Market access and local growth: what do 150 years of data about Italy say?

This draft: May 2014
Preliminary and incomplete

Marcello Pagnini[1]

Using a set of parametric and non-parametric techniques in this paper I collect four pieces of evidence about spatial concentration in Italy. First, a large share of population size distribution across municipalities is well described by a Paretian distribution, well beyond the segment of the upper tail of the distribution. Second, the concentration of population in Italy is lower than that would be predicted by a Zipf law. Third, in the period between 1861 and 2011, spatial concentration of population increased till 1971, after that year concentration started to decline. Fourth, this evolution of population size distribution is common to the Northern and Southern regions of the country. In the second part of the paper, an econometric analysis is carried out to investigate the role of market access in fostering the increase in spatial concentration. Consistently, with the tenets of the economic geography literature, I find a positive effect of market access on local population growth for the whole 1861-2011 period, even considering a wide set of controls. It is still to be assessed whether the new economic geography is able to explain the rising and then falling trends in spatial concentration in Italy.

JEL CODES: R11, R12, R23
Keywords: spatial concentration, Pareto Distribution, Market potential, local growth

[1] Bank of Italy, Regional Economic Research Division, Bologna branch. The opinions expressed in this paper are of the author and do not involve the Bank of Italy.

Introduction

One of the major prediction in the economic geography literature concerns the bell shaped relation between agglomeration and trade costs (see Puga, 1999 among others). When those costs are prohibitively high economic activities are dispersed. When trade costs start to follow centripetal forces set in determining an increase in agglomeration. Provided the fall in the trade costs continues, it is possible that dispersion forces regained momentum and that the new spatial equilibrium will be characterized by more dispersion.

In this paper we want to investigate this claim of the economic geography by looking at the long term evolution of population size distribution in Italy in the last 150 years. The availability of long term data is essential to examine this topic as one can observe a country on the pre-industrialization stage when trade costs were very high and then observing subsequently the effects on agglomeration of the dramatic fall in trade costs brought about by the evolution of technology. Finally, it is also possible to investigate the effects of the continuing fall in trade costs during the post industrialization phase.

Section 1: Methodological issues

Suppose there are N locations or cities ranked according to their population S_i $(i=1, .., N)$ in a way that $S_1 > \ldots > S_N$. Now we say that this set of locations would follow a power law or Pareto law whenever the probability that S> x is described by the following countercumulative distribution function: $P(S > x) = k / x^{\zeta}$ where k is a positive constant. Consequently the density function will be: $f(x) = k\zeta / x^{1+\zeta}$ .The exponent or shape parameter ξ measures the concentration of location size distribution and will vary inversely with it. If the Pareto describes the entire set of locations then we have $k=S_N$, alternatively one can assume that this law would describe only the locations in the upper tail of the distribution beyond a given threshold $S_* > S_N$ and with $k = S_*$. When the shape parameter $\zeta = 1$ then we have a Zipf law implying that the probability that a location has a population greater than *x* is proportional to *1/x*.

In what follows we are interested in establishing whether the Pareto law fits well the Italian city size distribution across time; the population threshold above which this law holds; the value of the shape parameter and its variations across time; whether the Zipf law holds for the Italian economy.

There are different ways to estimate a Power law distribution here we follow a method based on an approximation known as the log-rank-log size rule. Let *r* denote the rank of cities in terms of their size, then a power law can be estimated by running the following linear equation:

$$\ln(r - .5) = \alpha - \zeta \ln x_r + q(\ln x_r - \gamma)^2 \quad (1)$$

where $\gamma \equiv \mathrm{cov}((\ln(x_j)^2, \ln(x_j)) / (2\,\mathrm{var}(\ln(x_j)))$. This specification is manly due to Gabaix and Ibragimov (2011), henceforth GI, see also Gabaix (2009), and was implemented in Rozenfled et al (2011) and Bee et al (2012) among others.

Several clarifications are in order. First, the log rank–log size rule can be derived from a power law specification, but it does not deliver the exact value of the shape parameter even when data are actually generated through a Power law. Hence, with a specific reference to the Zipf law this implies that an estimated coefficient in the interval [*.8,1.2*] can be considered a good approximation to that law (see Ioannides and Gabaix, 1994 for a discussion). The term .5 was introduced to correct a bias related to small samples. Moreover, it can be proved that the standard error of the parameter delivered from the OLS estimation of equation (1) is biased due to the positive correlation of residuals induced by the ranking. A partial solution to this problem consists in using an asymptotic standard error, $\zeta^{OLS}(N/2)^{.5}$ where N represents the number of observations.

The quadratic term in (1) captures deviations from a power law distribution. When the estimated q parameter is approximately 0 then data will asymptotically follow a power law distribution, alternatively when $|q|$ is large the Pareto does not fit well the data. GI showed that the statistic $(2N)^{.5} q_N / \zeta^2$ converges to a standard normal for large samples. Hence we can test whether the Pareto fits the data well based on that test. An alternative procedure in order to assess the goodness of fit of the Pareto distribution would involve a Kolgomorov-Smirnov test. The problem with this method is that the null of a Pareto distribution is tested against every other alternative and that it lacks power. To address this problem we use the Uniform Most Powerful Unbiased (UMPU) test recently proposed Castillo and Puig (1999) and implemented by Malevergne et al (2011). This test is based on a null of the Pareto distribution against a lognormal one that is usually considered as a valid alternative to describe city size distribution (see Eeckout, 2004).

