Marburg Geography

Working Papers on
Innovation and Space

Characteristics of Regional Industry-specific Employment Growth – Empirical Evidence for Germany

07.11

Matthias Duschl and Thomas Brenner

Philipps Universität Marburg

Impressum:

Working Papers on Innovation and Space
Philipps-Universität Marburg

Herausgeber:

Prof. Dr. Dr. Thomas Brenner
Deutschhausstraße 10
35032 Marburg
E-Mail: thomas.brenner@staff.uni-marburg.de

Erschienen: 2011

Characteristics of Regional Industry-specific Employment Growth – Empirical Evidence for Germany

Matthias Duschl[1] and Thomas Brenner

both from Section Economic Geography and Location Research, Institute of Geography, Philipps-University, Marburg.

Abstract:

Regional growth dynamics significantly deviate from a normal process. Using industry-specific employment data for German regions, we find that the asymmetric Subbotin distribution is able to account properly for extreme positive and especially negative growth events. This result confirms previous studies on growth rates of firms and countries and fills an important research gap at the meso-level of regions. Furthermore, we show that regional growth patterns emerge to a considerable degree from the aggregation of micro-level firm growth rates distributions and that the knowledge intensity of the respective industries increases the regions' risk of being effected by extreme growth events.

Keywords: regional employment growth, stochastic characteristics, asymmetric Subbotin distribution, extreme negative growth events.

JEL Classifications: C46, C50, R11

[1] Corresponding Author: Matthias Duschl, Section Economic Geography and Location Research, Institute of Geography, Philipps-University, Marburg.
E-Mail: Matthias.duschl@staff.uni-marburg.de

1 Introduction

The analysis of stochastic characteristics of economic phenomena played both in the past and in the present a prominent role in economic research. The seminal work of Gibrat (1931) opened up the way for comprehensive investigations of the statistical properties of the size distribution of firms (e.g. Hart & Prais 1956; Simon & Bonini 1958; Lucas 1978, Axtell 2001) and its relation to firm dynamics in terms of an autoregressive stochastic process (e.g. Mansfield 1962; Evans 1987; Hall 1988). Later studies began to focus on the shape of the distribution of growth rates itself (e.g. Stanley et al. 1996; Amaral et al. 1997; Bottazzi & Secchi 2003). Recent empirical evidence on the basis of firm level data from several countries (e.g. Reichstein & Jensen 2005 for Denmark; Bottazzi et al. 2002 for Italy, Bottazzi et al. 2011 for France, or Duschl et al. 2011 for Germany) as well as at the disaggregated level of industries (Bottazzi et al. 2001 for the pharmaceutical industry) shows that the expectation of normal distributed growth rates is consistently rejected. Rather than the bell-shape of a normal curve, an exponential tent-like shaped distribution is observed, with tails that are much fatter than the ones of a normal distribution. In other words, growth events at the extremes occur with a higher probability (Amaral et al. 1997; Bottazzi & Secchi 2006a). Similar findings can be reported for countries, i.e. at a much higher level of economic aggregation. Whilst Quah (1996 and 1997) and Jones (1997) amongst others studied the income distributions of national economies, more recently the focus has shifted on the countries' income growth rates (e.g. Lee et al. 1998; Canning et al. 1998; Amaral et al. 2001; Maasoumi et al. 2007; Castaldi & Dosi 2009).

Tent-shaped and fat-tailed distributions of growth rates are found to be an extremely robust feature for both firms and countries. In addition, they show a much higher regularity and homogeneity than the respective size distributions. Accepted as a stylized fact, these distributions were further extended to and confirmed for the growth of whole industries (Sapio & Thoma 2006; Castaldi & Sapio 2008) or even scientific journals (Havemann et al. 2005). A logical next step is to look at the intermediate level between firms and countries, namely at the level of regions. Whereas the distribution of economic activities across regions and spatial clustering tendencies of different industries are well studied (e.g. Ellison & Glaeser 1997; Dumais et al. 2002; Brenner 2004 and 2006; Bottazzi et al. 2008), the distributional characteristics of the regions' growth rates still remain an important research gap –

neither regional economies as a whole, nor regional industry-specific growth processes have been investigated yet. Therefore, the aim of this paper is to analyse and to explain the stochastic properties of regional economic growth from an explicit disaggregated perspective.

Exploring the stochastic characteristics of regional industry-specific employment growth could reveal underlying mechanisms which govern economic growth. Explaining the emergence of the observed stochastic patterns at the meso-level of regions by considering the micro-level of firms and industrial specificities could improve our understanding of contemporary regional economic growth processes. In short, the probability distribution of the growth rates serves in this paper both as the *explanandum* and *explanans* for regional growth.

Using a time panel from 1999 to 2007 of industry-specific employment data for German regions, we identify the best fitting theoretical distribution for describing regional growth rates. Starting from a two-parameter Laplace distribution, we then increase the number of free parameters and test, if the more general symmetric as well as asymmetric Subbotin distribution, as described by Bottazzi and Secchi (2011), fits significantly better the empirical observations. It is found that the asymmetric Subbotin distribution most adequately describes industry-specific regional growth rates: regional economies are heavily prone to experience extreme positive and especially negative growth events. In an attempt to explain the emergence of the stochastic properties, we find that regional growth patterns are to a considerable degree the result of aggregation of the micro-level firm growth rates distributions. However, even in industries with the largest number of firms, fat tails remain a prominent feature and are not vanished out by aggregation. Finally, comparing different types of industries, the most striking fact is that knowledge-intensive industries have significantly fatter tails than non-knowledge intensive industries.

The paper is structured as follows. In the second section the theoretical framework is outlined and predictions for the distribution of regional industry-specific employment growth rates are presented. The methodology to identify the best distribution is developed in the third section, and data issues are discussed in the fourth section. In section five, the methodology is applied to German regions and economic explanations for the observed behaviour are provided. Section six summarizes and gives some tentative conclusions for economic theory and modelling.

2 Theory and predictions

2.1 Firm-like behaviour of regional economies

A significant departure from normality suggests that some other mechanisms than the central limit theorem are at work (Canning et al. 1998: 340). Therefore, confronted with the empirically observed firm growth rate distributions, Bottazzi & Secchi (2006a) developed a stochastic model of firm growth. In contrast to Gibrat (1931), they assume that economic opportunities, which sum up to the firms' growth rate in a given time period[1], are limited in their amount available to the firms and are not independent: a kind of increasing returns mechanism induces a cumulative and self-reinforcing process in the assignment of economic opportunities. Successful firms, which have realized more of those opportunities in the past, exhibit a higher probability of taking up new ones. Due to the limitedness of economic opportunities, competing firms directly affect each other – if the market share of one firm grows, the market shares of its competitors need to shrink. The emergence of a tent-shaped distribution with fat tails on both sides can result from the increased probability that a large number of opportunities is assigned to a few firms only, whereas most other firms do not receive any growth opportunity at all.

Employment in a region is the sum of the employment in the firms located in this region. Hence, regions also compete for limited economic growth opportunities. And if regions, which have received such growth opportunities, are exposed to higher probabilities to receive even further opportunities, then the stochastic model of firm growth would apply *mutatis mutandi* to regional economies and the distribution of their growth rates should consequently depart from normality. But as Castaldi & Dosi (2009: 489) argue, inter-regional competition differs in many aspects from inter-firm competition. First, it is reasonable to consider territorial competition as a competition for mobile production factors and firm locations, which may cause dynamic and turbulent processes (Cheshire & Malecki 2004: 260). Secondly, limited growth opportunities could be overpassed by collective efforts in the regional policy arena, for example by trying to achieve higher growth rates in the overall regional system through cooperation and coordination or by diversifying the regional production portfolio into new

[1] Economic opportunities as well as the resulting growth rates may be either positive or negative by nature.

areas of activity. Due to the latter issue it is important to focus on the regional growth dynamics of single industries, since industry-specific markets are the very locus of competition for economic opportunities. Moreover, also the drivers of positive and negative feedback loops in the context of growth opportunities accumulation turn out to be industry-specific (Castaldi & Dosi 2009; Castaldi & Sapio 2008). Referring to the language and concepts of the New Economic Geography (e.g. Krugman 1991), localization economies, which are by definition industry-specific, might induce increasing returns and cause an inter-regional dependence of growth rates. As already Marshall (1890) points out, when a new firm locates into a region, the other firms of the same region and industry might benefit from labour pooling, economies of scales in intermediate inputs and knowledge spillovers. The exit of a firm might decrease the chance of survival for other co-located firms of that industry in an analogous way. Finally and in accordance with endogenous growth theories (e.g. Romer 1990), the diffusion of growth opportunities stemming from innovation throughout the economy has both a geographical and technological dimension: feedback loops are bounded to work within regions and the capacity to absorb knowledge spillovers depends amongst others on the industrial structure of the regions (see Döring & Schnellenbach 2006 for an overview on the issue of knowledge spillovers). Briefly stated, the disaggregated level of industries is an adequate observational unit for the analysis of regional growth processes.

