Gibrat's law for countries

González-Val, Rafael and Sanso-Navarro, Marcos

23 June 2008

Gibrat's Law for Countries

Rafael González-Val* Marcos Sanso-Navarro

Universidad de Zaragoza

June 2008

Abstract

If the population of a given country evolves according to Gibrat's Law, its growth rate will be independent of its initial size. This short paper further investigates this empirical regularity by the application of a suitable panel unit root test and non-parametric methods. The evidence regarding its fulfilment is weaker than that previously found.

Keywords: Gibrat's Law, Country Size, Non-stationary Panels, Non-parametric Methods.

JEL codes: C12, C14, C23, F00, R12.

1 Introduction

Gibrat's Law - also known as the Law of Proportional Growth - establishes that the growth rate of a variable is independent of its initial size. It has been the subject of a large number of empirical studies about its validity for both cities and firms (surveys about these can be found in Sutton, 1997; and Santarelli et al., 2006, respectively). Moreover, the observation of this empirical regularity for city size distributions has motivated theoretical developments in regional and urban economics (Gabaix, 1999; Duranton, 2006).

Rose (2006) has gone further and analyzed whether Gibrat's Law also fits another phenomenon associated with population size: the number of inhabitants of a country.

*Corresponding author. Address: Departamento de Análisis Económico. Facultad de Ciencias Económicas y Empresariales. C/ Gran Vía, 2. 50005, Zaragoza (Spain). Tel.: (+34) 976761000 Ext. 4728, Fax: (+34) 976761996. e-mail: rafaelg@unizar.es

Although theories that explain the size distribution of cities do not apply to the country size distribution, he concludes that these size distributions are similar. This finding is quite surprising since urban structure models rely on the assumption of the free mobility of workers. The latter is less usual for countries since international emigrants usually face transport costs as well as cultural and legal barriers. In addition, a strong implication of Gibrat's Law in country sizes is that per capita income growth differences would only be explained by differences in labour productivity.

Rose (2006) analyzed the fulfilment of this empirical regularity for country sizes using both visual (scatter plots and histograms) and econometric (β-convergence regressions and normality tests) tools. We try to contribute to this recently established trend by developing further alternative tests that challenge previous findings. This will be done by the use of up-to-date techniques that are applied in other contexts where Gibrat's Law is also relevant.

Clark and Stabler (1991) suggested that testing Gibrat's Law is equivalent to testing for the presence of a unit root. This idea has also been emphasized by Gabaix and Ioannides (2004) who expect *"that the next generation of city evolution empirics could draw from the sophisticated econometric literature on unit roots"*. Given the structure of the data analyzed, we apply the panel unit root test recently proposed in Pesaran (2007). Apart from controlling for the possible dependence among countries, it has nice size and power performance when dealing with a cross-sectional dimension greater than the temporal one, as is the case here.

We also implement non-parametric tests. On one hand, kernel regression estimates (Ioannides and Overman, 2003; Eeckhout, 2004) that establish a functional form-free relationship between population growth and country size for the entire distribution will be used. On the other, transition matrices (Quah, 1993) that determine the long-run tendency of the population distribution across countries will be calculated. They will allow us to obtain information about the degree of intra-distributional mobility.

2 Data and Methodology

2.1 Data

The size of a country can be measured using several magnitudes. However, the study of Gibrat's Law in country sizes was motivated by the finding in Rose (2006) that the distributions of city and country populations are similar. For this reason, we will focus our analysis on the latter variable. The data has been extracted from the Penn World

Table Version 6.2. (see Heston et al., 2006 for details). It is annual and covers 187 countries during the period 1950-2004. Therefore, our analysis of the world population distribution and its evolution will mainly refer to the second half of the 20th century.

2.2 Panel unit root testing

The modeling of city growth with autocorrelated errors and the assumption that Gibrat's Law holds in Clark and Stabler (1991) can be directly applied to country population growth. Within this framework, testing Gibrat's Law in country sizes is equivalent to testing for the presence of a unit root in the natural logarithm of population. That is, if we reject the null hypothesis that the (log) country population time series have a unit root, we are also rejecting the null hypothesis that its population evolves according to Gibrat's Law. We test for a unit root by exploiting the panel structure of the available country population data.