Finally, a hotly debated issue in the literature concerns the way to estimate the lower bound, S_* above which the Pareto distribution should be a valid description of the data. This topic is pivotal both for the goodness of fit of the Pareto distribution as well as for the estimation of the parameter ζ.[2] Here we follow a sort of recursive procedure suggested by Clauset (2009).[3] The idea is that one can start from a sample of the largest locations and then add one observation at each stage of the procedure up to the point where the full sample is obtained (ie $S_* = S_N$). Then at each stage, it is possible to estimate the shape value based on the log rank-log size regression and the corresponding GI test for the goodness of fit of the Pareto distribution. Our estimation of the truncation

---

[2] See the contributions by Eeckout (2004), Clauset (2008), Malevergne et al (2011), Combes et al (2012) and Ioannides and Skouras (2013).

[3] See also Fazio and Modica (2012), I have actually mutuated the term 'recursive' from the title of their contribution..

point of the distribution is the value of S_* above which the GI test is passed. A similar procedure will be followed using the UMPU test.

Section 2: the Data

Our data come from the census of the Italian population conducted by the Italian National Institute of Statistics (ISTAT) at each decade starting from 1861 onwards. The census did not take place in 1891 due to financial problems and in 1941 due to the war while in 1936 an additional one was carried out. The latest census refers to 2011. All in all, our data include 15 census covering the period 1861-2011. Population data are available at the level of municipality. In 1991 there were 8,100 municipalities covering all the country's land.

Several changes occurred in the data also as a consequence of political events. First, Italy acquired some portion of land, due to the wars in 1866 and 1870, consequently the number of municipalities rose between 1861 and 1871, including the municipality of Rome that became the Italy's capital. Likewise a group of municipalities located in the North East was added to the data set in 1921 census following the Italy's territorial gains related to the outcome of the first war world. A second source of variation concerns both the change in the municipality borders and consequently of the population living within each location and the births or deaths of municipalities that were not related to the war. In Istat (1994) data were harmonized to take account of these variations in a way to obtain relatively comparable information across census in the interval between 1861 and 1991. The baseline for this reconstruction was the 8,100 municipalities as defined in 1991. Whenever possible we extend this harmonization to the 2001 and 2011 census. We end up with a final sample of municipalities for the different census years described in Table 1.

Section 3: the evolution of municipality size distribution

In this section we implement the recursive method described above and discuss results with reference to 1991 in the first place and to the entire sample period subsequently.

The value of the estimated shape parameter is reported as a function of the population threshold or truncation point in Figure 1 (the minimum population threshold is 33 corresponding to the full sample of 8,085 municipalities while the maximum is equal to 96,614 with a sample size of the 50 largest municipalities). Moreover, we run the goodness of fit tests based on GI and UMPU. According to the former we cannot reject the null that the empirical distribution is well described by the Paretian distribution starting from a population size of 3878. For that particular truncation point we get a $|q=.050213| < qc=.0520753$ and hence we do not reject the Paretian at the one per cent confidence level. For population thresholds above 3878 the test always returns the same outcome in favour of the Pareto distribution. As for the UMPU test we check when its p value is larger than one per cent. This is always the case for a population threshold above 9559 inhabitants, starting from that value onwards the Paretian has to be always preferred with respect to the lognormal distribution.

Municipalities showing a population above the two thresholds for the Pareto and UMPU test cover, respectively, the 85 and 67 per cent of total Italian population in 1991. Hence the goodness of fit of the Pareto law is not restricted to the upper tail of the Italian population size distribution but extends to a large part of its body. The Figure 1 also shows that the shape parameter clearly varies with the cut-off and that in particular its value decreases as we restrict attention to the largest cities. Finally confidence intervals fan out as a consequence of reducing sample size and increasing the cut-off. Although these results are usually considered as a problem for the Pareto law, we find that the range variation of the shape parameter as well as of the confidence intervals are relatively narrow compared to those of alternative distributions. Hence we confirm the use of the Pareto assumption to describe the size distribution of the Italian municipalities and its evolution across time.

Table 2 and Figures 2-3 report for the different census waves the estimated shape parameter, some statistics for the regression based on equation (1), the population cutoffs based on IG and UMPU tests and the confidence intervals.

Two major findings quite neatly stand out. First, the degree of concentration of population in Italy is lower than that it could be expected on the basis of the Zipf law, ie when $\zeta \approx 1$. Our findings reject this law even when considering the wider interval of parameter values [*.8*,*1.2*] that IG indicate as potentially consistent with the validity of the Zipf. Actually, those values were outside the limits set by the confidence interval for most of the census years. Second, consistently with the fall in trade costs following the unification of the Country, the degree of concentration had been increasing till the 1971. After that census year, the tendency toward concentration was reverted and replaced by a moderate growth in dispersion.