By taking a closer look at the micro-economic processes within a regional economic system, we argue that a region’s firms compete over a finite set of industry-specific growth opportunities, but also interact and learn at the industrial level, leading to self-reinforcing dynamics in opportunity catching. This reasoning leads us to:

> ***Hypothesis 1****: Self-reinforcing mechanisms still prevail at the regional level. Consequently, fat tails as signs of a higher probability of extreme positive and negative growth events emerge as a key feature of regional industry-specific employment growth.*

2.2 The link between the micro- and the meso-level

Firms, industries as well as regional and national economies are all complex systems comprising a large number of interacting subunits (Amaral et al. 2001: 127). As it is initially stated, the growth rates of firms are not normally distributed. If only a few of the regions' firms change their size each year, the characteristics of the regional growth rates could be a mere artefact of the aggregation of firms' growth rates. However, if the number of firms is high and their growth rates are independent from each other, it is reasonable to assume that the central limit theorem is at work, leading to normal distributed growth rates at a higher level of aggregation (Fu et al. 2005; Fagiolo et al. 2008: 641). Hence, we should observe for regions:

> ***Hypothesis 2a****: The higher the number of firms belonging to an industry, the more is the regional growth rates distribution of that industry similar to a normal distribution.*

The opposite hypothesis would state that correlated and interdependent firm level growth rates aggregate in a non-trivial way: instead of compensating out in the aggregation process, they may amplify and produce extreme regional growth events (Castaldi & Sapio 2008: 525).

Another issue concerning the micro-meso link is the dependence of the dispersion of the growth rates on the employment number in a region. It was Hymer & Pashigian (1962) who first discovered that the variance of firms' growth rates was inversely related to their size. Stanley et al. (1996) state that the spread of the distribution of growth rates decreases with increasing firm sales as a power law:

$$\sigma(S) = c * S^{\beta} \tag{1}$$

with S being the size of the units of observation, $\sigma(S)$ the dispersion of their growth rates conditional on size, c a constant and β the scaling exponent. Empirical studies for firms in different industries (Amaral et al. 1997) and for national economies (Lee et al. 1998) estimate a scaling exponent β ranging mostly between -0.15 and -0.20. This robust power-law scaling regularity is declared to be a universal feature of the growth of complex economic organizations involving large numbers of interacting subunits (Stanley et al. 1996; Amaral et al. 2001). Two explanations for the observed scaling behaviour can be put forward. The first is supported by Canning et al. (1998) and others, who consider the scaling exponent β as an indicator for the strength of micro-economic interactions between the subunits of a system, here the firms of a region[2]. If β is 0, a perfect correlation between the subunits exists, since there is no dependence of the dispersion of growth rates on the employment number. If β is -0.5, any correlation between the subunits is absent, meaning that the volatility of a system falls with the square root of its size. Finally, if β lies between the two limiting cases, the dispersion of the growth rates decays as a power law, but not as fast as the square root, indicating the presence of some considerable positive interactions between the firms of a region. Second, Bottazzi & Secchi (2006b) argue that the observed scaling relationship can be explained as a diversification effect, since larger economic entities like firms tend to operate in a higher number sub-markets of an industry due to scope economies of diversification. Up-scaling this argumentation to the level of regions, the dispersion of overall growth may be reduced over-proportionally for regions with higher employment numbers. Both explanations lead us to:

> ***Hypothesis 2b****: The higher the number of average regional employment in an industry, the smaller is the dispersion of the conditional distribution of the respective growth rates.*

2.3 Relationship between type of industry and stochastic characteristics of regional growth

Two arguments make a finer break-down of the regional economy into single industries necessary. First, stochastic properties at a higher level of aggregation may be a sheer outcome of the aggregation process (Bottazzi & Secchi 2003: 218; Brock 1999: 431). As argued above, true economic mechanisms of regional economic growth operate at the level of specific industries. Hence, it has to be tested whether the findings survive at different levels of aggregation. Second, the stochastic properties may also differ across heterogeneous industries. A systematic comparison of different types of industries could provide new insights for explaining the characteristics of regional growth dynamics.

Three lines of distinction seem to be relevant. First, industries differ in their knowledge intensity in production (Smith 2002). Using data on German labour market regions, Schlump & Brenner (2010)

---

[2] Scaling law models rely on the assumption of an identical size of the subunits, which strictly speaking does not hold in the reality of regional economies. This restricts to some degree the explanatory power of the scaling exponent as an indicator of the strength of the interactions between the firms within a region.

prove that the effects of universities and public research and development on regional employment growth differ significantly across industries. Fornahl and Brenner (2009) point out that especially industries, which rely on a scientific knowledge base, show a higher geographical concentration of innovation activities. As innovations are inherently disturbing, it is expected that these industries are more prone to experience extreme growth events. Secondly, also patterns of innovation per se vary across industries. Pavitt (1984) demonstrates the differences in underlying production technologies, competition processes and learning modes. The extent of opportunities for innovations and thus growth obviously depends on the particular technological regime of an industry (Dosi 1988). Finally, we are the first in the literature to explicitly include service industries in the study of the stochastic characteristics of growth. Thus, it is natural to contrast the growth dynamics of service and manufacturing sectors. To sum up, we get:

> ***Hypothesis 3****: Stochastic characteristics of regional employment growth depend on the type of industry under consideration.*

3 Method of fitting the best theoretical distribution

In search for a more general and flexible distributional model that describes the empirical distribution of growth rates *g*, the Subbotin distribution family was introduced into economics by Bottazzi et al. (2002):

$$f(g;b,a,m) = \frac{1}{2 * a * b^{\frac{1}{b}} * \Gamma(1+\frac{1}{b})} * \exp\left(-\frac{1}{b} * \left|\frac{g-m}{a}\right|^{b}\right) \tag{2}$$

with $\Gamma(.)$ standing for the gamma function. Three parameters define the distribution: the location parameter *m*, which indicates the existence of a general trend in the data, the scale parameter *a*, which determines the spread or dispersion of the distribution, and the shape parameter *b*. Both the normal (*b* = 2) and Laplace (*b* = 1), also known as the symmetric exponential distribution, are particular cases of the Subbotin family of probability densities. Thus this distribution family allows for a continuous variation from non-normality to normality, with a smaller shape parameter *b* representing fatter tails of the density. Furthermore, it can be extended to a 5-parameter family of distributions, which is able to cope with asymmetries in the data. In addition to the location parameter m, the asymmetric Subbotin density distribution has two scale parameters a_{left} and a_{right} for the values below and above m and two shape parameters b_{left} and b_{right} describing the tail behaviour on the left and right side of the distribution:

$$f(g;\ b_l, b_r, a_l, a_r, m) \tag{3}$$
$$= \frac{1}{\mathrm{C}} * \exp\left(-\left[\frac{1}{b_l} * \left|\frac{g-m}{a_l}\right|^{b_l} * \theta(m-g) + \frac{1}{b_r} * \left|\frac{g-m}{a_r}\right|^{b_r} * \theta(g-m)\right]\right)$$

where $\theta(g)$ is the Heaviside theta function and $\mathrm{C} = a_l {b_l}^{1/b_l - 1}\Gamma(1/b_l) + a_r {b_r}^{1/b_r - 1}\Gamma(1/b_r)$ a normalization constant (Bottazzi & Secchi 2011).

The aim of the study conducted here is to find out which of the theoretical distributions fits best to reality. To anticipate: the normal distribution will be robustly rejected by several standard parametric tests. Consequently, we restrict the analysis to the Laplace, Subbotin and asymmetric Subbotin distribution[3]. We apply a similar procedure as described in Brenner (2006). First, the likelihood value is calculated for each parameter set of the respective theoretical distributions using the regional employment growth rates. This is done for each single industry separately. The likelihood value L is the probability that the empirical situation occurs according to the theoretical distribution. The maximum likelihood $\hat{L}$ is the maximal value of Lthat can be reached for any parameter set of the respective distribution[4]. The parameter sets with the maximal likelihood value provide those distributions that best describe reality.

The Laplace distribution contains two parameters, the Subbotin three and the asymmetric Subbotin five. Furthermore, the Laplace is a special case of the Subbotin, and the Subbotin can be extended to the asymmetric Subbotin. Therefore, the more general distribution will always fit reality better than the restricted or nested one, so that $\hat{L}_{Subbotin\ asym.} \geq \hat{L}_{Subbotin} \geq \hat{L}_{Laplace}$ is satisfied for all industries. Finally, the likelihood ratio test is performed to check whether the more general theoretical

[3] Several tests with the normal distribution confirmed the findings in the literature for our case.
[4] The maximum likelihood estimation for the Laplace distribution can be solved analytically. In the case of both versions of the Subbotin distributions, the estimation becomes a more complex issue which has to be solved numerically (see AGRO 1995 or MINEO 2003 for the symmetric and Bottazzi & Secchi 2011 for the asymmetric Subbotin). Subbotools v1.1, which is provided by BOTTAZZI (2004), was used here for the maximum likelihood estimation.

distribution with more free parameters describes reality significantly better than the more restricted one with less free parameters. To this end, the value

$$\lambda = 2 * \ln(\hat{L}_{more\ general}) - 2 * \ln(\hat{L}_{more\ restricted}) \quad (4)$$

is calculated for each pair of theoretical distributions. λ measures this difference in the fitting of the data. Statistical theories tell us that λ can be expected to follow a χ^2-distribution with degrees of freedom equal to the difference between the numbers of free parameters of the compared distributional models (Mittelhammer 1995). Hence, the hypothesis that the more general distribution is not more adequate than the more restricted distribution can be tested. Only when it is rejected, the more general distribution is reported.