The first question we should be aware of is the possible presence of cross-sectional dependence. This is because it has been well established in the literature that panel unit root and stationarity tests that do not explicitly allow for this feature among individuals present size distortions (see Banerjee et al., 2005). We show that this is relevant in our case using the simple test of Pesaran (2004). It is based on the average of pair-wise correlation coefficients of the OLS residuals obtained from standard augmented Dickey-Fuller (1979, DF) regressions for each individual i (e_{it}). Let $\hat{\rho}_{ij}$ be the sample estimate of the pair-wise correlation coefficient for countries i and j calculated over T periods:

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^{T} e_{it}e_{jt}}{\left(\sum_{t=1}^{T} e_{it}\right)^{\frac{1}{2}} \left(\sum_{t=1}^{T} e_{jt}\right)^{\frac{1}{2}}} \tag{1}$$

Pesaran (2004)'s test does not depend on any particular spatial weight matrix when the cross-sectional dimension (N) is large. Its null hypothesis is cross-sectional independence and is asymptotically distributed as a two-tailed standard normal distribution. The test statistic is calculated as:

$$CD = \sqrt{\frac{2T}{N(N-1)} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij}\right)} \longrightarrow N(0,1) \tag{2}$$

Having shown that the members of our panel are cross-sectionally correlated, we will test for the presence of a unit root in the natural logarithm of country population (y_{it}) taking this into account. Following the proposal of Pesaran (2007), DF regressions will be further augmented with the cross-sectional mean and some of its lags in order to proxy for a single unobserved factor. The resulting individual DF test statistics can then be averaged in a similar fashion to Im et al. (2003)'s panel unit root test (CIPS test). Following Choi (2001), we will also combine the p-values of the individual tests (CZ test). Critical values are obtained with Monte Carlo simulations for a given specification of the deterministic component and depend on both the cross-sectional and temporal dimensions. In the case of serially correlated errors, it is suggested to include (up to p) additional lags of the augmentation terms.

2.3 Non-parametric methods

2.3.1 Kernel regression

In line with Ioannides and Overman (2003) and Eeckhout (2004), we specify the normalized logarithmic growth rate of a given country i (g_i) in a non-parametric way:

$$g_i = m(s_i) + \varepsilon_i \tag{3}$$

Thus, this variable is expressed as a function $m(\cdot)$ of the natural logarithm of its relative size (s_i). The latter is defined as its ratio over the contemporary world sample average. Consequently, instead of assuming a linear relationship between these two variables, as in the conventional β-convergence regressions framework, $m(\cdot)$ is estimated as a local average. This is done using a Kernel function $K(\cdot)$, assumed to be symmetric, weighted and continuous. ε_i is the perturbation term.

Population growth rates have been calculated yearly over the entire sample period. Since we are working with normalized rates, Gibrat's Law would be observed if the estimated mean is a straight line along zero and its variance is around one. Deviations from these values should lead us to reject its fulfilment in our sample.

The applied estimation method was originally developed by Nadayara and Watson. It is described in Härdle (1990) as:

$$\hat{m}(s) = \frac{n^{-1}\sum_{i=1}^{n} K_h(s - s_i)\, g_i}{n^{-1}\sum_{i=1}^{n} K_h(s - s_i)} \tag{4}$$

K_h reflects the dependence of the Kernel function on the bandwidth (h). We have fixed a value for this latter parameter equal to 0.5 and used the Epanechnikov kernel. n is the number of observations.

The population growth rate's variance is estimated as:

$$\hat{\sigma}^2(s) = \frac{n^{-1}\sum_{i=1}^{n} K_h(s - s_i)(g_i - \hat{m}(s))^2}{n^{-1}\sum_{i=1}^{n} K_h(s - s_i)} \tag{5}$$

2.3.2 Transition matrices

Gibrat's Law and transition matrices (Quah, 1993) are related in the following sense: persistence in size distribution would be observed when population growth is independent of size. This is equivalent to saying that transitions would be rarely observed.