To check whether our results could be driven by the time pattern of the cut-offs (see Table 2) or in any case by presence of the small sized locations, we drop from the sample the municipalities with a population below 5000 inhabitants (Fig. 4) . Results are very similar to previous ones. Moreover, we radically changed our concentration measure by resorting to a Theil Index (see Combes et al , 2011). The time pattern of concentration detected through that index was basically the same as that illustrated above.

The Italian economy is well known to be characterized by huge differences across its territories. To investigate how the latter could have reflected onto the evolution of the population size distribution, we replicate our analysis by splitting the country into four areas (North West, North East, Centre and the South, including Sicily and Sardinia). Quite surprisingly, Figures 5, 6 and 7 basically reproduced the same time patterns in the population size distribution as those detected for the Italian economy as a whole, ie a propensity of population to agglomerate up to the 1971 and a subsequent fall in concentration thereafter.

Finally, some authors sometimes criticized these measures of concentration as being excessively dependent on parameterization. In view of this criticism and to provide an additional robustness check for our findings, we resort to a fully non parametric technique that was recently proposed by Combes et al (2012) and that it is completely new in the context of the literature on city size distribution. The basic idea is that of comparing two distributions, in our case the populations of a group of municipalities in two subsequent census years without making any assumption about the shape of the distribution and potentially using all the points in the distributions.

Three parameters describe the change in the distribution between the two census years. The first one, let us all it T, describes the shift (rightward for T>0 or leftward, T<0) for the entire distribution (for instance, population in all municipalities would grow at same rate when T>0). The second one, denoted by S, a shift in the truncation point in the lower tail of the distribution (S>0). Finally, the third one captures the dilation in the distribution (in that case for D>1 large cities would grow at a higher rate than small sized municipalities and the opposite would be true for D<1). Thanks to bootstrap techniques, it is also possible to test for T and S to be significantly different from zero and for D from one. Moreover, we can also obtain a goodness of fit test based on a sort of Pseudo R2. Results are reported in Table 2. All in all these findings show that the changes in the municipality size distribution were dominated by translation and dilation. As for the latter, the D parameter is significant for all the census years and its level evolve in a similar way as to that followed by previous indicators. Actually, larger cities were increasingly favoured in terms of population dynamics up to 1971, after that date the evolution of the distribution was still indicating an increasing concentration but at a much lower rate than in the previous decade.

Section 4. Explaining the evolution of municipality size distribution: the role of market access

As explained in the introduction, the economic geography literature claim that the relation between trade costs and agglomeration can be described by a bell shaped curve. When trade costs are prohibitively high, dispersion force prevail leading to a relatively uniform size distribution of the locations. As trade costs start to decline, agglomeration forces set in and lead to a new equilibrium characterized by an increase in spatial inequality. Provided trade costs continue to fall, a point can be reached where dispersion forces regained momentum determining an equilibrium with a less unequal spatial distribution of economic activities.

In what follows, we examine whether this bell shaped relation can be used to explain the evolution of municipally size distribution in Italy in the last 150 years that was described in the Section 3. More specifically we will investigate the relationship between population dynamics and local market access as measured by the Harris Market potential. To go more deeply into this matter we resort to a simple model by Redding and Sturm (2008).

Consider an economy with N locations, denoted by i $(i=1,\dots,N)$. A given mass of L workers are endowed with a one unit of labour that they offer inelastically. Workers are free to move across locations. Each location is endowed with a given stock of local amenities, H_i, that is inelastically supplied. A share, $1-\mu$ of workers' expenditures is spent on local non tradable amenities, while the share μ is used to buy tradable horizontally differentiated varieties of the same good. To produce each variety requires a fixed amount of labour and a constant marginal cost. Tradable varieties are produced under monopolistic competition. Finally there will be a iceberg trade costs to sheep one variety from location *i* to *j* ($T_{ij}>1$). Real wages are equalized across locations thanks to migration flows that also determine endogenously the local population L_i in equilibrium. Given this set-up, it is possible to show that:

$$L_i^* = x(FMA_i)^{\frac{\mu}{(1-\mu)\sigma}}(CMA_i)^{\frac{\mu}{(1-\mu)(\sigma-1)}}H_i \quad (2)$$

where $FMA_i = \sum_j (w_j L_j)(P_j^M)^{\sigma-1}(T_{ij})^{1-\sigma}$ and $CMA_i = \sum_j n_j(p_j T_{ij})^{1-\sigma}$ are respectively firms' and consumers' market access. Assume that L_i^* represent the long term equilibrium population level and that the economy takes time to reach it as described in the following adjustment process: $L_{it+1} = L_{it}^{*\lambda} L_{it}^{1-\lambda}$ $(0 < \lambda < 1)$. Taking logs and substituting for L_i^* from (*) we get :

$$\Delta_{t+1,t} \log L_i = \lambda \log(x) + \lambda \log\left(x(FMA_{it})^{\frac{\mu}{(1-\mu)\sigma}}(CMA_{it})^{\frac{\mu}{(1-\mu)(\sigma-1)}} \right) + \lambda \log(H_{it}) - \lambda \log L_{it}$$