The likelihood ratio test procedure answers the question of which theoretical distribution describes the empirical data better. However, it does not answer the question of whether the distribution describes the empirical data adequately. To test whether the empirical distribution and the asymmetric Subbotin distribution, the most general theoretical distribution of the analysed distributional portfolio and consequently the distribution with the highest fit, deviate from each other significantly, the Kolmogorov-Smirnov test is used.

4 Data issues

4.1 Empirical data

The data used in this approach was collected by the German Federal Institute of Labour (IAB) at the 30th of June in each year. The data set contains a time panel from 1999 to 2007 of the number of employees and the number of firm establishments for each 4-digit industry[5] and for each municipality in Germany. Industries are denoted by *i*, regions by *r*. In order to check the robustness of the results to the choice of the level of industrial aggregation and the definition of regions, we perform the analysis separately for each 4-digit, 3-digit and 2-digit industry as well as for three different definitions of

[5] Industries are classified according to the WZ-03 classification, which was the standard classification of industries in Germany for the analysed time period.

regions: labour market regions as defined by the IAB (Binder & Schwengler 2006), labour market regions as defined by Eckey et al. (2006) and administrative districts. Altogether, nine combinations of industries and regions result as levels of investigation (see Tab. 1).

Table 1 Levels of investigation

← *Definition of regions* →						
	LMR-IAB	**2-digit**	**LMR-IAB**	**3-digit**	**LMR-IAB**	**4-digit**
	270 regions	*59 industries*	*270 regions*	*212 industries*	*270 regions*	*459 industries*
	LMR-Eckey	**2-digit**	**LMR-Eckey**	**3-digit**	**LMR-Eckey**	**4-digit**
	150 regions	*59 industries*	*150 regions*	*212 industries*	*150 regions*	*459 industries*
	Districts	**2-digit**	**Districts**	**3-digit**	**Districts**	**4-digit**
	413 regions	*59 industries*	*413 regions*	*212 industries*	*413 regions*	*459 industries*

← Level of industrial aggregation →

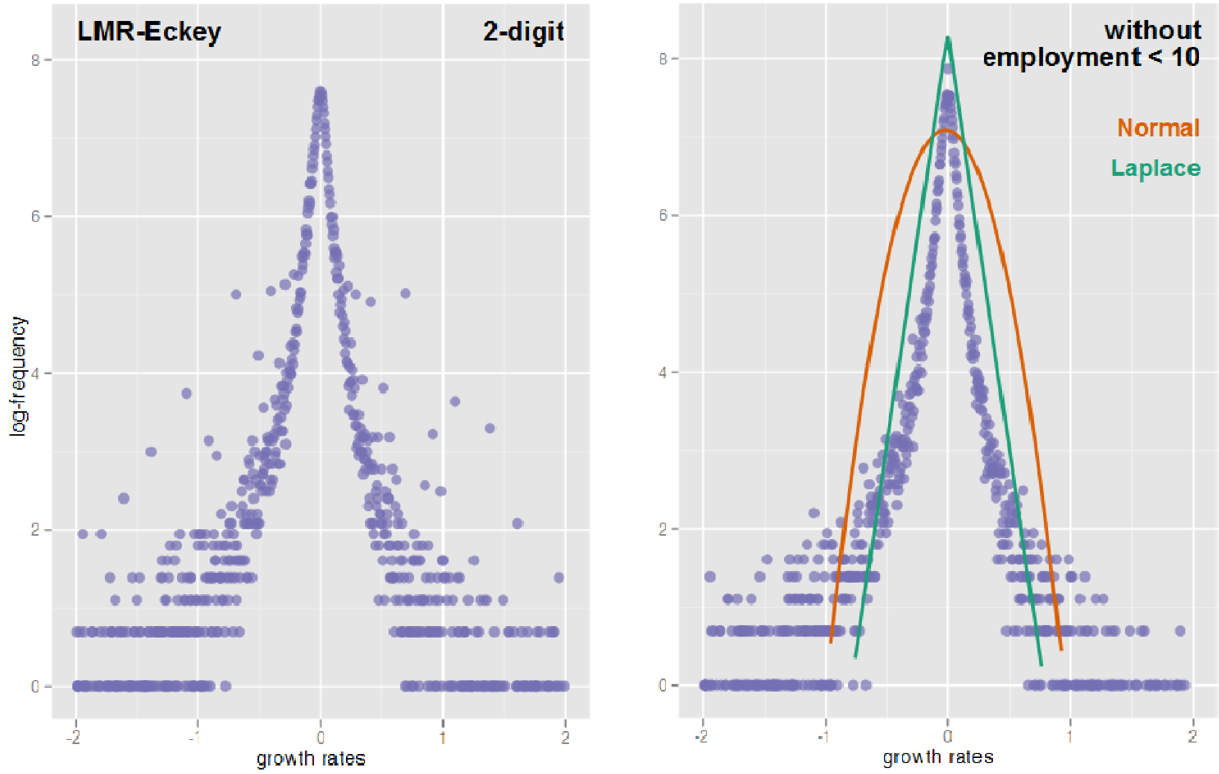
4.2 Calculation of growth rates

Industry-specific regional employment growth rates $g_{i,r,t}$ are calculated by taking the differences of the natural logarithms of the regional employment stocks between two successive years *t* in an industry:

$$g_{i,r,t} = \ln\,(employment_{r,i,t+1}) - \ln\,(employment_{r,i,t}) \quad (5)$$

Fig. 1 displays a frequency distribution of the binned regional growth rates, pooled together from all years and industries at the 2-digit level for Eckey's labour market regions. Plotted on a log-log scale, a tent-like shape known from the Laplace distribution emerge at the centre. However, the tails on both sides are much fatter in comparison to the normal as well as the Laplace distribution. This implies that growth events at the extremes are more probable than could have been expected according to the normal or Laplace distribution. Furthermore, it is already visible that the left tail is even more pronounced.

Figure 1 Frequency distribution of pooled $g_{i,r,t}$ (left panel) and pooled $g_{i,r,t}$ excluding all cases with $employment_{i,r,t} < 10$ (right panel)[6]



The over-dispersed points in the left panel of Fig. 1 are an artefact of regions, where employees in a certain industry are only few in numbers, since the growth of small quantities can manifest itself only in a limited number of different growth rates. Therefore, we truncate the data at minimum 10 industry-specific employees working in a region. As the right panel of Fig. 2 shows, the overall shape remained unaffected, but the noise was reduced strongly.

Pooling together growth rates from different years requires that the assumption of temporal stationarity of the underlying growth process holds true (Bottazzi & Secchi 2006a: 239). Indeed, the growth rates of two successive years show only a slight positive correlation – the Spearman *rho* correlation coefficients, calculated for each industry and each successive years, range in their median values between 0.05 and 0.13 across all nine levels of investigation (see Tab. X1 in the Appendix). Analogously, we test for the dependence of the growth rates on the employment number. The correlation was in general slightly negative, which means that regions with fewer employees tend to grow faster. However, the dependence of the growth rates on the employment number is weak again,

[6] For visual reasons, the plots are truncated on the x-axis at -2 and 2. Nevertheless, the plotted interval includes more than 99.9% of all cases.

with median values for Spearman *rho* ranging between -0.06 and -0.10. Hence, we decided to pool together the growth rates from different years and from regions with different employment numbers.

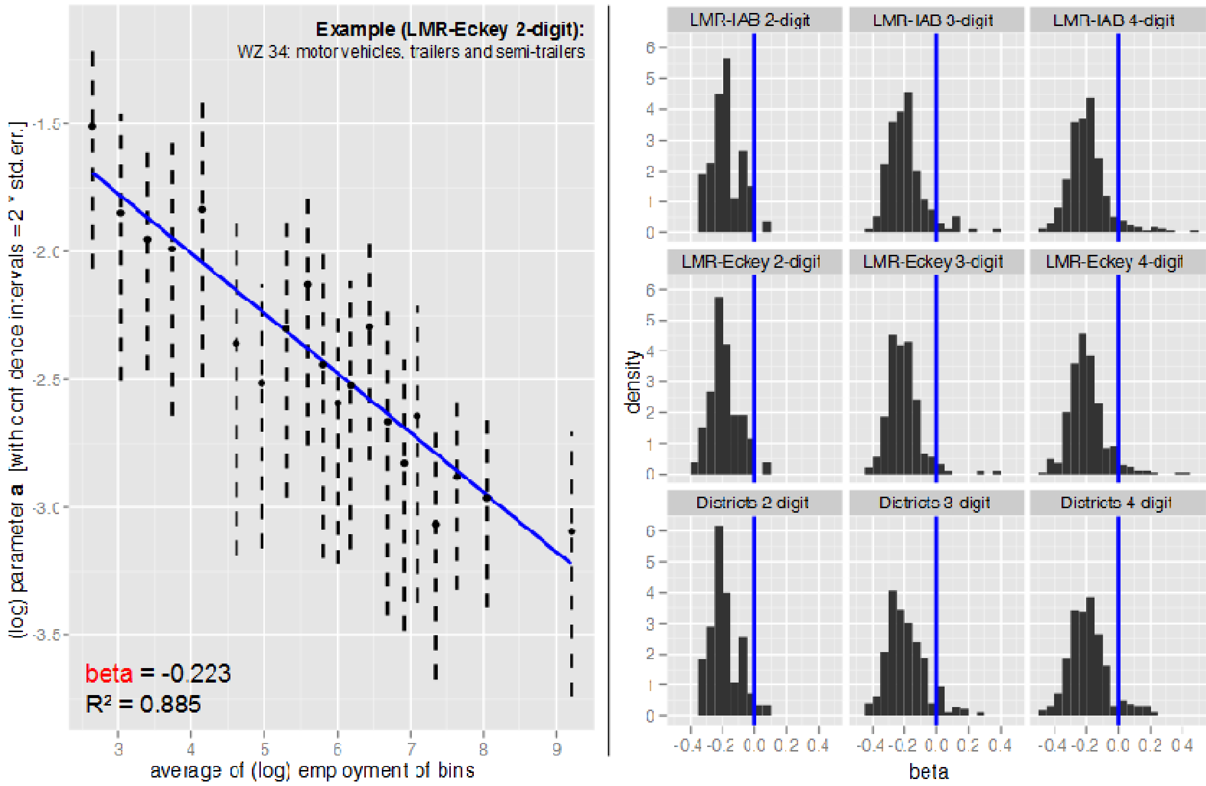
4.3 The relationship between the dispersion of growth rates and regional employment number

