An improvement over the normal distribution for log-growth rates of city sizes: Empirical evidence for France, Germany, Italy and Spain

Puente-Ajovin, Miguel and Ramos, Arturo

Universidad de Zaragoza

27 October 2015

An improvement over the normal distribution for log-growth rates of city sizes: Empirical evidence for France, Germany, Italy and Spain

MIGUEL PUENTE-AJOVÍN[a] ARTURO RAMOS[b]

October 27, 2015

Abstract

We study the decennial log-growth population rate distributions of France (1990-2009), Germany (1996-2006), Italy (1951-1961, 2001-2011) and Spain (1950-1960, 2001-2010).

It is obtained an excellent parametric description of these log-growth rates by means of a modification of the normal distribution in that the tails are mixed by means of convex linear combinations with exponential distributions, giving rise to the so called "double mixture exponential normal".

The normal distribution is not the one empirically observed for the same datasets.

JEL: C46, R11, R12.

Keywords: urban log-growth rates distribution, exponential distribution, normal distribution, European population log-growth rates

[a]Department of Economic Analysis, Universidad de Zaragoza (SPAIN) mpajovin@gmail.com. We thank constructive comments from Fernando Sanz-Gracia on a previous version of the manuscript, although all remaining errors are ours. This work is supported by the project ECO2013-45969-P of the Spanish Ministry of Economy and Competitiveness and by the Aragon Government, ADETRE Consolidated Group.

[b]Department of Economic Analysis, Universidad de Zaragoza (SPAIN) aramos@unizar.es.

1 Introduction

The parametric study of the log-growth rates of city size distribution has received traditionally scarce attention in the Urban Economics literature. However, in recent times have appeared the studies of Schluter and Trede (2013); Ramos (2015) where the cases of Germany and the US, respectively, have been treated. In the second of these references it has been found that an excellent description of three different types of US log-growth data is obtained with the so called "double mixture exponential Generalized Beta 2 (dmeGB2)" distribution. This distribution is the exponential version of the "dmPGB2" of Ramos and Sanz-Gracia (2015). Since we worked in Puente-Ajovín and Ramos (2015) with population data of four major European countries, namely France, Germany, Italy and Spain, it seemed to be natural to study the log-growth rates of these last countries.

However, we have followed in this article a different line of reasoning in order to obtain the final results. That is, we have tried to check whether the normal distribution is a good description of the log-growth rates of the cited European countries. If not, try to modify the normal distribution in a minimal way in order to obtain a good enough parametric description of the mentioned quantities. For that, we rely on previous studies stating that the tails of the log-growth distributions (for firm size) are approximately exponential (Johnson et al., 1995; Stanley et al., 1996; Canning et al., 1998; Bottazzi and Secchi, 2003, 2011) and we also take the idea of convex linear combinations of distributions at the tails (Combes et al., 2012) to obtain a distribution that we call "double mixture exponential normal (dmen)". We will see that this parametric model is not rejected empirically for all of the studied samples, contrary to the ordinary normal distribution. We also checked the best model of Ramos (2015) and found that for the studied European countries it does not lead to a real improvement over the dmen, which means that the log-growth processes of the US and of the studied European countries do differ in practice.

The fact that the log-growth size distributions is not normal has implications with regards to the standard Gibrat's Law, see Ramos (2015) for details.

We use here part of the same databases of Puente-Ajovín and Ramos (2015), see therein for details, and we show in Table 1 the descriptive statistics of the used data for France, Germany, Italy and Spain. For France, we have used the last three available Census. For Germany, only the sample (1996-2006) is used due to the difficulty of constructing the log-growth rates for other years, since the German urban units do change notably in other periods due to mergers and/or splits. Out of the samples for Italy and Spain, we have used the last available ones, and those of mid-century (Italy, 1951-1961; Spain 1950-1960) in which the mean log-growth rates are negative.

[Table 1 near here]

The rest of the article is organized as follows. Section 2 introduces the parametric distributions used in this paper. Section 3 describes the empirical results obtained. Finally, Section 4 concludes.