To get closer to the specification that will be used in the empirical analysis we assume that $\log\left((FMA_{it})^{\frac{\mu}{(1-\mu)\sigma}}(CMA_{it})^{\frac{\mu}{(1-\mu)(\sigma-1)}} \right) = \alpha_1 \log(L_{it}) + \alpha_2 \log(HMA_{it}) + \varepsilon_{it}$

where $HMA_{it} = \sum_{j\neq i}(L_{jt}/d_{ij})$ is the Harris formula for the market access in which the population of the other N-1 locations is discounted by the inverse of the physical distance between *i* and *j*, α(>0) is a parameter and ε_{it} is an error term.[4] Finally we also assume that

[4] We don't use any internal distance measure to discount the size of location i. This choice is consistent with the theoretical model ($T_{ii}=1$) and is also reasonable in view of the small size of the municipalities in our sample.

$\log(H_{it}) = A_{it}\delta' + \eta_{it}$ where A_{it} is a vector of variables related to the availability of local amenities to be specified later, δ' is a vector of parameters and η_{it} is an error term.

We end up with the following specification:

$$\Delta_{t+1,t} \log L_i = \beta_0 + \beta_1 \log L_{it} + \beta_2 \log(HMA_{it}) + A_{it}\gamma' + \nu_{it} \quad (3)$$

where $\beta_0 = \lambda \log(x)$, $\beta_1 = \lambda(\alpha_1 - 1)$, $\beta_2 = \lambda\alpha_2, \gamma' = \lambda\delta'$ and finally $\nu_{it} = \lambda\mu_{it} + \eta_{it}$.

Section 6.The econometric analysis

Equation (3) is estimated by pooling all the waves of population census data at municipally level carried out between 1861 and 2011. The dependent variable is the annualized percent rate of the population between two consecutive census years. We use the annualized rate because the intervals between two census may have a different length in our data (see Section 3).

Each explanatory variable is considered at the value it takes on at the beginning of each period. The formula for HMA_{it} has been already explained in the previous Section. Physical distance between each pair of municipalities, d_{ij}, is computed from their geographic coordinates and by using the great circle formula.

As controls for the stock of local amenities, we use data on the rain fall during the winter and summer season and average temperature in January and July estimated for each municipality. Following Rappaport (2007), these variables are entered with a linear and a quadratic term. For the latter, the linear sample mean of each variable is subtracted before squaring.

Considering the detailed spatial scale of the data, it is also important to control for urban sprawl across municipalities. To this aim, we consider the top largest twenty locations in each census wave and compute the minimum distance of each municipality from one of those top 20 cities. We then build a set of dummy variables that equal 1 when the municipality is within a given distance band from the group of the top 20 cities. We use different distance intervals for the group of small municipalities, ie those that are not in the top 20 group, and for the largest cities (for instance no city in the group of the top 20 is located at a distance shorter than 50 kms with respect to another large city). For instance the dummy band10 indicates distance between 0 and 10 kilometres, band20 between 10 and 20 and so on. For the large cities bc_band50 indicated sìdistances between 0 and 50 Kilometres and so on These variables are especially important to control for the expansion of the cities into the hinterland in which workers could decide to locate in order to avoid the high housing and congestion costs in the city and from which they could commute on a daily basis to the downtown city.

Finally, although it was not included in the model, we also introduce a variable reporting the electoral turnout in the political elections in Italy again measured at the level of municipality (see Albanese e De Blasio, 2013). Through that it is possible to control for the influence of social and human capital on local growth.

We use three estimation methods OLS, municipality fixed effects (FE) and IV gmm based on difference equations only (see Table 4). When using the OLS the specification is augmented with a set of time invariant variables that include: municipality longitude and latitude and its interaction, a dummy for municipalities on a island and another for those located along the coastline and 20 region fixed effects. All the specifications include period fixed effects, standard errors of parameters are always clusterized at provincial level for all specifications. For the sake of brevity, Table 4 only reports the results for the main variable of interest. Results clearly indicate that market access has a positive effect on population growth at local level. This result is robust across the three estimation methods. For the IV GMM however our estimates do not pass the Hansen test and also reject the absence of residual autocorrelation of the second order. This could mean that we have too many weak instruments.

All in all, this evidence points to the fact that locations with a larger market potential or a better market access will exhibit a higher population growth. This finding is consistent with results obtained by several contributions in this stream of literature.[5] Moreover, it is robust to the introduction of controls picking up the role of local amenities that are especially relevant for examining population dynamics, human and social capital as well as urban sprawl. The latter result is also extremely important as our results could have be driven by the fact that rising congestion costs in urban areas could have positively affected population growth in closer small locations.

Section 5. Spatial agglomeration patterns and the changing balance between agglomeration and dispersion forces

In this section we will deal with the following question. Provided that market access fostered agglomeration for the whole period, might the structural breaks in the relation between market potential and local growth explain the rising and subsequently falling patterns in spatial agglomeration?

For instance consider local amenities, in the early stage of development their role in driving people location choices can be modest as with law per capita income local amenities would be superior goods. As incomes rise, part of the population could afford a location choice based on amenity considerations. Hence we could expect that this would augment the impact of local amenities on local growth and maybe dwarf that related to market access.