A further important issue regarding the possibility of pooling together data of different regions is the dependence of the dispersion of the growth rates on the employment number of a region – this so-called scaling relationship was expected in hypothesis 2b to be negative. In order to test for it in the case of regions, we allocated the logarithms of the regional employees into equipopulated bins. In the literature the standard deviation is commonly used as a measure for the conditional dispersion of growth rates. Since the standard deviation is a poor indicator for dispersion in case of a non-normal distribution, we directly estimated the scale parameter a of the Subbotin distribution for the associated growth rates inside each employment bin. Figure 3 (left panel) shows for one exemplary industry on a log-log scale the estimated Subbotin scale parameters a versus the average employment numbers for each bin. Likewise to firms and countries a clear power-law scaling can be observed, which can be described by the following linear relationship on log-scale:

$$\log\left(a_{i,q}\right) = c + \beta_i * \log\left(\overline{employment}_{i,q}\right) + \varepsilon_i \tag{6}$$

with q denoting the bins or quantiles. Industries which contain more than 800 observations were divided into 20 bins and industries with at least 400 observations into 10 bins. However, industries with less than 400 observations were removed entirely from the sample in order to ensure statistical reliability. The density distributions of the estimated industry-specific scaling exponents β_i are illustrated on the right panel of Fig. 2. Almost all estimated scaling exponents β_i show the expected negative sign and vary between the two limiting cases $\beta_i = 0$ and $\beta_i = -0.5$. In the majority of the cases, β_i concentrates around a value of -0.2, thus confirming the findings for firms and countries (Amaral et al. 1997) and indicating the presence of some considerable positive interactions between the regions' firms.

Figure 2 Scaling relationship for an exemplary industry (left panel) and distributions of the estimated industry-specific scaling exponent β_i (right panel)



Being confronted with a significant dependence of the dispersion of the conditional distribution of growth rates on the regional employment number, we rescaled the growth rates according to:

$$\tilde{g}_{r,i} = \frac{g_{r,i}}{employment_{r,i}{}^{\beta_i}} * \overline{employment}^{\beta_i} \tag{7}$$

The calculation of the industry-specific rescaled regional growth rates $\tilde{g}_{r,i}$ is based on the Subbotin density distribution. For normalization, the average employment number of all industries is used. Only after rescaling all values of $\tilde{g}_{r,i}$, they can be interpreted as different realizations of the same stochastic process. These rescaled growth rates are finally allowed to be pooled together.

5 Empirical results

5.1 Fitting of the theoretical distribution for regional industry-specific growth rates

A normal distribution of the growth rates as the underlying theoretical distribution is rejected by several standard parametric tests[7] for virtually all industries, independent of the regional definition and industrial level analysed (see Tab. X2 in the Appendix). Therefore, we restricted the analysis to the Laplace, Subbotin and asymmetric Subbotin distribution. The method explained in section 3 allows for identifying separately for each industry the best distribution out of the potential candidates by counterbalancing the improvements in the goodness of fit with the loss of degrees of freedom due to an increase of the number of parameters (Fagiolo et al. 2008). Tab. 2 reports the (relative) numbers of occurrences of the best fitting distribution identified by the likelihood ratio test procedure.

Table 2 Results of the likelihood ratio test for the identification of the best fitting distribution

	Laplace		**Subbotin symmetric**		**Subbotin asymmetric**		**N[1]**	***KS-Test [2] (H0 rejected)***
LMR-IAB 2-digit	1	(2%)	2	(4%)	50	(94%)	53	59%
LMR-IAB 3-digit	3	(2%)	20	(10%)	165	(88%)	188	72%
LMR-IAB 4-digit	4	(1%)	37	(10%)	312	(89%)	353	73%
LMR-Eckey 2-digit	5	(10%)	4	(7%)	43	(83%)	52	20%
LMR-Eckey 3-digit	5	(3%)	32	(18%)	144	(79%)	181	31%
LMR-Eckey 4-digit	12	(4%)	44	(13%)	273	(83%)	329	36%
Districts 2-digit	1	(2%)	3	(5%)	51	(94%)	55	71%
Districts 3-digit	1	(1%)	16	(8%)	175	(91%)	192	79%
Districts 4-digit	0	(0%)	29	(8%)	333	(92%)	362	85%

Note: [1] Number of industries analyzed. This number varies between different definitions of regions, because industries with less than 400 observations were excluded from the sample in order to ensure statistical reliability (see section 4.3). For the different regional definitions, the number of observations is theoretically the same, however the more detailed the regional delimitations are the more missing values across the regions naturally occur.

[2] Kolmogorov-Smirnov Test with H0 that empirical distribution is identical to theoretical distribution, here the asymmetric Subbotin, at significance level of 0.05.

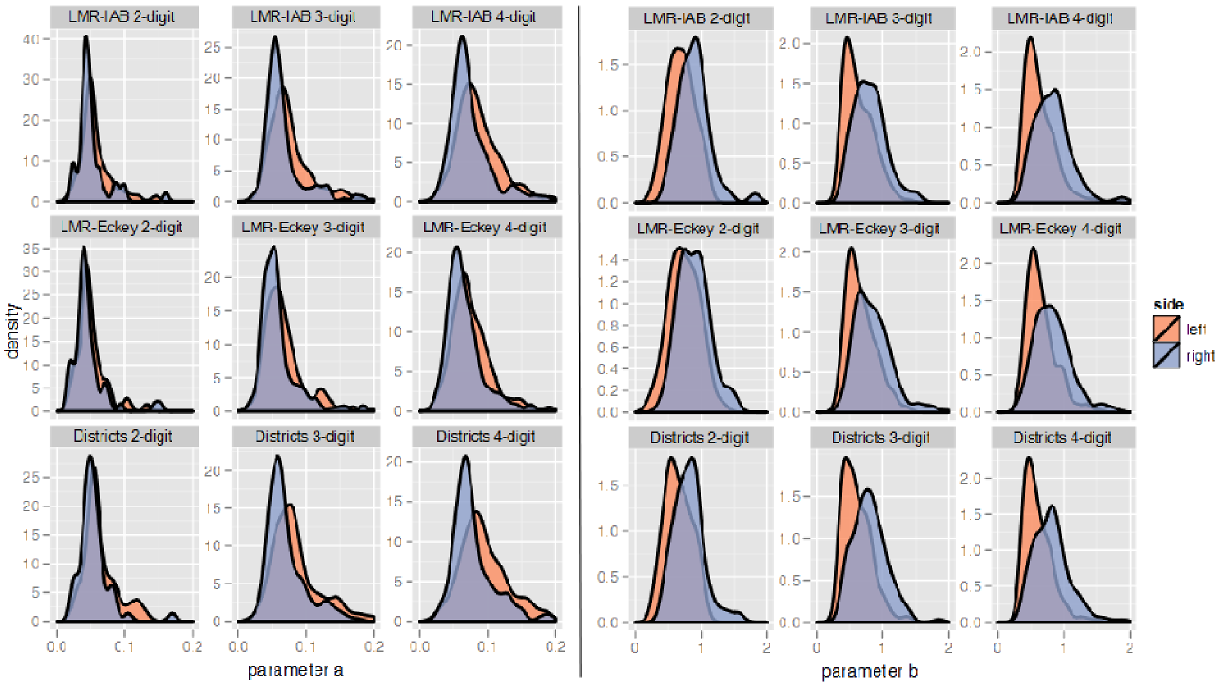
The asymmetric Subbotin distribution results to be the best choice out of the theoretical distributions for describing regional industry-specific growth in more than four out of five cases. In order to assess its adequacy to describe the empirical growth rates distribution, the relative rejection frequency of the

[7] We employed the Kolmogorov-Smirnov normality test, the Jarque-Bera test for a deviation from normality in terms of kurtosis or skewness or both (JARQUE & BERA 1980), and the Anscombe-Glynn test for a deviation from normality in terms of an excess kurtosis (ANSCOMBE & GLYNN 1983).

Kolmogorov-Smirnov test is reported in the last column of Tab. 2. Since the number of observations is relatively high, already small deviations suffice to reject H0. Acknowledging this, a considerable improvement of the fit to the empirical data compared with the normal distribution, which was rejected consistently, can be achieved. This holds especially true in the case of labour market regions as defined by Eckey, for which the asymmetric Subbotin is only rejected in 20% (for 2-digit industries) to 36% (for 4-digit industries) of the cases, whereas the other regional definitions show a higher rejection rate. The difference between the three definitions of regions can be explained, on the one hand, by a higher number of observations for districts (413 spatial units) and IAB labour market regions (270 spatial units) in comparison to Eckey (150 spatial units), which increases the test power and consequently the probability of a null rejection. On the other hand, however, the Eckey definition might be most adequate in terms of describing the intra-regional feedback loops mechanisms which lead to the asymmetric Subbotin distribution. The same argument could be put forward to explain the lower rejection rates at the 2-digit industrial aggregation level.