2 Description of the presented distributions

In this section we will introduce the distributions used along the paper[1] for the (two consecutive periods) log-growth rates, denoted by

$$g_{i,t} = \log x_{i,t} - \log x_{i,t-1} \in (-\infty, \infty)$$

[1]From a practical point of view, it is our interest in this paper to obtain a very good parametric fit of the log-growth rate distributions of the studied European countries. For that, we have first tried several distributions well-known in the economics literature: the normal, the asymmetric exponential power (AEP) of Bottazzi and Secchi (2011), which generalizes the Laplace distribution of, e.g., Johnson et al. (1995); Stanley et al. (1996); Canning et al. (1998); Bottazzi and Secchi (2003) and references therein, the α-stable distribution, see, e.g., Zolotarev (1986); Uchaikin and Zolotarev (1999); Gaffeo (2011) and references therein (the calculations for the α-stable distribution have been performed using the STABLE software of Robust Analysis Inc., see `http://www.robustanalysis.com/`) the generalized hyperbolic distribution (Barndorff-Nielsen (1977); Barndorff-Nielsen and Halgreen (1977); Barndorff-Nielsen and Stelzer (2005)), the normal-Laplace distribution of Reed (2003); Reed and Jorgensen (2004); Manas (2009) and the (non-standardized) Student-t distribution, see, e.g., Johnson et al. (1995) and references therein. The results for the distributions, which are not those with the best performance, and not presented here are available from the authors upon request.

where $x_{i,t}$ is the population of city i at time t. When a fixed t is taken we will simply write $g \in (-\infty, \infty)$ for the variable of all log-growth rates of the cross-sections taken.

2.1 Normal distribution

Firstly, we recall the normal distribution for the log-growth rates g. The probability density distribution (pdf) and cumulative distribution function (cdf) are, respectively,

$$\begin{aligned} f_{\mathrm{n}}(g,\mu,\sigma) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(g-\mu)^2}{2\sigma^2}\right) \\ \mathrm{cdf}_{\mathrm{n}}(g,\mu,\sigma) &= \frac{1}{2}\left(1+\mathrm{erf}\left(\frac{g-\mu}{\sqrt{2}\sigma}\right)\right) \end{aligned} \quad (1)$$

where μ is real and $\sigma > 0$ are the mean and the standard deviation of the variable g according to this distribution. Also, erf is the error function associated to the standard normal distribution.

2.2 The double mixture exponential normal (dmen)

For our new distribution "double mixture exponential normal (dmen)" we first define some basic functions which will be employed by the former.

Then, let us consider

$$\begin{aligned} u(g,\zeta) &= \exp(-\zeta g) \\ l(g,\rho) &= \exp(\rho g) \end{aligned}$$

The function $u(g,\zeta)$ will model the decreasing exponential part of the upper tail of our new distribution, where $\zeta > 0$, and $l(g,\rho)$ corresponds to the increasing exponential lower tail, with $\rho > 0$. The functions u, l are not normalized at this stage like in Ioannides and Skouras (2013); Ramos (2015). Note that if the variable x follows a Pareto distribution and $y = \ln x$, then y follows an exponential distribution.

The new distribution we introduce here, has two tails which are exponential with a convex linear combination mixture of normal, and body of this last type. The switch between the tails and the body occurs at two exact thresholds ϵ (lower tail-body) and $\tau > \epsilon$ (body-upper tail). For the lower tail, the combining coefficient will be denoted by $\nu \in (0,1)$, and by $\theta \in (0,1)$ for the upper tail. We require continuity of the density function at the threshold points and overall normalization to one. They are also imposed equal weight of the distributions of the mixing at the tails, like in Ioannides and Skouras (2013), in order that the parameters ν, θ control the proportion of each component of the combination in the lower (resp. upper) tail.

The resulting composite density is given by:

$$
\begin{aligned}
& f_{\mathrm{dmen}}(g,\rho,\epsilon,\nu,\mu,\sigma,\tau,\zeta,\theta) \\
& = \begin{cases} b_2[(1-\nu)\, d_2\, f_{\mathrm{n}}(g,\mu,\sigma) + \nu\, e_2\, l(g,\rho)] & g < \epsilon \\ b_2\, f_{\mathrm{n}}(g,\mu,\sigma) & \epsilon \le g \le \tau \\ b_2[(1-\theta)\, c_2\, f_{\mathrm{n}}(g,\mu,\sigma) + \theta\, a_2\, u(g,\zeta)] & \tau < g \end{cases}
\end{aligned}
$$

where the constants are given as follows:

$$
\begin{aligned}
d_2^{-1} &= 1-\nu + \frac{\exp(-\rho\epsilon)\,\nu\,\rho\,\mathrm{cdf}_{\mathrm{n}}(\epsilon,\mu,\sigma)\, l(\epsilon,\rho)}{f_{\mathrm{n}}(\epsilon,\mu,\sigma)} \\
e_2^{-1} &= \frac{(1-\nu)\,\exp(\epsilon\rho)}{\rho\,\mathrm{cdf}_{\mathrm{n}}(\epsilon,\mu,\sigma)} + \frac{\nu\, l(\epsilon,\rho)}{f_{\mathrm{n}}(\epsilon,\mu,\sigma)} \\
c_2^{-1} &= 1-\theta + \frac{\zeta\,\theta\,\exp(\tau\zeta)\,(1-\mathrm{cdf}_{\mathrm{n}}(\tau,\mu,\sigma))\, u(\tau,\zeta)}{f_{\mathrm{n}}(\tau,\mu,\sigma)} \\
a_2^{-1} &= \frac{(1-\theta)\,\exp(-\tau\zeta)}{\zeta\,(1-\mathrm{cdf}_{\mathrm{n}}(\tau,\mu,\sigma))} + \frac{\theta\, u(\tau,\zeta)}{f_{\mathrm{n}}(\tau,\mu,\sigma)} \\
b_2^{-1} &= e_2\,\frac{\exp(\epsilon\rho)}{\rho} + \mathrm{cdf}_{\mathrm{n}}(\tau,\mu,\sigma) - \mathrm{cdf}_{\mathrm{n}}(\epsilon,\mu,\sigma) + \frac{a_2}{\zeta\,\exp(\tau\zeta)}
\end{aligned}
$$

This distribution depends on eight parameters $(\rho,\epsilon,\nu,\mu,\sigma,\tau,\zeta,\theta)$ to be estimated below by Maximum Likelihood (ML). It can be deduced as a limiting case (McDonald, 1984; McDonald and Xu, 1995) of the theoretical model proposed in Ramos (2015), see also Ramos and Sanz-Gracia (2015).

3 Results

In this Section we recall briefly the empirical results concerning the samples of the four European countries studied.

We have computed the log-growth rates between each two consecutive cross-sections of our data. In order to avoid infinite values we have removed the observations for which at least one of the population values is zero. The descriptive statistics of the data so obtained is given in Table 2.

[Table 2 near here]

We have estimated the parameters of the distributions by maximum likelihood estimation (MLE), using the software MATLAB® and MATHEMATICA®. We report on Table 3 the estimated values of the parameters for the dmen and the corresponding standard errors (SE) computed according to Efron and Hinkley (1978) and McCullough and Vinod (2003). The MLE estimators for the normal are exact and equal to the mean and standard deviation (SD) of the data, see simply Table 2, and to be compared to those of Table 4, computed according to the dmen for each studied sample. Observe that these last values are almost identical, meaning that at least the two first moments of the log-growth rate distributions are accurately described by the new dmen.

[Table 3 near here]

[Table 4 near here]

In order to assess the goodness of fit of the two distributions explicitly shown in this paper, we use three standard statistical tests: the Kolmogorov–Smirnov (KS) test, the Crámer–von Mises (CM) test and the Anderson–Darling (AD) test. The results are shown on Table 5. Very briefly, the normal distribution is *strongly* rejected always by the three tests. Meanwhile, the dmen is not rejected in almost 100% of the cases, except in the case of France (1990-1999) and the KS test.

[Table 5 near here]

Also, we have computed the Akaike Information Criterion (AIC) and Bayesian or Schwarz Information Criterion (BIC) (Burnham and Anderson, 2002, 2004), very well adapted to the maximum likelihood estimation we have performed before, see Table 6. Even with the dmen having eight parameters instead of two of the normal, the preferred model amongst these is always the former.