To answer this question…

[to be written]

[5] See Duranton and Puga (2013) for a recent and very clear survey of these contributions.

Section 6. Final remarks

Using a set of parametric and non-parametric techniques in this paper I detected four basic evidences about spatial concentration in Italy. First, a large share of population size distribution across municipalities is well described by a Paretian distribution, well beyond the segment of the upper tail of the distribution. Second, the concentration of population in Italy is lower than that would be predicted by a Zipf law. Third, in the period between 1861 and 2011, spatial concentration of population increased till 1971, after that year concentration started to decline. Fourth, this evolution of population size distribution is common to the Northern and Southern regions of the country. In the second part of the paper, an econometric analysis is carried out to investigate the role of market access in fostering the increase in spatial concentration. Consistently, with the tenets of the economic geography literature, I find a positive effect of market access on local population growth for the whole 1861-2011 period, even considering a wide set of controls. It is still to be assessed whether the new economic geography is able to explain the long term evolution of spatial concentration in Italy.

Table1 Descriptive statistics on population in Italy: 19861-2011

year	N. of municipalities	Total population in 000s	Mean pop.	75 perc. pop.	Median pop.	25 perc. Pop.	Mean pop. density	75 perc. Pop. Dens.	Median pop. Dens.	25 perc. Pop dens.
1861	6679	22176	3320	3320	1864	1085	131	156	96	55
1871	7703	27300	3544	3526	2032	1192	135	160	103	59
1881	7703	28952	3758	3711	2119	1236	143	168	107	61
1901	7703	32963	4279	4107	2313	1329	158	184	116	66
1911	7703	35842	4653	4380	2441	1382	168	196	122	68
1921	8085	39397	4873	4487	2424	1329	172	199	122	67
1931	8085	41043	5076	4673	2425	1309	178	200	119	65
1936	8085	42398	5244	4732	2440	1305	181	201	119	65
1951	8085	47516	5877	5075	2587	1344	200	213	122	67
1961	8085	50624	6261	4881	2438	1246	211	208	111	61
1971	8085	54137	6696	4849	2260	1103	234	212	101	52
1981	8085	56557	6995	5207	2295	1061	256	226	101	49
1991	8085	56778	7023	5398	2317	1039	268	238	101	48
2001	8085	56973	7047	5662	2350	1024	278	251	104	47
2011	8085	59403	7347	6099	2436	1026	297	275	108	46

Table 2 Paretian parameter estimated by using eq. 1 in the text

TEST IG

year	Population cutoff	n. Municipalities	Shape parameter	Parameter standard error	R^2
1861	1697	3642	1.514	0.021	0.993
1871	1646	4680	1.509	0.019	0.993
1881	1687	4758	1.490	0.018	0.993
1901	1748	4925	1.440	0.017	0.992
1911	1812	4964	1.413	0.017	0.992
1921	1819	5080	1.371	0.016	0.991
1931	1913	4950	1.352	0.016	0.991
1936	1881	4978	1.335	0.016	0.990
1951	2023	4892	1.297	0.015	0.991
1961	1937	4808	1.225	0.015	0.993
1971	2241	4069	1.174	0.016	0.995
1981	3051	3299	1.192	0.018	0.995
1991	3878	2779	1.229	0.021	0.994
2001	4494	2492	1.263	0.022	0.993
2011	5075	2355	1.293	0.023	0.992

TEST UMPU

year	Population cutoff	n. Municipalities	Shape parameter	Parameter standard error	R^2
1861	2804	2128	1.616	0.033	0.998
1871	2830	2626	1.616	0.030	0.998
1881	2768	2911	1.591	0.028	0.999
1901	2937	3031	1.540	0.027	0.999
1911	3201	2914	1.517	0.027	0.998
1921	3370	2883	1.476	0.026	0.998
1931	3545	2804	1.465	0.027	0.998
1936	3608	2784	1.449	0.026	0.998
1951	3703	2935	1.396	0.025	0.998
1961	3536	2934	1.307	0.023	0.999
1971	3853	2606	1.231	0.023	0.999
1981	5629	1863	1.258	0.028	0.998
1991	8922	1163	1.323	0.036	0.996
2001	9559	1142	1.363	0.038	0.996
2011	10080	1183	1.395	0.038	0.996