A closer look at the estimated parameters for the scale (a_{left} and a_{right}) as well as the shape (b_{left} and b_{right}) of the asymmetric Subbotin distribution reveals further insights into the stochastic properties of regional industry-specific growth process (see Fig. 3).

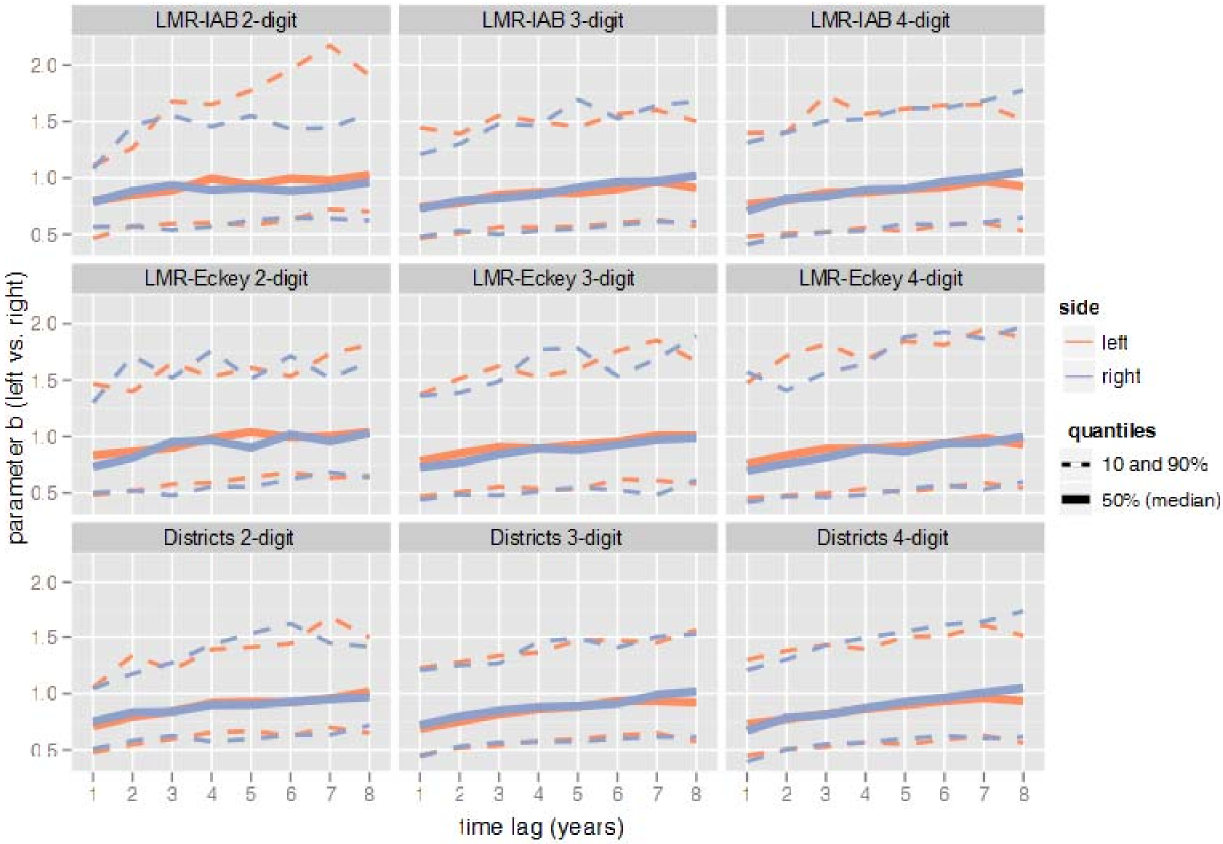
Figure 3 Density distributions of the estimated parameters a_{left}, a_{right}, b_{left} and b_{right} of the asymmetric Subbotin distribution.



The left panel of Fig. 3 contrasts the respective right and left side scale parameters a_{left} and a_{right}, which were estimated separately for all industries. Growth rates that are below the general trend m tend to have a slightly higher dispersion than the growth rates above the trend. The Wilcoxon signed-rank test confirms a significant difference for all nine combinations of industrial levels and regional definitions. Differences between both sides of the distribution are even more pronounced in case of the shape parameters (see right panel of Fig. 3): b_{left} is significantly smaller than b_{right}. This means that regions are more prone to experience extreme negative growth events than extreme positive events. But both the tails on the left and on the right side are fatter than the respective tails of the normal distribution (b = 2) and, in most industries, even than the tails of the Laplace distribution (b = 1).

It could be argued that the influence of mechanisms responsible for fat tails may fade away when longer time horizons are considered. Consequently, a progressive normalization of the growth rates' distribution should take place as an effect of the central limit theorem, which holds under the assumption that growth events become more independent over time (Bottazzi & Secchi 2006a). To test different time spans, we constructed the growth rates for a one year up to an eight year time-lag, starting always with the employment numbers of 1999. Fig. 4 shows that the shape parameter b for both sides increases slightly with time, but remains still far from the normal value 2. Rather the median values of b_{left} and b_{right} approach the Laplace distribution. In summary, extreme growth events are prominent features of regional industry-specific growth, independent of time scales, industrial aggregation levels or regional definitions.

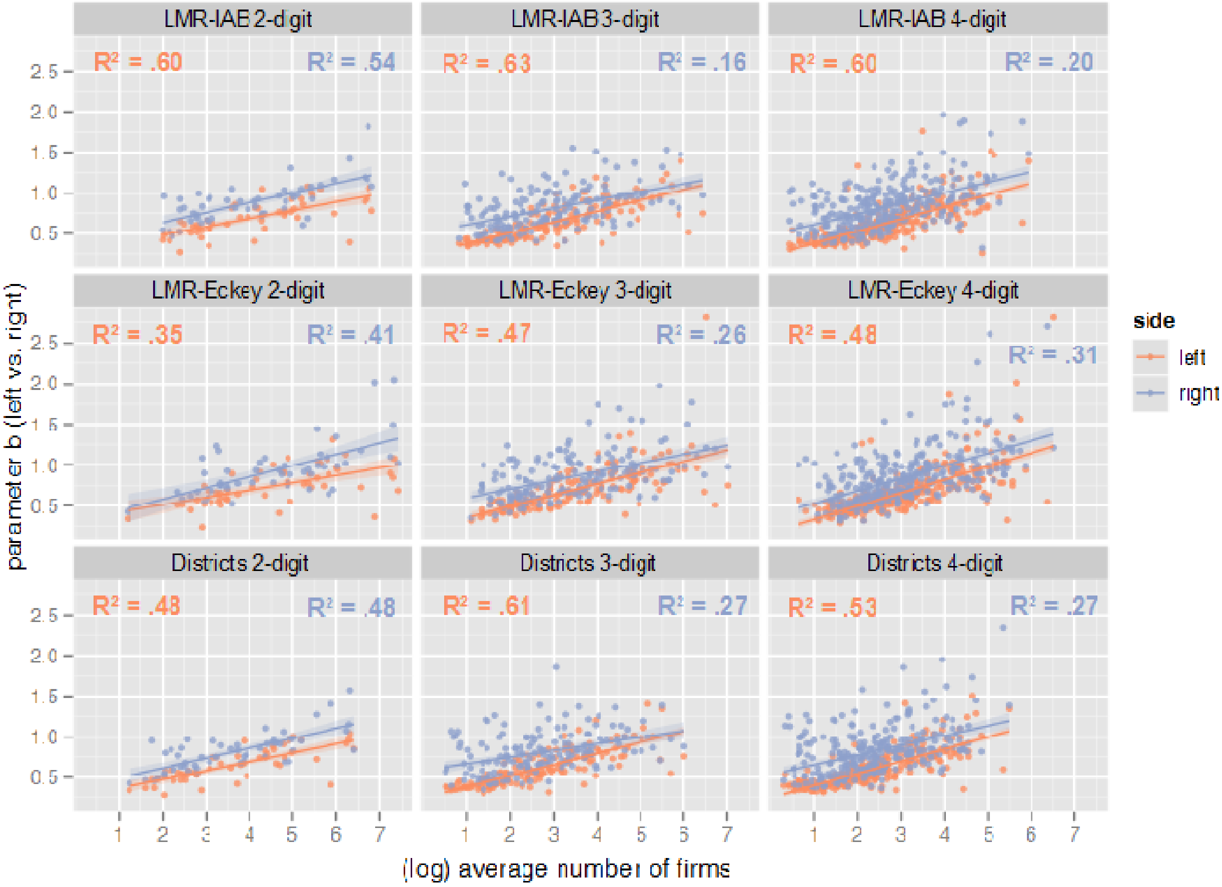
Figure 4 Dependence of the shape parameters b_{left} and b_{right} (and their confidence bands) on the time lag for $g_{i,r,t}$



5.2 The link between the micro- and the meso-level

In hypothesis 2a it was expected that the higher the number of firms of an industry is, the faster the distribution of the regional growth rates approaches to a normal shape. Since the Subbotin distribution gradually moves onto a normal one when the shape parameter b assumes 2, we plotted in Fig. 5 the (log) average number of firms of an industry against b_{left} as well as b_{right}.