[Table 6 near here]

As a complement of the KS, CM, AD, AIC and BIC criteria, we show in Figure 1 an informal graphical approximation of the obtained fits for the sample of Germany (1996-2006) and the normal on the one hand and the dmen on the other hand. This sample has been selected in order to show the extreme difference of the fits of our studied samples. For the normal (left-hand panel) the fits can be improved notably at the tails and at the body of the distribution, meanwhile for the dmen (right-hand panel) and the sample of Germany (1996-2006) the fit is almost perfect both at tails and body, even accounting for the amplification effect of the discrepancies at the tails (González-Val et al., 2013). Let us remark that on the plots of the tails the cdf for the lower tail or $1 - \mathrm{cdf}$ for the upper tail are nearly exponential, and therefore the graphs (empirical of estimated with the dmen) are almost linear, in agreement with previous knowledge for log-growth of firm sizes (Johnson et al., 1995; Stanley et al., 1996; Canning et al., 1998; Bottazzi and Secchi, 2003, 2011).

[Figure 1 near here]

4 Conclusions

In the preceding Section we have seen that a very appropriate parametric model for the log-growth rate distributions of the city size of France, Germany, Italy and Spain is the newly introduced (in Subsection 2.2) dmen.

The variances given by the dmen in all of our cases of study are finite, so we have found an example of distribution for the log-growth rates of city size for the mentioned European countries, always not rejected empirically and with finite variances. This is an alternative to the normal distribution.

However, the dmen is not the only possibility of describing the log-growth rates of these European countries. If one replaces the normal distribution in the mixing of the dmen by the logistic distribution (Johnson et al., 1995; Kleiber and Kotz, 2003) one obtains very similar results in performance and fit. However, what distinguishes the log-growth rates of the US and that of the studied European countries is that in the former case the dmeGB2 is more appropriate, but for the latter it is not necessary to generalize so much the distribution at the body and it is enough to take the normal distribution or if one prefers, just the logistic. This makes the log-growth process of the US and Europe (at least of the four major countries studied) different, although the city size distribution can be described in similar terms (Puente-Ajovín and Ramos, 2015; Ramos and Sanz-Gracia, 2015). These results are somewhat puzzling and deserve a closer look in further research, maybe taking into account a new look into the economic urban growth literature, see, e.g., Glaeser et al. (1992); Glaeser and Shapiro (2003); Glaeser et al. (2006); Glaeser and Redlick (2009); Duranton and Puga (2014).

References

Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 353(1674):401–419.

Barndorff-Nielsen, O. and Halgreen, C. (1977). Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 38:309–311.

Barndorff-Nielsen, O. and Stelzer, R. (2005). Absolute moments of generalized hyperbolic distributions and approximate scaling of normal inverse Gaussian Lévy processes. *Scandinavian Journal of Statistics*, 32:617–637.

Bottazzi, G. and Secchi, A. (2003). Why are distributions of firm growth rates tent-shaped? *Economics Letters*, 80:415–420.

Bottazzi, G. and Secchi, A. (2011). A new class of asymmetric exponential power densities with applications to economics and finance. *Industrial and Corporate Change*, 20(4):991–1030.

Burnham, K. P. and Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. New York: Springer-Verlag.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, 33:261–304.

Canning, D., Amaral, L. A. N., Lee, Y., Meyer, M., and Stanley, H. E. (1998). Scaling the volatility of GPD growth rates. *Economics Letters*, 60:335–341.

Combes, P. P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012). The productivity advantages of large cities: Distinghishing agglomeration from firm selection. *Econometrica*, 80:2543–2594.

Duranton, G. and Puga, D. (2014). The growth of cities. In Durlauf, S. N. and Aghion, P., editors, *Handbook of Economic Growth*. North-Holland.

Efron, B. and Hinkley, D. V. (1978). Assessing the accuracy of the maximum likelihood estimator: Observed versus expected Fisher information. *Biometrika*, 65(3):457–482.

Gaffeo, E. (2011). The distribution of sectoral TFP growth rates: International evidence. *Economics Letters*, 113:252–255.

Glaeser, E. L., Gyourko, J., and Saks, R. E. (2006). Urban growth and housing supply. *Journal of Economic Geography*, 6(1):71–89.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100(6):1126–1152.

Glaeser, E. L. and Redlick, C. (2009). Social capital and urban growth. *International Regional Science Review*, 32(3):264.

Glaeser, E. L. and Shapiro, J. M. (2003). Urban growth in the 1990s: Is city living back? *Journal of Regional Science*, 43(1):139–165.

González-Val, R., Ramos, A., and Sanz-Gracia, F. (2013). The accuracy of graphs to describe size distributions. *Applied Economics Letters*, 20(17):1580–1585.