Table 3 Test on distribution change in each period

period	T	S	D	sd_T	sd_S	sd_D	t_T	t_S	t_D	pseudo R2
1861-1871	0.075	0.000	0.986	0.004	0.001	0.005	20.006	-0.388	-2.627	0.941
1871-1881	0.045	0.000	1.006	0.001	0.000	0.001	48.062	0.815	4.562	0.971
1881-1901	0.090	0.000	1.023	0.002	0.000	0.002	41.297	-0.555	10.194	0.990
1901-1911	0.053	0.000	1.020	0.001	0.000	0.002	43.877	1.071	11.865	0.979
1911-1921	-0.007	0.000	1.034	0.002	0.000	0.003	-2.787	1.066	12.518	0.907
1921-1931	0.003	0.000	1.029	0.002	0.000	0.003	1.710	0.094	11.249	0.691
1931-1936	0.002	0.000	1.020	0.001	0.000	0.002	2.164	0.825	12.420	0.752
1936-1951	0.054	0.000	1.039	0.002	0.000	0.002	34.029	-0.779	19.699	0.694
1951-1961	-0.049	0.001	1.051	0.002	0.001	0.002	-20.019	0.814	22.284	0.963
1961-1971	-0.060	0.000	1.078	0.003	0.000	0.002	-23.779	0.591	31.740	0.965
1971-1981	0.006	0.000	1.050	0.003	0.001	0.003	1.998	-0.412	18.857	0.837
1981-1991	0.005	0.000	1.028	0.002	0.001	0.002	2.031	-0.452	14.547	0.696
1991-2001	0.008	0.000	1.018	0.003	0.001	0.002	3.271	-0.364	7.571	0.613
2001-2011	0.029	0.000	1.026	0.002	0.000	0.002	13.390	-0.456	12.098	0.814
1921-1961	0.010	0.001	1.147	0.004	0.001	0.004	2.674	0.979	33.401	0.984
1971-2011	0.082	-0.011	1.097	0.097	0.055	0.049	0.839	-0.209	1.969	0.818

Tables 4. Dependent variable: annualized percentage change in population: estimation period 1871-2011 (1)

	OLS(2)	FE	IVGMM (3)
Log (HMA_{t-1})	0.008***	0.028***	0.033**
	(0.002)	(0.006)	(0.011)
Log(L_{it-1})	0.002***	0.000	-0.005**
	(0.000)	(0.001)	(0.002)
Electoral Turnout	0.007***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)
band10==1	0.007***	0.005**	0.005*
	(0.001)	(0.002)	(0.002)
band20==1	0.004***	0.002*	0.000
	(0.001)	(0.001)	(0.001)
band30==1	0.002**	0.002	0.001
	(0.001)	(0.001)	(0.001)
band40==1	0.001**	0.001	-0.000
	(0.000)	(0.001)	(0.001)
band50==1	0.001*	0.001	0.001
	(0.000)	(0.001)	(0.001)
bc_band50==1	-0.004*	-0.009***	-0.021***
	(0.001)	(0.003)	(0.004)
bc_band100==1	0.001	0.003	0.011*
	(0.001)	(0.002)	(0.005)
bc_band150==1	-0.002	0.001	-0.007
	(0.002)	(0.003)	(0.005)
_cons	-0.378***	-0.361***	-0.368**
	(0.075)	(0.067)	(0.121)
Municipality FE	N	Y	
Year FE	Y	Y	Y
N	110216	110216	102091
R^2 adjusted	0.215	0.128	

(1) In all specifications the set of controls includes the amount of rain fall during the winter and summer season and the January and July temperatures in each municipality. (2) Specification includes a set of time invariant controls, ie a set of dummies for municipalities located on an island, the portion of land covered by mountains, one for municipalities located along the coastline, the (log of) municipality latitude and longitude and their interaction.(3)Endogenous variables are the log of population, market potential am the electoral turnout. Instruments include lags of all the time variant explanatory variables. Estimations are in difference only.

Fig.1: Shape parameters computed at different population cutoffs in 1991. Dashed curves represent IG confidence intervals, vertical line refer to population cutoffs defined by the IG and UMPU tests.

