0549***	-0.260***	-0.110***	-0.207***	-0.228***	-0.321***	-0.235***	-0.0763***	-0.177***
	(0.0287)	(0.0164)	(0.0272)	(0.0317)	(0.0254)	(0.0193)	(0.0462)	(0.0360)	(0.0185)	(0.0175)
constant	-1.536***	-1.933***	-1.476***	-1.398***	-1.078***	-1.173***	-0.942***	-1.093***	-0.948***	-0.850***
	(0.0919)	(0.0605)	(0.0961)	(0.123)	(0.0819)	(0.0580)	(0.129)	(0.117)	(0.0504)	(0.0501)
Obs.	90	58	62	63	72	80	36	90	30	98
Adj. R^2	0.329	0.270	0.635	0.167	0.670	0.586	0.687	0.440	0.367	0.697

Notes: Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$ with observation weighted by the number of firms in each cell. Manufacturing firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.

Table C4: Main regression - Age dummies and interaction

	AT	AU	BE	BR	DK	ES	FI	FR	GB	HU
log size	-0.345***	-0.248***	-0.157*	-0.186***	-0.227***	-0.295***	-0.221***	-0.350***	-0.0999***	-0.299***
	(0.0829)	(0.0152)	(0.0859)	(0.0215)	(0.0561)	(0.0636)	(0.0372)	(0.0212)	(0.0208)	(0.0340)
ageclass 1-2 # log size	0.0805	-0.00728	-0.0878	0.0589**	0.0853	0.0859	-0.0677	0.0189	0.0472	0.161***
	(0.0848)	(0.0323)	(0.0891)	(0.0254)	(0.0703)	(0.0696)	(0.0472)	(0.0407)	(0.0285)	(0.0500)
ageclass 3-5 # log size	0.148*	0.0265	-0.0926	0.0549**	0.0446	0.0906	-0.0269	0.145***	-0.000560	0.196***
	(0.0879)	(0.0358)	(0.0904)	(0.0257)	(0.0666)	(0.0706)	(0.0434)	(0.0420)	(0.0380)	(0.0475)
ageclass 6-10 # log size	0.0965	0.0654**	-0.0566	0.0408	0.0243	0.119*	0.0486	0.109***	0.0818***	0.177***
	(0.0842)	(0.0295)	(0.0884)	(0.0267)	(0.0690)	(0.0669)	(0.0417)	(0.0384)	(0.0266)	(0.0436)
ageclass 11+ # log size	0.121	0.0503	-0.0115	0.0246	0.0791	0.121*	0.0699	0.131***	-	0.123***
	(0.0858)	(0.0309)	(0.0890)	(0.0245)	(0.0642)	(0.0654)	(0.0552)	(0.0350)	-	(0.0387)
ageclass missing # log size	-	-	-	0.0468*	-	-	-	-	0.0998**	-
	-	-	-	(0.0253)	-	-	-	-	(0.0466)	-
ageclass 1-2	-0.170	-0.119*	-0.395*	-0.322***	-0.481**	-0.296*	0.149	0.0513	-0.280**	-0.652***
	(0.250)	(0.0673)	(0.218)	(0.0732)	(0.200)	(0.154)	(0.175)	(0.135)	(0.108)	(0.172)
ageclass 3-5	-0.492**	-0.197***	-0.554**	-0.412***	-0.387*	-0.395**	-0.233*	-0.458***	-0.221*	-0.895***
	(0.247)	(0.0610)	(0.222)	(0.0761)	(0.223)	(0.160)	(0.119)	(0.158)	(0.116)	(0.170)
ageclass 6-10	-0.411*	-0.382***	-0.725***	-0.445***	-0.445*	-0.612***	-0.405***	-0.449***	-0.743***	-1.007***
	(0.242)	(0.0873)	(0.215)	(0.0778)	(0.240)	(0.151)	(0.116)	(0.138)	(0.110)	(0.165)
ageclass 11+	-0.647**	-0.375***	-1.058***	-0.489***	-0.641***	-0.812***	-0.552***	-0.671***	-	-0.907***
	(0.247)	(0.108)	(0.227)	(0.0727)	(0.185)	(0.153)	(0.204)	(0.132)	-	(0.156)
ageclass missing	-	-	-	-0.630***	-	-	-	-	-0.619***	-
	-	-	-	(0.0855)	-	-	-	-	(0.155)	-
constant	-0.709***	-0.108***	-0.562***	-0.435***	-0.712***	-0.395***	-0.768***	-0.587***	-1.035***	-0.308**
	(0.237)	(0.0282)	(0.206)	(0.0626)	(0.155)	(0.137)	(0.101)	(0.0644)	(0.0760)	(0.136)
Obs.	89	49	89	102	84	59	85	60	60	90
Adj. R^2	0.704	0.856	0.795	0.916	0.594	0.832	0.680	0.883	0.566	0.729