Ioannides, Y. M. and Skouras, S. (2013). US city size distribution: Robustly Pareto, but only in the tail. *Journal of Urban Economics*, 73:18–29.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous univariate distributions. Volume 2*. John Wiley & Sons.

Kleiber, C. and Kotz, S. (2003). *Statistical size distributions in Economics and actuarial sciences*. Wiley-Interscience.

Manas, A. (2009). French butchers don't do quantum physics. *Economics Letters*, 103:101–106.

McCullough, B. D. and Vinod, H. D. (2003). Verifying the solution from a nonlinear solver: A case study. *American Economic Review*, 93(3):873–892.

McDonald, J. B. (1984). Generalized functions for the size distribution of income. *Econometrica*, 52(3):647–665.

McDonald, J. B. and Xu, Y. J. (1995). A generalization of the beta distribution with applications. *Journal of Econometrics*, 66:133–152.

Puente-Ajovín, M. and Ramos, A. (2015). On the parametric description of the French, German, Italian and Spanish city size distributions. *The Annals of Regional Science*, 54(2):489–509.

Ramos, A. (2015). Are the log-growth rates of city sizes normally distributed? Empirical evidence for the US. *Working Paper, available at Munich RePec* `http://mpra.ub.uni-muenchen.de/65584/`.

Ramos, A. and Sanz-Gracia, F. (2015). US city size distribution revisited: Theory and empirical evidence. *Working Paper, available at Munich RePec* `http://mpra.ub.uni-muenchen.de/64051/`.

Reed, W. J. (2003). The Pareto law of incomes–an explanation and an extension. *Physica A*, 319:469–486.

Reed, W. J. and Jorgensen, M. (2004). The double Pareto-lognormal distribution–a new parametric model for size distributions. *Communications in Statistics-Theory and Methods*, 33(8):1733–1753.

Schluter, C. and Trede, M. (2013). Gibrat, Zipf, Fisher and Tippet: City size and growth distributions reconsidered. *Working Paper 27/2013 Center for Quantitative Economics WWU Münster*.

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

Uchaikin, V. V. and Zolotarev, V. M. (1999). *Chance and stability. Stable distributions and their applications*. VSP.

Zolotarev, V. M. (1986). *One-dimensional stable distributions*. American Mathematical Society.

Table 1: Descriptive statistics of the data samples used

Sample	Obs.	Mean	SD	Min.	Max.
France 1990	36,644	1,611	14,157	1	2,175,200
France 1999	36,643	1,679	14,173	1	2,147,857
France 2009	36,674	1,793	14,895	1	2,257,981
Germany 1996	12,309	6,663	44,188	3	3,458,763
Germany 2006	12,309	6,687	44,048	7	3,404,037
Italy 1951	8,100	5,866	31,138	74	1,651,393
Italy 1961	8,100	6,250	39,131	90	2,187,682
Italy 2001	8,100	7,021	39,326	33	2,546,804
Italy 2011	8,094	7,490	41,505	34	2,761,477
Spain 1950	7,901	3,480	26,033	64	1,618,435
Spain 1960	7,910	3,802	33,652	51	2,259,931
Spain 2001	8,077	5,039	43,080	7	2,938,723
Spain 2010	8,114	5,795	47,530	5	3,273,049

Table 2: Descriptive statistics of the log-growth rates for the consecutive samples used

Sample	Obs	Mean	SD	Min	Max
Fr 1990-1999	36,643	0.046	0.127	-1.386	1.786
Fr 1999-2009	36,643	0.099	0.150	-2.060	2.692
Ge 1996-2006	12,309	0.007	0.112	-0.827	1.006
It 1951-1961	8,100	-0.047	0.161	-0.861	1.873
It 2001-2011	8,081	0.043	0.117	-0.580	3.303
Sp 1950-1960	7,901	-0.053	0.176	-1.360	1.579
Sp 2001-2010	8,074	0.038	0.244	-1.458	3.258

Table 3: ML estimators and standard errors (SE) for the dmen and the studied log-growth rate samples. The estimators for the normal distribution are the mean and standard deviation of the log-growth data, see Table 2