Figure 5 Relationship between the log number of firms and the shape parameters b_{left} and b_{right}



For all levels of investigation a strong positive correlation is present, with the explained variance (R^2) in a simple linear regression model ranging from 0.16 to 0.63. The revealed relationship between the tail behaviour and the number of firms of an industry supports the idea that the observed pattern on the meso-level of regions to a considerable degree is the result of aggregation of all micro-level firm growth rate distributions within a single region. R^2 is in general lower for the right tail of the distributions, suggesting that factors other than the number of firms of an industry are more important in explaining positive growth events. Moreover, it can be observed that even for most of those industries with the highest average regional firm numbers between a few hundreds up to almost 1600, b_{left} and b_{right} stay visibly below the value of the normal distribution. Fat tails remain due to some underlying correlating mechanisms, which therefore must work at the regional level. Further investigation is needed to disentangle systematically both the effects stemming from pure aggregation of the micro entities and from correlating mechanisms working at the higher level of regions.

5.3 Relation between type of industry and stochastic characteristics of regional growth

To assess the relationship between the stochastic characteristics of regional employment growth and the type of industry, we grouped all industries according to the following classifications:

- *Economic sector*: the secondary or manufacturing sector (industries from WZ 15 to 45) can be distinguished from the tertiary or service sector (WZ 50 to 99). The primary sector was excluded because it contains too little observations.
- *Knowledge intensity*: all industries can be classified whether they are knowledge intensive or non-knowledge intensive. The assignment here is based on Legler & Frietsch (2006) with 3-digit industries as the highest level of disaggregation.
- *Pavitt taxonomy*: four categories can be differentiated on the 2-digit level, namely science-based, specialised suppliers, scale intensive and supplier dominated industries. We used the revised Pavitt taxonomy by Bogliacino & Pianti (2010), because they also address service industries (up to WZ 74).

Using these classification schemes, we then follow a two-step procedure. First, we regress the estimated distributional parameters b_{left}, b_{right}, a_{left} and a_{right} of the asymmetric Subbotin with the average (log) number of firms. This is necessary because the average (log) number of firms is both correlated with the distributional parameters (see the previous section) and the types of industries. For this purpose, a robust MM-type linear regression is employed, since heteroscedasticity and a non-normal distributed error-term, mainly due to the effect of outliers in the dependent variables, are present (see Tab. X3 in the Appendix). We are not interested in the standard results of this regression but in the residuals for each industry. The residuals provide information about whether each industry has higher or lower parameter values b_{left}, b_{right}, a_{left} and a_{right} than average. In a second step, we test whether the groups of our various industrial classifications show significant deviations in the parameters from the whole economy's average. To this end, we perform the Kruskal-Wallis rank sum test of the null hypothesis that the median values are equal in each group (Kruskal & Wallis 1952). This test can be considered as a non-parametric alternative to the single-factor analysis of variance (Bortz et al. 2008: 222). Conditional median values of the residuals and significant differences are reported in Tab. 3a for the manufacturing and service sectors, in Tab. 3b for the knowledge intensity classification scheme, and in Tab. 3c for the revised Pavitt taxonomy.

Table 3a Conditional median values of the residuals and significant differences (Kruskal-Wallis test)

	Economic sector	$\sim a_{left}$	$\sim a_{right}$	$\sim b_{left}$	$\sim b_{right}$
LMR-IAB 2-digit	*manufacturing*	0.000	0.002	0.007	-0.029
	service	0.000	0.001	-0.021	0.033
LMR-IAB 3-digit	*manufacturing*	-0.001	-0.001*	-0.008	-0.044
	service	0.000	0.006*	0.000	-0.001
LMR-IAB 4-digit	*manufacturing*	-0.002	-0.004**	-0.009	-0.043
	service	0.003	0.003**	0.005	0.026
LMR-Eckey 2-digit	*manufacturing*	0.004	0.000	-0.003	-0.007
	service	-0.002	0.001	-0.056	0.005
LMR-Eckey 3-digit	*manufacturing*	-0.001	0.000*	-0.005	-0.044
	service	0.000	0.002*	-0.022	-0.005
LMR-Eckey 4-digit	*manufacturing*	-0.001	-0.001	-0.012	-0.038
	service	-0.001	-0.001	-0.009	0.011
Districts 2-digit	*manufacturing*	-0.002	-0.001	0.012	-0.021
	service	0.001	0.000	-0.017	-0.003
Districts 3-digit	*manufacturing*	-0.002	-0.006***	-0.003	-0.06*
	service	0.003	0.005***	-0.003	0.031*
Districts 4-digit	*manufacturing*	-0.004	-0.006**	-0.015	-0.058
	service	0.001	0.003**	-0.006	0.009

Note: *** p<0.001, ** p<0.01 and * p<0.05

Table 3b Conditional median values of the residuals and significant differences (Kruskal-Wallis test)

	Knowledge intensity	$\sim a_{left}$	$\sim a_{right}$	$\sim b_{left}$	$\sim b_{right}$
LMR-IAB 2-digit	*non knowledge intensive*	0.000	-0.001*	0.016*	0.021
	knowledge intensive	0.000	0.004*	-0.038*	-0.059
LMR-IAB 3-digit	*not knowledge intensive*	-0.001	-0.001	0.007*	0.012**
	knowledge intensive	0.001	0.005	-0.036*	-0.117**
LMR-IAB 4-digit	*not knowledge intensive*	0.001	0.000	0.013**	0.014**
	knowledge intensive	-0.002	0.001	-0.043**	-0.138**
LMR-Eckey 2-digit	*not knowledge intensive*	0.001	-0.001	0.044***	-0.007
	knowledge intensive	-0.002	0.004	-0.072***	-0.024
LMR-Eckey 3-digit	*not knowledge intensive*	0.000	-0.001**	0.005**	-0.004
	knowledge intensive	0.001	0.005**	-0.042**	-0.077
LMR-Eckey 4-digit	*not knowledge intensive*	-0.001	-0.001	0.007**	0.007*
	knowledge intensive	-0.003	-0.001	-0.034**	-0.094*
Districts 2-digit	*not knowledge intensive*	0.002	-0.002	0.025	0.027*
	knowledge intensive	-0.003	0.003	-0.030	-0.045*
Districts 3-digit	*not knowledge intensive*	-0.001	-0.002	0.011**	0.020**
	knowledge intensive	0.000	0.002	-0.047**	-0.12**
Districts 4-digit	*not knowledge intensive*	0.001	0.000	0.006**	0.018***
	knowledge intensive	-0.010	-0.003	-0.042**	-0.133***

Note: *** p<0.001, ** p<0.01 and * p<0.05

Table 3c Conditional median values of the residuals and significant differences (Kruskal-Wallis test)

	Revised Pavitt taxonomy	**~ a_{left}**	**~ a_{right}**	**~ b_{left}**	**~ b_{right}**
LMR-IAB 2-digit	*supplier dominated*	0.002**	-0.001*	0.007	0.035
	scale intensive	-0.012**	-0.009*	-0.011	0.003
	specialized suppliers	0.001**	0.001*	-0.070	-0.100
	science based	0.044**	0.039*	-0.055	-0.003
LMR-IAB 3-digit	*supplier dominated*	0.000	-0.003***	0.007	0.022
	scale intensive	-0.008	-0.009***	0.000	-0.084
	specialized suppliers	0.004	0.006***	-0.022	-0.084
	science based	0.026	0.035***	-0.039	-0.105
LMR-IAB4-digit	*supplier dominated*	0.000	-0.001**	0.016*	0.031
	scale intensive	-0.008	-0.009**	0.009*	-0.031
	specialized suppliers	0.000	0.004**	-0.015*	-0.098
	science based	0.036	0.030**	-0.04*	-0.134
LMR-Eckey 2-digit	*supplier dominated*	0.005**	-0.002*	0.043	0.008
	scale intensive	-0.009**	-0.007*	-0.013	-0.029
	specialized suppliers	0.002**	0.003*	-0.036	-0.044
	science based	0.038**	0.028*	-0.097	0.020
LMR-Eckey 3-digit	*supplier dominated*	-0.001	-0.001***	0.005	0.045
	scale intensive	-0.006	-0.010***	0.003	-0.065
	specialized suppliers	0.001	0.005***	-0.022	-0.036
	science based	0.013	0.023***	-0.036	-0.074
LMR-Eckey 4-digit	*supplier dominated*	-0.002	-0.001**	-0.003*	0.033*
	scale intensive	-0.007	-0.004**	0.023*	-0.032*
	specialized suppliers	0.001	0.005**	-0.029*	-0.075*
	science based	0.013	0.020**	-0.048*	-0.087*
Districts 2-digit	*supplier dominated*	0.003**	0.000*	0.021	0.060*
	scale intensive	-0.011**	-0.012*	0.007	-0.064*
	specialized suppliers	-0.005**	0.003*	-0.017	-0.065*
	science based	0.042**	0.025*	-0.057	0.007*
Districts 3-digit	*supplier dominated*	0.002*	-0.004**	0.009	0.043*
	scale intensive	-0.012*	-0.006**	-0.009	-0.043*
	specialized suppliers	0.000*	0.000**	0.004	-0.096*
	science based	0.044*	0.031**	-0.044	-0.057*
Districts 4-digit	*supplier dominated*	0.000**	-0.001**	0.012*	0.026**
	scale intensive	-0.013**	-0.006**	-0.007*	-0.021**
	specialized suppliers	-0.001**	0.000**	-0.011*	-0.106**
	science based	0.038**	0.034**	-0.044*	-0.093**

Note: *** p<0.001, ** p<0.01 and * p<0.05

Hypothesis 3 can be confirmed partially – some stochastic characteristics of regional employment growth depend on the type of industry. The general dispersion of the growth rates as measured by the parameter a_{left} and a_{right} differs across the several industrial classification schemes. Service industries (in five out of nine cases significant) and to a lesser degree knowledge intensive industries (in two out of nine cases significant) tend to be accompanied by a higher dispersion of their above the average growth rates than corresponding manufacturing industries and non-knowledge intensive industries. No significant differences are found for a_{left} for these two classification schemes. Grouped according to Pavitt's taxonomy, a higher heterogeneity of the industries can be observed – science based industries show the highest dispersion and scale intensive industries the lowest dispersion of growth rates on both sides of the distribution. Hence, we find that knowledge-intensive, science based and service industries show higher fluctuations in their regional growth dynamics. For science based industries this holds for growth as well as decline. In the case of knowledge-intensive and service industries the higher fluctuation is only given for positive growth events.