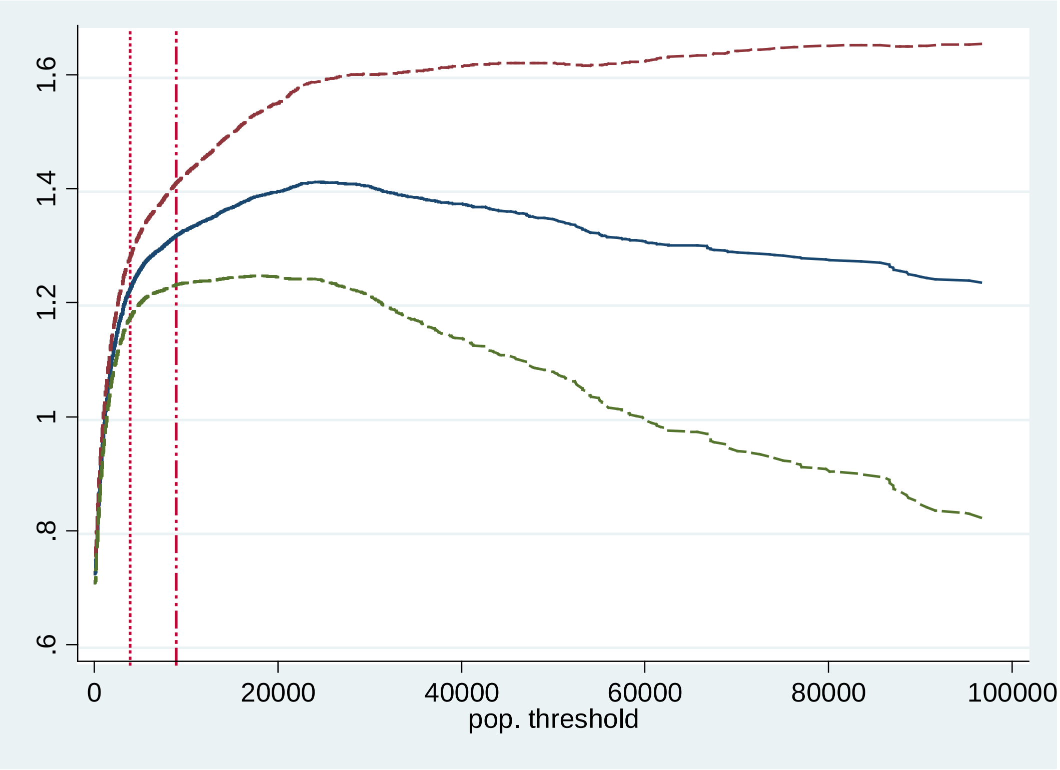

Fig.2 Shape parameters computed for all the census years (1861-2011) and population cutoffs defined by IG test. Dashed lines represent IG confidence intervals

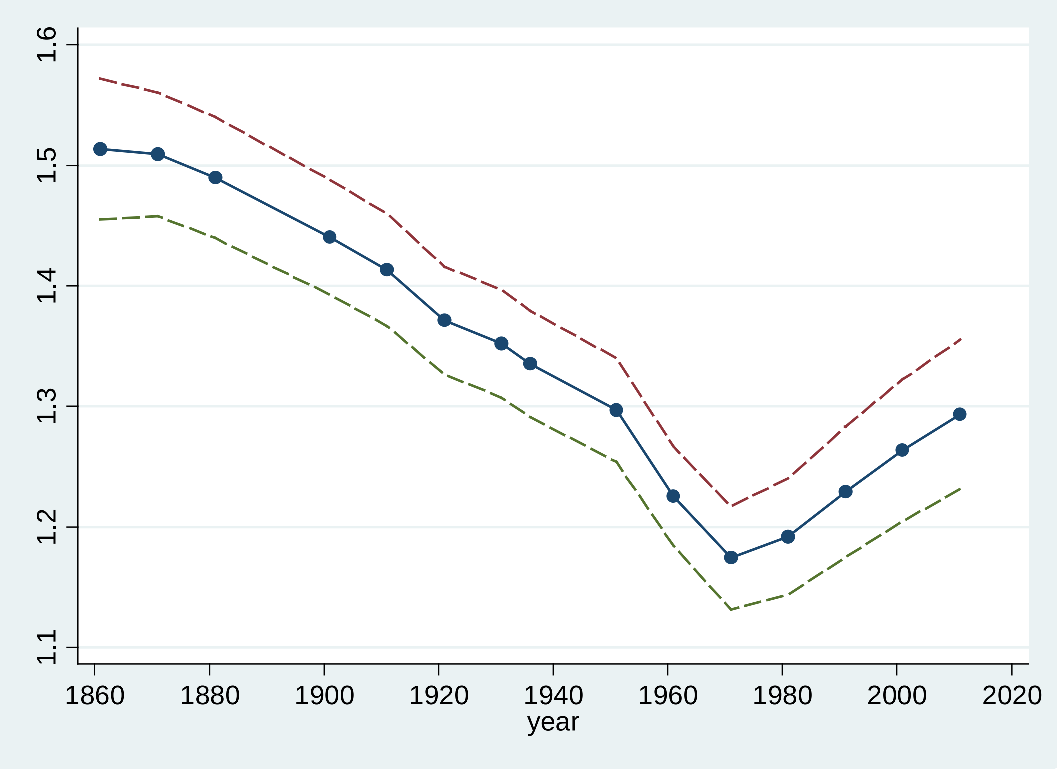

Fig.3 Shape parameters computed for all the census years (1861-2011) and population cutoffs defined by UMPU test. Dashed lines represent IG confidence intervals.

Fig.4 Shape parameters computed for all the census years (1861-2011) and population cutoff defined at 5000 inhabitants. Dashed lines represent IG confidence intervals.

Fig. 5 Shape parameters by macroaregions, computed for all the census years (1861-2011) and population cutoffs defined by IG test. Dashed lines represent IG confidence intervals (moving anticlockwise macroregions are NW, NE, CE; SO)