Sample	dmen		
	ρ (SE)	ϵ (SE)	ν (SE)
Fr 1990-1999	9.590 (0.214)	-0.085 (0.004)	0.484 (0.016)
Fr 1999-2009	4.131 (0.205)	0.047 (0.017)	0.041 (0.003)
Ge 1996-2006	10.012 (0.207)	0.031 (0.003)	0.338 (0.020)
It 1951-1961	12.722 (0.631)	-0.203 (0.011)	0.371 (0.113)
It 2001-2011	15.239 (0.389)	-0.019 (0.004)	0.565 (0.060)
Sp 1950-1960	5.718 (0.404)	-0.243 (0.009)	0.349 (0.029)
Sp 2001-2010	7.308 (0.556)	-0.292 (0.016)	0.496 (0.059)
	μ (SE)	σ (SE)	
Fr 1990-1999	0.055 (0.001)	0.0981 (0.0005)	
Fr 1999-2009	0.108 (0.001)	0.1215 (0.0006)	
Ge 1996-2006	-0.005 (0.001)	0.0873 (0.0007)	
It 1951-1961	-0.059 (0.002)	0.1420 (0.0014)	
It 2001-2011	0.029 (0.001)	0.1005 (0.0011)	
Sp 1950-1960	-0.076 (0.002)	0.1140 (0.0013)	
Sp 2001-2010	0.015 (0.003)	0.1627 (0.0017)	
	τ (SE)	ζ (SE)	θ (SE)
Fr 1990-1999	0.026 (0.001)	8.995 (0.078)	0.572 (0.015)
Fr 1999-2009	0.053 (0.002)	7.151 (0.068)	0.466 (0.013)
Ge 1996-2006	0.133 (0.004)	8.357 (0.339)	0.534 (0.022)
It 1951-1961	0.124 (0.014)	4.767 (0.306)	0.301 (0.025)
It 2001-2011	0.167 (0.008)	7.326 (0.435)	0.360 (0.028)
Sp 1950-1960	-0.048 (0.012)	5.454 (0.129)	0.507 (0.020)
Sp 2001-2010	-0.062 (0.005)	3.913 (0.068)	0.574 (0.024)

Table 4: Means and standard deviations (SD) according to the estimated dmen and the studied log-growth rate samples.

Sample	Mean	SD
Fr 1990-1999	0.046	0.126
Fr 1999-2009	0.099	0.147
Ge 1996-2006	0.007	0.112
It 1951-1961	-0.047	0.160
It 2001-2011	0.043	0.113
Sp 1950-1960	-0.053	0.174
Sp 2001-2010	0.038	0.241

Table 5: p-values (statistics) of the Kolmogorov–Smirnov (KS), Cramér–Von Mises (CM) and Anderson–Darling (AD) tests for the used samples and density functions. Non-rejections are marked in bold

Sample	normal		
	KS	CM	AD
Fr 1990-1999	0 (0.054)	0 (32.200)	0 (188.373)
Fr 1999-2009	0 (0.051)	0 (29.538)	0 (188.566)
Ge 1996-2006	0 (0.050)	0 (10.045)	0 (64.962)
It 1951-1961	0 (0.038)	0 (3.295)	0 (22.600)
It 2001-2011	0 (0.051)	0 (6.917)	0 (43.700)
Sp 1950-1960	0 (0.086)	0 (19.519)	0 (114.201)
Sp 2001-2010	0 (0.083)	0 (20.033)	0 (118.089)
	dmen		
	KS	CM	AD
Fr 1990-1999	**0.074 (0.008)**	**0.410 (0.144)**	**0.348 (1.018)**
Fr 1999-2009	**0.485 (0.005)**	**0.444 (0.133)**	**0.375 (0.967)**
Ge 1996-2006	**0.865 (0.006)**	**0.931 (0.040)**	**0.958 (0.271)**
It 1951-1961	**0.869 (0.007)**	**0.832 (0.057)**	**0.927 (0.315)**
It 2001-2011	**0.469 (0.010)**	**0.695 (0.079)**	**0.739 (0.508)**
Sp 1950-1960	**0.368 (0.011)**	**0.423 (0.139)**	**0.486 (0.793)**
Sp 2001-2010	**0.838 (0.007)**	**0.763 (0.068)**	**0.819 (0.429)**

Table 6: Maximum log-likelihoods, AIC and BIC for the used distributions and log-growth rates samples. The lowest values of AIC and BIC for each sample are marked in bold