Further interesting characteristics of the growth process reside in the probability of a regional economy to be affected by extreme growth events, which is expressed by the fat tails of the distribution. Manufacturing and service industries show in general no significant differences in their tail behaviours. Results for Pavitt's taxonomy are mixed. Negative fat tails seem to be more pronounced in science based industries and less pronounced in supplier dominated and scale intensive industries, whereas positive fat tails are more present in specialized suppliers industries and less present in supplier dominated industries. Hence, especially supplier dominated industries consistently indicate a tendency to be less affected by extreme growth events. Finally, as strongly robust evidence it is found that knowledge-intensive industries have significantly fatter tails than non-knowledge intensive industries. In other words, industries where knowledge and innovation assume an important role in production are more exposed to experience extreme negative as wells as extreme positive growth events.

6 Conclusions

In this paper we find that the asymmetric Subbotin distribution is the theoretical distribution that describes best industry-specific employment growth in regions. The estimated distributional parameters show that the growth rates follow a tent-like shape with tails significantly fatter than the normal and Laplace distribution across all industries analysed. Despite the aggregate nature of regions, extreme growth events are more probable to occur than could be expected under a *normal* data-generating process. A comparison of both sides of the distribution revealed that particularly extreme negative growth events are a prominent feature of regional growth dynamics. This “rich statistical structure in dynamics” (Dosi et al. 2010: 1872) of economic entities like regions have important implications for both theory and modelling.

Economic theory must go beyond simply focusing on and explaining the average dynamics of regional growth. In an evolutionary context, persistent interactions between firms as the actual economic agents lead to an observed non-normal behaviour at higher levels of the economy (Dosi et al. 2010). The pronounced extreme negative growth events clearly demonstrate that regional economies operate in a decade of multiple crises. The necessity of a shift in the focus from the mean to the tails of the distribution is reflected by recent advances in the regional development literature, where concepts like regional vulnerability and resilience have been introduced (e.g. Pike et al. 2010; Simmie & Martin 2010). If stability in regional growth is the aim of policy, a better understanding of the causes of extreme events and the ability of regional economies to cope with them is of high relevance.

Explaining regional economic growth with more realistic models is a further major challenge for economic geographers and regional scientists. Statistical inference from economic models based on the assumption that the observations are drawn from a normal distribution may deliver invalid implications (Fagiolo et al. 2008). Since normal distributed errors are often inappropriate to study natural phenomena (Koenker & Bassett 1978), two approaches seem to be promising for explaining regional growth. On the one hand, new insights into the dynamics of growth can be gained by looking at the entire shape of the distribution via quantile regression techniques, as suggested by Coad (2007) or Reichstein et al. (2010). On the other hand, the estimation of the models can be performed directly on the assumption of fat-tailed errors (Mineo 2003; Fagiolo et al. 2008).

Finally, in an attempt to explain the emergence of the stochastic properties of regional industry-specific employment growth we find that the economic micro-structure in terms of the number of firms of an industry matter. The observed patterns at the meso-level of regions are to a considerable degree the result of the aggregation of all micro-level firm growth rates distributions within a single region. However, even the growth rates of those industries, which encompass the highest numbers of firms, are far from being normal distributed. The observed fat tails have to be attributed to some underlying correlating and self-reinforcing mechanisms working at the meso-level of regions. Here, the remaining variance of the distributional parameters of the industry-specific growth rates are put into relation to various industrial classification schemes. The most clear-cut result is found for the shape parameter *b*: knowledge-intensive industries are significantly more exposed to experience extreme negative as wells as extreme positive growth events. Many further potential economic mechanisms (Alfarano & Milkovic 2008: 274) may be able to explain the emergence of fat-tailed distributions at the regional level. More light could be shed on the determinants of the distributional characteristics by explicitly taking into account phenomena like industry life cycles, demand shocks, global integration, the relative frequency of lumpy growth events like the entry of new or the exit of existing firms, or the degree of importance of innovative and therefore inherently disturbing activities. Besides the before mentioned industry-specific factors, spatial measures as direct proxies for correlation mechanisms at the level of regions may have explanatory power with respect to the distributional parameters. The strength and configuration of spatial interactions between regions, the structural characteristics of regions (for instance, urban versus rural areas), or the extend of spatial concentration of innovative activities (Fornahl & Brenner 2009) are expected to influence the dispersion of regional industry-specific employment growth as well as the exposure of regions to experience extreme growth events.

Literature

Agro G (1995) Maximum likelihood estimation for the exponential power function parameters. *Communications in Statistics-Simulation and Computation* 24: 523-536
Alfaranao S, Milakovic M (2008) Does classical competition explain the statistical features of firm growth? *Economic Letters* 101: 272-274
Amaral L, Buldyrev S, Havlin S, Maass P, Salinger M, Stanley E, Stanley M (1997) Scaling behavior in economics: The problem of quantifying company growth. *Physica A* 244: 1-24
Amaral L, Gopikrishnan P, Plerou V, Stanley E (2001) A model for the growth dynamics of economic organizations. *Physica A* 299: 127-136
Anscombe F, Glynn W (1983) Distribution of kurtosis statistic for normal statistics. *Biometrika* 70: 227-234
Axtell R (2001) Zipf distribution of US firm sizes. *Science* 293: 1818-1820
Binder J, Schwengler B (2006) Neuer Gebietszuschnitt der Arbeitsmarktregionen im Raum Berlin und Brandenburg. Kritische Überprüfung der bisher gültigen Arbeitsmarktregionen und Vorschläge für einen Neuzuschnitt. *IAB-Forschungsbericht* 4/2006, Nürnberg
Bottazzi G, Cefis E, Dosi G (2002) Corporate Growth and Industrial Structure. Some Evidence from the Italian Manufacturing Industry. *Industrial and Corporate Change* 11: 705-723
Bottazzi G, Coad A, Jacoby N, Secchi A (2011) Corporate growth and industrial dynamics: Evidence from French manufacturing. *Applied Economics* 43: 103-116
Bottazzi G, Dosi G, Fagiolo G, Secchi A (2008) Sectoral and geographical specificities in the spatial structure of economic activities. *Structural Change and Economic Dynamics* 19: 189-202
Bottazzi G, Secchi A (2003) Common Properties and Sectoral Specificities in the Dynamics of U.S. Manufacturing Companies. *Review of Industrial Organization* 23: 217-232
Bottazzi G (2004) Subbotools user's manual. Sant'Anna School of Advanced Science: *LEM Working Paper Series* 2004/14
Bottazzi G, Secchi A (2006a) Explaining the distribution of firm growth rates. *RAND Journal of Economics* 37: 235-256
Bottazzi G, Secchi A (2006b) Gibrat's Law and diversification. *Industrial and Corporate Change* 15: 847-875
Bottazzi G, Secchi A (2011) A new class of asymmetric exponential power densities with application to economics and finance. *Industrial and Corporate Change*, Advance Access, July 4, 2011
Bogliacini F, Pianta M (2010) Innovation and Employment: a Reinvestigation using Revised Pavitt classes. *Research Policy* 39: 799-809
Bortz J, Lienert G, Boehnke K (2008) *Verteilungsfreie Methoden in der Biostatistik*. Springer, Heidelberg
Brenner T (2004) *Local Industrial Clusters - Existence, Emergence, and Evolution*. Routledge, London
Brenner T (2006) Identification of Local Industrial Clusters in Germany. *Regional Studies* 40: 991-1004
Brock W (1999) Scaling in Economics: a Reader's Guide. *Industrial and Corporate Change* 8: 409-446
Canning D, Amaral, Lee Y, Meyer M, Stanley E (1998) Scaling the volatility of GDP growth rates. *Economics Letters* 60: 335-341
Castaldi C, Sapio S (2008) Growing like mushrooms? Sectoral evidence from four large European economies. *Journal of Evolutionary Economics* 18: 509-527
Castaldi C, Dosi G (2009) The patterns of output growth of firms and countries: Scale invariances and scale specificities. *Empirical Economics* 37: 475-495
Cheshire P, Malecki E (2004) Growth, development, and innovation: A look backward and forward. *Papers in Regional Science* 83: 249-267
Coad A (2007) A Closer Look at Serial Growth Rate Correlation. *Review of Industrial Organization* 31: 69-82
Döring T, Schnellenbach J (2006) What Do We Know about Geographical Knowledge Spillovers and Regional Growth? A Survey of the Literature. *Regional Studies* 40: 375-395
Dosi G (1988) Sources, Procedures, and Microeconomic Effects of Innovation. *Journal of Economic Literature* 26: 1120-1171
Dosi G, Lechevalier S, Secchi A (2010) Introduction: Interfirm heterogeneity - nature, sources and consequences for industrial dynamics. *Industrial and Corporate Change* 19: 1867-1890
Dumais G, Ellison G, Glaeser E (2002) Geographic concentration as a dynamic process. *Review of Economics and Statistics* 84: 193-204
Duschl M, Schimke S, Brenner T, Luxen D (2011): Firm Growth and the Spatial Impact of Geolocated External Factors: Empirical Evidence for Germany Manufacturing Firms. Marburg Working Paper on Innovation and Space 03/11.
Eckey H, Kosfeld R, Türck M (2006) Abgrenzung deutscher Arbeitsmarktregionen. *Raumforschung und Raumordnung* 64: 299-309
Ellison G, Glaeser E (1997) Geographic concentration in U.S. manufacturing industries: a dartboard approach. *The Journal of Political Economy* 105: 889-927