In order to estimate a transition matrix, we should assume that the world population distribution evolves according to a homogeneous first-order stationary Markov process. The relative country size distribution at a given point in time is divided into discrete cells whose cut-off points are defined by specific values. Our grid is formed by five states, the upper bounds for each of them being 0.25, 0.5, 0.75, 2 and ∞. In 1950, they correspond to cell shares starting from the bottom of 57, 14, 11, 9 and 9% of the total number of countries, respectively.

Denoting F_t as the distribution of population across countries at time t, we assume that it could be described according to the following Law of Motion (see Quah, 1993):

$$F_{t+1} = M \cdot F_t \tag{6}$$

M maps one distribution onto another. That is, it tracks where a given point in F_t ends up in F_{t+1} in probability terms. Given the chosen number of states, M is the 5×5 Markov chain transition matrix whose (a, b) entry is the probability that a country in state a changes to state b.

Expression (6) is a useful first step when analyzing the dynamics in $\{F_t\}$. Its iteration yields (an estimation of) future distributions:

$$F_{t+k} = (M \cdot M \cdot \ldots \cdot M) \cdot F_t = M^k \cdot F_t \tag{7}$$

That is, one can characterize the ergodic distribution of cross-country population by calculating the limit of (7) as $k \to \infty$.

3 Results

3.1 Unit root testing

Table 1 presents both the cross-sectional dependence and panel unit root test results for the whole sample. Panel A reports those obtained when including only a constant term as the deterministic component in the augmented DF auxiliary regression. Note that up to four lags have been included in order to control for the possible presence of autocorrelation in the error term. The null hypothesis of no cross-sectional correlation is always rejected regardless of the number of lags included at the 1% significance level.

The unit root null hypothesis - equivalent to the fulfilment of Gibrat's Law - is rejected by both the CIPS and the CZ tests when a single lag is added at the 1% significance level. When a second lag is included, this rejection disappears for both tests. However, further rejections are encountered when including a third and a fourth lag.

Panel B displays the results when including both a constant and a trend as the deterministic components. It can be observed that previous results do not significantly depend on the specification of the deterministic terms since the conclusions remain almost unchanged with respect to those included in the upper panel of Table 1. The main difference is that the rejection of the null vanishes with the fourth lag. Therefore, it can be concluded that evidence against Gibrat's Law is found when using the panel unit root tests in Pesaran (2007).

3.2 Kernel regression

Kernel estimation results are plotted in Figures 1 and 2 for both the mean and variance of the normalized population growth, respectively. All the available information has been pooled, corresponding to 10,098 observations. Bootstrapped 95 % confidence bands calculated using 500 random samples with replacement are also displayed.

It can be observed in Figure 1 that population growth conditional on country relative size has an inverted U-shape around zero if the upper and lower tails of the distribution are not considered. The null hypothesis of this mean being equal to zero can be rejected at the 5 % significance level for some values near the lower tail. This may be related to the fact that small countries such as Djibouti, Grenade, Vanuatu or Brunei have been included in the analysis. For them, a small change in population size corresponds to a big number in percentage terms, the estimated growth rate being statistically different from the sample mean.

In figure 2, it can be observed that the population growth variance is greater for the smallest countries in the sample. Again, we can only reject that this measure is equal to one for some small relative country sizes. Consequently, we only find some evidence against Gibrat's Law in countries in the lower tail of the size distribution.

3.3 Transition matrices

The results of the transition matrices estimation are reported in Table 2. Panel A shows the one-step decennial transition matrix, obtained by averaging the observed transitions for each of the five decades in the period 1950-2000. When using this 10-year horizon, the main feature obtained is persistence. It can be observed that some values in the diagonal exceed 0.90. Specifically, most countries with less than 0.25 times the mean and 94 per cent of the biggest countries remained in the same state during the following decade. Intermediate states are less persistent. Nonetheless, all diagonal entries for them are greater than 0.75. Ergodic probabilities show some intra-distributional mobility since a tendency is observed for countries to move towards the upper-mean states of the distribution.