Sample	normal		
	log-likelihood	AIC	BIC
Fr 1990-1999	23,609	-47,213	-47,196
Fr 1999-2009	17,477	-34,951	-34,934
Ge 1996-2006	9,432	-18,859	-18,845
It 1951-1961	3,322	-6,640	-6,626
It 2001-2011	5,855	-11,705	-11,691
Sp 1950-1960	2,518	-5,031	-5,017
Sp 2001-2010	-80.254	164.508	178.501
	dmen		
	log-likelihood	AIC	BIC
Fr 1990-1999	25,992	**-51,968**	**-51,900**
Fr 1999-2009	20,455	**-40,895**	**-40,827**
Ge 1996-2006	10,241	**-20,466**	**-20,407**
It 1951-1961	3,702	**-7,388**	**-7,332**
It 2001-2011	6,522	**-13,028**	**-12,972**
Sp 1950-1960	3,508	**-7,000**	**-6,944**
Sp 2001-2010	945.513	**-1,875**	**-1,819**

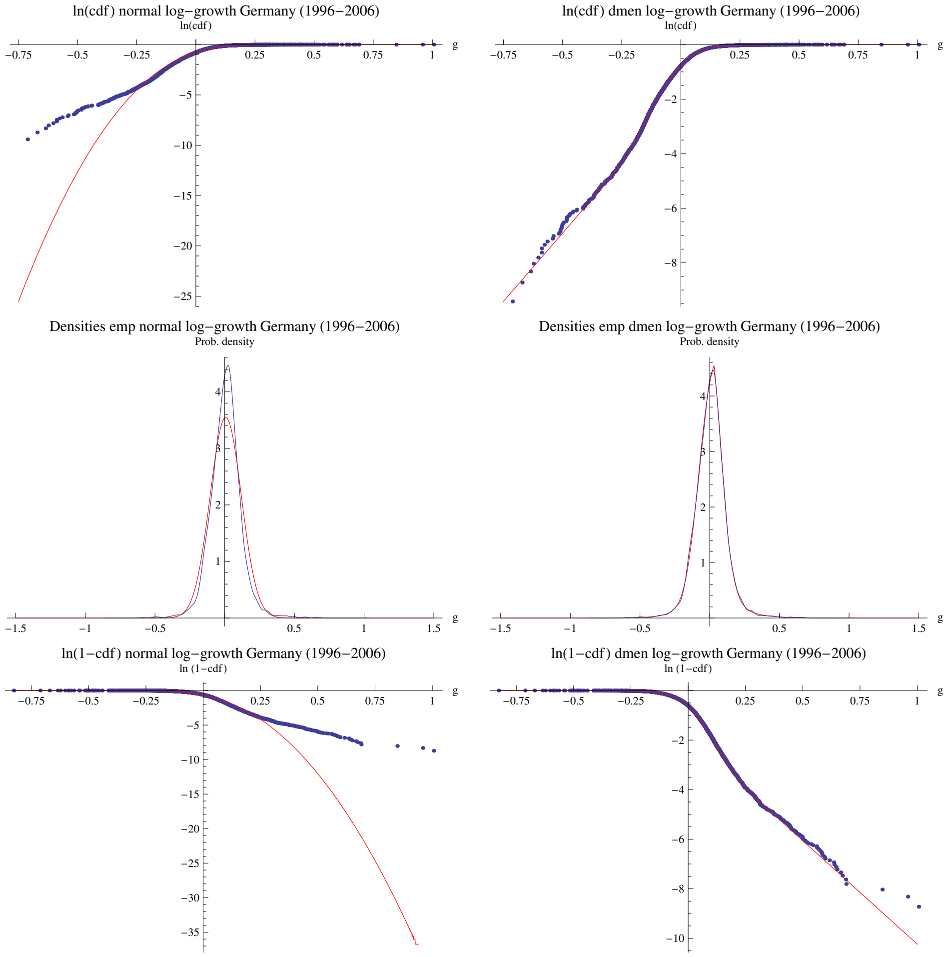

Figure 1: First row: empirical and estimated normal and dmen $\ln(\mathrm{cdf})$ for the lower tail. Second row: empirical (Gaussian adaptive kernel density) and estimated normal and dmen density functions. Third row: empirical and estimated normal dmen $\ln(1 - \mathrm{cdf})$ for the upper tail. Both columns: log-growth rates of German *Gemeinden* 1996-2006. Empirical in blue, estimated in red in all cases.