Evans D (1987) Tests of alternative theories of firm growth. *The Journal of Political Economy* 95: 657-674
Fagiolo G, Napoletano M, Roventini A. (2008) Are output growth-rate distributions fat-tailed? Some evidence from OECD countries. *Journal of Applied Econometrics* 23: 639-669
Fornahl D, Brenner T. (2009) Geographic concentration and innovative activities in Germany. *Structural Change and Economic Dynamics* 20: 163-182
Fu D, Pammolli F, Buldyrev S, Riccaboni M, Matia K, Yamaski K, Stanley E. H. (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
Gibrat R (1931) *Les Inégalités Économiques*. Recueil Sirey, Paris
Hall B (1988) The relationship between firm size and firm growth in the US manufacturing sector. *NBER Working Paper* No. w1965
Hart P, Prais S (1956) The analysis of business concentration: a statistical approach. *Journal of the Royal Statistical Society. Series A* 119: 150-191
Havemann F, Heinz M, Wagner-Döbler R (2005) Firm-Like Behavior of Journals? Scaling Properties of Their Output and Impact Growth Dynamics. *Journal of the American Society for Information Science and Technology* 5: 3-12
Hymer S, Pashigian P (1962) Firm size and rate of growth. *The Journal of Political Economy* 70: 556-569
Jones C (1997) In the Evolution of the World Income Distribution. *Journal of Economic Perspectives* 11: 19-36
Jarque C, Bera A (1980) Efficient test for normality, homoscedasticity and serial independence of residuals. *Economic Letters* 6: 255-259
Koenker R, Bassett G (1978) Regression quantiles. *Econometrica* 26: 33-50
Krugman P (1991) Increasing Returns and Economic Geography. *The Journal of Political Economy* 99: 483-399
Lee Y, Amaral L, Canning D, Meyer M, Stanley E (1998) Universal Features in the Growth Dynamics of Complex Organizations. *Physical Review Letters* 81: 3275-3278
Legler H, Frietsch R (2006) Neuabgrenzung der Wissenswirtschaft –forschungsintensive Industrien und wissensintensive Dienstleistungen. In: BMBF (eds): *Studien zum deutschen Innovationssystem*, 22-2007
Lucas R (1978) On the size distribution of business firms. *The Bell Journal of Economics* 9: 508-523.
Maasoumi E, Racine J, Stengos T (2007) Growth and convergence: A profile of distribution dynamics and mobility. *Journal of Econometrics* 136: 483-508
Mansfield E (1962) Entry, Gibrat's law, innovation, and the growth of firms. *The American Economic Review* 52: 1023-1051
Marshall A (1890) *Principles of Economics*. Macmillan, London
Mineo A (2003) On the estimation of the structure parameter of a normal distribution of order p. *Statistica anno* LXIII: 109-122
Mittelhammer R (1995) *Mathematical Statistics for Economics and Business*. Springer, New York
Pavitt K (1984) Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy* 13: 343-373
Pike A, Dawley S, Tomaney J (2010) Resilience, adaptation and adaptability. *Cambridge Journal of Regions, Economy and Society* 3: 59–70
Kruskal W, Wallis W (1952) Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 47: 583-621
Quah D (1996) Twin Peaks: Growth and Convergence in Models of Distribution Dynamics. *The Economic Journal* 106: 1045-1055
Quah D (1997) Empirics of Growth and Distribution: Stratification, Polarization, and Convergence Clubs. *Journal of Economic Growth* 2: 27-59
Reichstein T, Jensen M (2005) Firm size and firm growth rate distributions - The case of Denmark. *Industrial and Corporate Change* 14: 1145-1166
Reichstein T, Dahl M, Ebersberger B, Jensen M (2010) The devil dwells in the tails. *Journal of Evolutionary Economics* 20: 219-231
Sapio S, Thoma G. (2006) The Growth of Industrial Sectors: Theoretical Insights and Empirical Evidence from U.S. Manufacturing. Sant'Anna School of Advanced Science: LEM Working Paper Series 2006/09
Schlump C, Brenner T (2010) University Education, Public Research and Employment Growth in Regions - An Empirical Study of Germany. *Ber. z. dt. Landeskunde* 84: 115-136
Simon H, Bonini C (1958) The size distribution of business firms. *The American Economic Review* 46: 607-617
Simmie J, Martin R (2010) The economic resilience of regions: towards an evolutionary approach. *Cambridge Journal of Regions, Economy and Society* 3: 27-43
Smith K (2002) What is the 'Knowledge Economy'? Knowledge Intensity and Distributed Knowledge Bases. *UNU/INTECH Discussion Papers* 2002-6
Stanley M, Amaral L, Buldyrev S, Havlin S, Leschhorn H, Maass P, Salinger M, Stanley E (1996) Scaling behaviour in the growth of companies. *Nature* 379: 804-806

Appendix

Table X1 Temporal autocorrelation and dependence of regional growth rates on employment number – median Spearman *rho* coefficient

	Temporal autocorrelation[1]	**Dependence on employment number[2]**
LMR-IAB 2-digit	0.11	-0.09
LMR-IAB 3-digit	0.08	-0.09
LMR-IAB 4-digit	0.06	-0.10
LMR-Eckey 2-digit	0.13	-0.06
LMR-Eckey 3-digit	0.09	-0.08
LMR-Eckey 4-digit	0.06	-0.10
Districts 2-digit	0.10	-0.08
Districts 3-digit	0.07	-0.09
Districts 4-digit	0.05	-0.10

Notes: [1] Spearman correlation between each consecutive year pair in each industry (median of rho is reported)
[2] Spearman correlation between the employment number and the growth rates in each industry (median of rho is reported)

Table X2 Normality tests for $\tilde{g}_{r,i}$ – relative frequencies of significant rejections of the null hypothesis at 0.05 significance level

	Kolmogorov-Smirnov	**Jarque-Bera**	**Anscombe-Glynn**
LMR-IAB 2-digit	100 %	100 %	100 %
LMR-IAB 3-digit	99.5 %	100 %	100 %
LMR-IAB 4-digit	99.4 %	100 %	100 %
LMR-Eckey 2-digit	100 %	100 %	100 %
LMR-Eckey 3-digit	99.5 %	100 %	100 %
LMR-Eckey 4-digit	99.0 %	100 %	100 %
Districts 2-digit	100 %	100 %	100 %
Districts 3-digit	100 %	100 %	100 %
Districts 4-digit	99.8 %	100 %	100 %

Table X3 *p-values* for Kolmogorov-Smirnov normality test on residuals and Breusch-Pagan test for heteroscedasticity for the linear models of the number of firms against the distributional parameters

	KS normality test on residuals				**BP test for heteroscedasticity**			
	~ a_{left}	**~ a_{right}**	**~ b_{left}**	**~ b_{right}**	**~ a_{left}**	**~ a_{right}**	**~ b_{left}**	**~ b_{right}**
LMR-IAB 2-digit	0.01*	0.00*	0.32	0.38	0.33	0.74	0.06	0.09
LMR-IAB 3-digit	0.00*	0.00*	0.05	0.14	0.37	0.49	0.00*	0.30
LMR-IAB 4-digit	0.00*	0.00*	0.00*	0.06	0.03*	0.63	0.00*	0.00*
LMR-Eckey 2-digit	0.00*	0.00*	0.48	0.08	0.24	0.49	0.03*	0.01*
LMR-Eckey 3-digit	0.00*	0.00*	0.01*	0.01*	0.71	0.75	0.00*	0.01*
LMR-Eckey 4-digit	0.00*	0.00*	0.00*	0.00*	0.77	0.47	0.00*	0.00*
Districts 2-digit	0.11	0.00*	0.48	0.17	0.77	0.94	0.06	0.06
Districts 3-digit	0.00*	0.00*	0.18	0.09	0.91	0.76	0.00*	0.42
Districts 4-digit	0.00*	0.00*	0.04	0.00*	0.17	0.91	0.00*	0.10

Note: * null hypothesis rejected at 0.05 significance level