Panel B reports the 54-year transition matrix for the period 1950-2004. Although less pronounced, we still observe some persistence in the extremes of the distribution. In addition, there is much more intra-distributional mobility of the countries in the intermediate states. Ergodic probabilities also point to a high intra-distributional mobility as well as a tendency to move towards the upper-mean states. Evidence of a loss of weight in the lower tail is also found.

Summarizing, we obtain further results against the empirical fulfilment of Gibrat's Law for country population using transition matrices in terms of intra-distributional mobility of countries and concentration changes. Finally, note that more mobility is observed to that usually encountered for cities.

4 Conclusions

This paper has implemented further tests of Gibrat's Law in country sizes to those in Rose (2006). Although our results are mixed, we obtain much more evidence against this empirical regularity in the distribution of country population. This is specially true when using panel unit root tests and analyzing a 54-year transition matrix. Therefore, we should conclude that the theoretical modeling of country population in accordance with Gibrat's Law should be less of a concern unless stronger evidence for it is found.

In addition, our results make us to doubt that there is a common explanation for the size distribution of cities and countries.

References

[1] Banerjee, A., M. Massimiliano and C. Osbat, 2005. Testing for PPP: should we use panel methods? Empirical Economics, Vol. 30, pp. 77–91.

[2] Choi, I., 2001. Unit root tests for panel data. Journal of International Money and Banking, Vol. 20, pp. 249–272.

[3] Clark, J. S. and J. C. Stabler, 1991. Gibrat's law and the growth of Canadian cities. Urban Studies, Vol. 28, pp. 635–639.

[4] Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, Vol. 74, pp. 427–431.

[5] Duranton, G., 2006. Some Foundations for Zipf´s Law: Product Proliferation and Local Spillovers. Regional Science and Urban Economics, Vol. 36, pp. 542-563.

[6] Eeckhout, J., 2004. Gibrat's Law for (All) Cities. American Economic Review, Vol. 94 (5), pp. 1429-1451.

[7] Gabaix, X., 1999. Zipf's law for cities: An explanation. Quarterly Journal of Economics, Vol. 114 (3), pp. 739-767.

[8] Gabaix, X. and Y. M. Ioannides, 2004. The evolution of city size distributions. Handbook of Urban and Regional Economics, Vol. 4, J. V. Henderson and J. F. Thisse, eds. Amsterdam: Elsevier Science, North-Holland, pp. 2341-2378.

[9] Härdle, W., 1990. Applied nonparametric regression. Econometric Society Monographs. Cambridge, New York and Melbourne: Cambridge University Press.

[10] Heston, A., R. Summers and B. Aten, 2006. Penn World Table Version 6.2. Center for International Comparisons of Production, Income and Prices. University of Pennsylvania.

[11] Im, K., M. H. Pesaran and Y. Shin, 2003. Testing for unit roots in heterogeneous panels. Journal of Econometrics, Vol. 115, pp. 53–74.

[12] Ioannides, Y. M. and H. G. Overman, 2003. Zipf's law for cities: an empirical examination. Regional Science and Urban Economics, Vol. 33, pp. 127-137.

[13] Pesaran, M. H., 2004. General Diagnostic Tests for Cross Section Dependence in Panels. Cambridge Working Papers in Economics No. 435, University of Cambridge, and CESifo Working Paper Series No. 1229.

[14] Pesaran, M. H., 2007. A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence. Journal of Applied Econometrics, Vol. 22, pp. 265-312.

[15] Quah, D. T., 1993. Empirical cross-section dynamics in economic growth. European Economic Review, Vol. 37, pp. 426-434.

[16] Rose, A. K., 2006. Cities and countries. Journal of Money, Credit, and Banking, Vol. 38 (8), 2225-2246.

[17] Santarelli, E., L. Klomp and A. R. Thurik, 2006. Gibrat's Law: An Overview of the Empirical Literature. in E. Santarelli (Ed.), Entrepreneurship, Growth, and Innovation: the Dynamics of Firms and Industries, Springer, New York, pp. 41-73.