	IT	JP	LU	NL	NO	NZ	PT	SE	TR	US
log size	-0.213***	-0.00660	-0.131***	-0.222	-0.162*	-0.0148	-0.387***	-0.199***	-0.115**	-0.185***
	(0.0480)	(0.0500)	(0.0250)	(0.144)	(0.0901)	(0.0693)	(0.131)	(0.0426)	(0.0542)	(0.0337)
ageclass 1-2 # log size	0.0435	-0.0168	0.0270	0.126	0.0679	-0.113	0.137	0.0306	0.0601	0.0346
	(0.0511)	(0.0588)	(0.0557)	(0.147)	(0.133)	(0.0812)	(0.154)	(0.0793)	(0.0576)	(0.0495)
ageclass 3-5 # log size	0.0570	-0.000707	-0.0863	0.0422	-0.0674	-0.216***	0.180	-0.0506	0.0150	0.111**
	(0.0594)	(0.0522)	(0.0917)	(0.158)	(0.0921)	(0.0786)	(0.135)	(0.0609)	(0.0758)	(0.0443)
ageclass 6-10 # log size	0.0690	-0.0346	-0.0170	0.122	-0.0243	-0.158**	0.215	0.0232	0.0398	0.113***
	(0.0582)	(0.0515)	(0.0658)	(0.149)	(0.0937)	(0.0725)	(0.132)	(0.0482)	(0.0571)	(0.0387)
ageclass 11+ # log size	0.0832*	-0.0539	-0.0665*	0.0771	0.0478	-0.148*	0.203	0.108**	0.00938	0.0822**
	(0.0491)	(0.0510)	(0.0345)	(0.151)	(0.0973)	(0.0748)	(0.146)	(0.0483)	(0.0572)	(0.0364)
ageclass missing # log size	-	-	-0.0256	0.0730	-0.0872	-	0.160	-	-	-0.0370
	-	-	(0.292)	(0.147)	(0.104)	-	(0.265)	-	-	(0.0682)
ageclass 1-2	-0.504***	-0.0299	-0.251*	-0.149	-0.312	-0.242	-0.371	-0.470*	-0.253	-0.213
	(0.161)	(0.182)	(0.140)	(0.304)	(0.395)	(0.190)	(0.359)	(0.238)	(0.217)	(0.134)
ageclass 3-5	-0.622***	-0.0988	-0.180	0.0504	-0.0506	-0.117	-0.825**	-0.137	-0.250	-0.533***
	(0.184)	(0.161)	(0.195)	(0.374)	(0.252)	(0.178)	(0.325)	(0.215)	(0.269)	(0.138)
ageclass 6-10	-0.786***	-0.0605	-0.611***	-0.179	-0.160	-0.334**	-1.071***	-0.487**	-0.384*	-0.616***
	(0.181)	(0.161)	(0.195)	(0.333)	(0.270)	(0.150)	(0.324)	(0.194)	(0.216)	(0.125)
ageclass 11+	-1.100***	-0.129	-0.630***	-0.171	-0.438	-0.485***	-1.174***	-0.996***	-0.377*	-0.682***
	(0.153)	(0.161)	(0.108)	(0.367)	(0.272)	(0.169)	(0.407)	(0.148)	(0.218)	(0.117)
ageclass missing	-	-	-1.119*	0.118	-0.0204	-	-0.614	-	-	0.196
	-	-	(0.573)	(0.319)	(0.286)	-	(1.063)	-	-	(0.282)
constant	-0.642***	-1.813***	-1.153***	-1.138***	-0.951***	-0.944***	-0.255	-0.634***	-0.562**	-0.500***
	(0.149)	(0.153)	(0.0704)	(0.283)	(0.241)	(0.146)	(0.316)	(0.123)	(0.203)	(0.0939)
Obs	90	58	62	63	72	80	36	90	30	98
Adj. R^2	0.777	0.476	0.652	0.324	0.495	0.777	0.215	0.571	0.674	0.646

Notes: Regression of the log volatility of growth $\sigma^j_{c,t}$ on log of firms size $S_{c,t}$ including age dummies and age dummies interacted with the log volatility of growth $\sigma^j_{c,t}$. The baseline age category is entering firms. Manufacturing firms only over a 3 years time window and pooling together observations from 2001, 2004 and 2007. Robust standard error in parenthesis with *** p<0.01, ** p<0.05, * p<0.1.