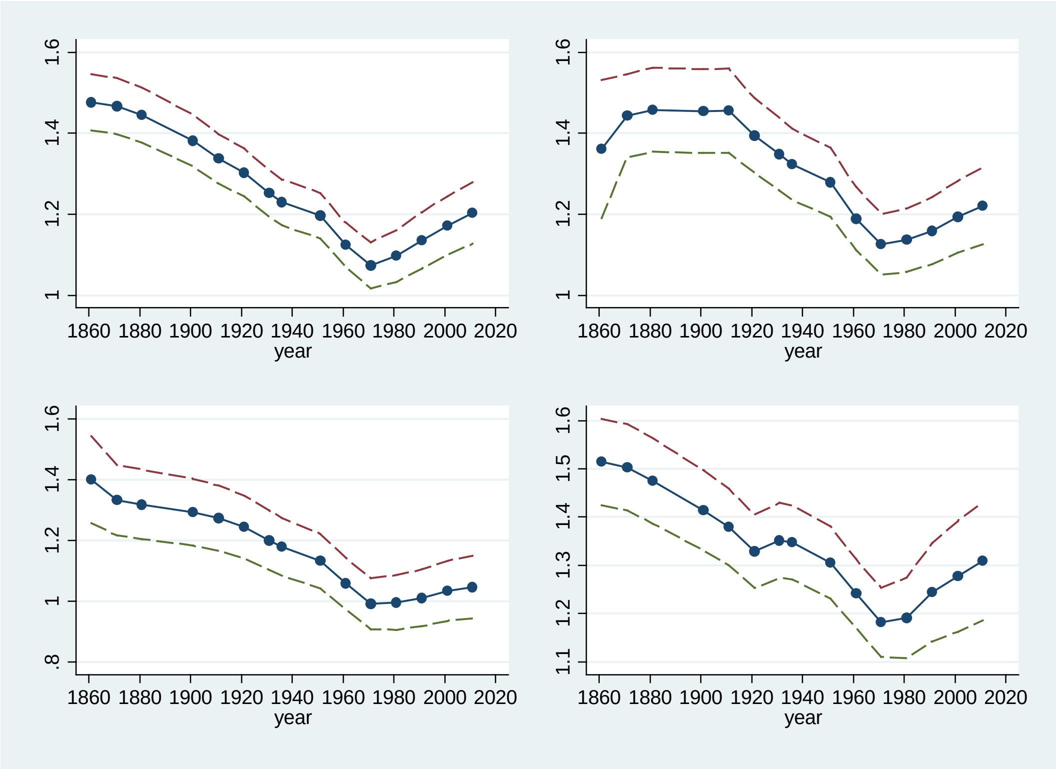

Fig. 6 Shape parameters by macroaregions, computed for all the census years (1861-2011) and population cutoffs defined by UMPU test. Dashed lines represent IG confidence intervals (moving anticlockwise macroregions are NW, NE, CE; SO)

Fig. 7 Shape parameters by macroaregions, computed for all the census years (1861-2011) and population cutoffs defined by population cuoff > 5000 inhabitants. Dashed lines represent IG confidence intervals (moving anticlockwise macroregions are NW, NE,CE;SO)

References

Albanese, G., de Blasio, G., 2013. Civic capital and Development: Italy, 1951-2001, Bank of Italy Economic History Working Papers n. 32

Bee, M., Riccaboni, M., & Schiavo, S. (2012). *A Trick of the (Pareto) Tail* (No. 1206). Department of Economics, University of Trento, Italia.

Combes, P. P., Lafourcade, M., Thisse, J. F., & Toutain, J. C. (2011). The rise and fall of spatial inequalities in France: a long-run perspective. *Explorations in Economic History*, *48*(2), 243-271.

Combes, P. P., Duranton, G., Gobillon, L., Puga, D., & Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, *80*(6), 2543-2594.

Duranton, G., & Puga, D. (2013). *The growth of cities*. Centre for Economic Policy Research, mimeo.

Fazio, G., & Modica, M. (2012). *Pareto or log-normal? A recursive-truncation approach to the distribution of (all) cities* (No. 2012-54). Scottish Institute for Research in Economics (SIRE).

Gabaix, X. (1999). Zipf's law for cities: an explanation. *The Quarterly journal of economics*, *114*(3), 739-767.

Gabaix, X., & Ibragimov, R. (2011). Rank− 1/2: a simple way to improve the OLS estimation of tail exponents. *Journal of Business & Economic Statistics*, *29*(1), 24-39.

Istat (1994), Popolazione residente dei comuni - Censimenti dal 1861 al 1991 (Circoscrizioni territoriali al 20 ottobre 1991), Roma.

Gabaix, X., & Ioannides, Y. M. (2004). The evolution of city size distributions. *Handbook of regional and urban economics*, *4*, 2341-2378.

Eeckhout, J. (2004). Gibrat's law for (all) cities. *The American Economic Review*, *94*(5), 1429-1451.

Ioannides, Y., & Skouras, S. (2013). US city size distribution: Robustly Pareto, but only in the tail. *Journal of Urban Economics*, *73*(1), 18-29.

Malevergne, Y., Pisarenko, V., & Sornette, D. (2011). Testing the Pareto against the lognormal distributions with the uniformly most powerful unbiased test applied to the distribution of cities. *Physical Review E*, *83*(3), 036111.

Puga, D. (1999). The rise and fall of regional inequalities. *European economic review*, *43*(2), 303-334.

Redding, S. J., & Sturm, D. M. (2008). The Costs of Remoteness: Evidence from German Division and Reunification. *The American Economic Review*, *98*(5), 1766-1797.

Rappaport, J. (2007). Moving to nice weather. *Regional Science and Urban Economics*, *37*(3), 375-398.

Rozenfeld, H. D., Rybski, D., Gabaix, X., & Makse, H. A. (2011). The Area and Population of Cities: New Insights from a Different Perspective on Cities. *The American Economic Review*, *101*(5), 2205-2225.