[18] Sutton, J., 1997. Gibrat's Legacy. Journal of Economic Literature. Vol. 35 (1), pp. 40-59.

Table 1. Cross-sectional dependence and panel unit root tests. Country population (in natural logs, N=187), 1950-2004.

Panel A. Constant, no trend				
	p = 1	**p = 2**	**p = 3**	**p = 4**
CD	6.96***	10.57***	6.39***	8.28***
CIPS	-2.85***	-1.81	-2.43***	-2.02*
CZ	-15.74***	-0.51	-9.53***	-3.50***
Panel B. Constant and trend				
	p = 1	**p = 2**	**p = 3**	**p = 4**
CD	10.94***	6.57***	7.25***	6.42***
CIPS	-3.22***	-1.97	-2.69***	-2.18
CZ	-13.77***	6.05	-5.40***	2.73

Note: All statistics are based on univariate AR(p) specifications. CD test developed in Pesaran (2004) for the null of cross-sectional independence distributed as a two-tailed standard normal. CIPS and CZ are the panel unit root tests in Pesaran (2007). Critical values are provided by the author. **, ** and * denote rejection of the null at the 1, 5 and 10 % significance level, respectively.

Table 2. Estimates of transition matrices and ergodic distributions. Country population (in natural logs), 1950-2004.

	Upper endpoint				
	∞	**2**	**0.75**	**0.50**	**0.25**
Initial	0.09	0.09	0.11	0.14	0.57
Panel A. Average 10-year transition matrix					
	∞	**2**	**0.75**	**0.50**	**0.25**
∞	0.94	0.06			
2	0.06	0.92	0.02		
0.75		0.11	0.76	0.12	
0.50			0.09	0.85	0.06
0.25				0.02	0.98
Ergodic	0.32	0.33	0.06	0.09	0.20
Panel B. 54-year transition matrix					
	∞	**2**	**0.75**	**0.50**	**0.25**
∞	0.71	0.29			
2	0.38	0.56	0.06		
0.75		0.33	0.19	0.38	0.10
0.50		0.12	0.35	0.31	0.23
0.25			0.01	0.10	0.89
Ergodic	0.43	0.34	0.05	0.05	0.13

Note: Constructed by the authors with data from the Penn World Table Version 6.2. The ratio of a country population with respect to the world sample average is used.

Figure 1. Mean population growth (standardized), 1950-2004. Kernel regression estimation and 95% confidence bands. Bandwidth=0.5.

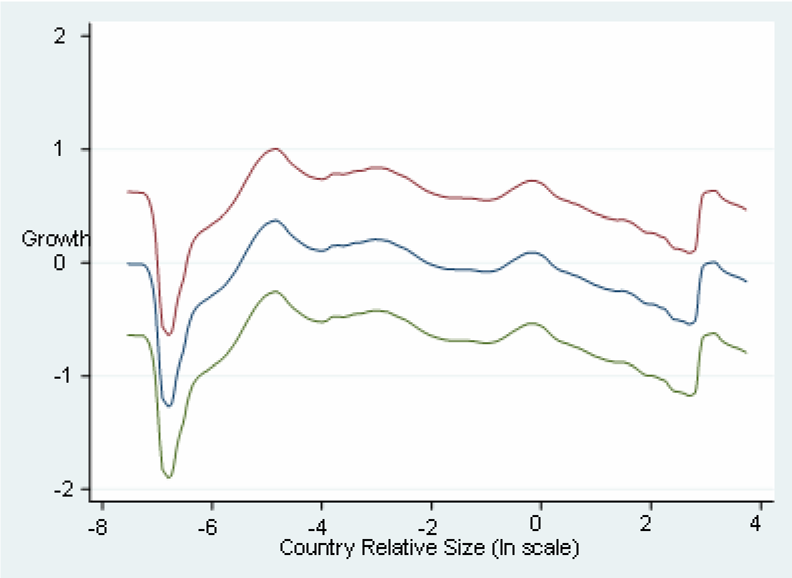

Figure 2. Variance of population growth (standardized), 1950-2004. Kernel regression estimation and 95% confidence bands. Bandwidth=0.5.