POPULATION GROWTH AND STRUCTURAL CHANGE: ITALY, 1951-2011

Antonio Accetturo
Michele Cascarano
Guido de Blasio[♣]

July 2016

Abstract

This paper presents some evidence on the dynamics of population across all Italian locations (municipalities or LLMs) over the period 1951-2011, with the aim to detect whether Gibrat's law (random growth, i.e. the orthogonality between growth and initial levels) holds. We find that random growth is consistently rejected. From 1951 to 2011, population shifted from initially smaller to relatively larger locations; while the growth rate of very large urban areas was not different from the average one. This result is confirmed for all subperiods and holds when we consider the Centre-North and the South separately. We interpret this result through the lens of structural change, as recently proposed by the literature. We find that structural change away from agriculture and, to a lesser extent, away from manufacturing had a big role in explaining population divergence across locations.

JEL: Gibrat law, Population divergence, Structural change
Keywords: R2, L16

[♣] Bank of Italy. E-mails: antonio.accetturo@bancaditalia.it, michele.cascarano@bancaditalia.it, guido.deblasio@bancaditalia.it

We wish to thank Alberto Dalmazzo and the participants to the Bank of Italy – Urban economics workshop (Rome, December 2015) for useful comments. The usual disclaimers apply. Further, the views expressed are the author's own and do not necessarily reflect those of the Bank of Italy.

1. Introduction

The aim of this paper is to provide some stylized facts on patterns of population growth of Italian cities in the last 60 years. We proceed in two steps. We first evaluate whether Gibrat's law (or random growth), i.e. the orthogonality between growth and initial levels, holds for locations between 1951 and 2011. Then we check the relation between population growth and the deep structural transformations that characterized the Italian economy in the last 60 years.

Our focus on population growth has two main reasons. The first is related to the deep connection between population dynamics and economic factors. While physical geography (mountains, coasts, rivers) may have played a crucial role in early settlements, in the long run, the evolution of the population across geographical locations is mostly led by economic motives.[1] The second refers to the fact that city level data on population are characterized by a good availability and higher quality. For this reason long run analyses in urban economics generally rely on population data, since other economic indicators are more likely to be affected by measurement errors and lack of comparability over time.

Testing Gibrat's law is important for both its positive and normative implications. The size of a city is the result of a complex amalgam between centripetal (agglomeration) and centrifugal (congestion) forces. If the Gibrat's law were confirmed, these two forces perfectly counterbalance in expectation and city size is simply determined by natural advantages or geographical constraints. In principle, this has relevant policy implications; if growth is random, irrespective of past productivity shocks, so to represent a kind of "natural law", all policies aimed at modifying the spatial distribution of population and economic activities are doomed to be effective only in the short run.

Early papers on the topic tended to confirm the hypothesis that city growth is random (Glaeser et al., 1995; Eaton and Eckstein, 1997; Ioannides and Overman, 2003; Eeckout, 2004). These studies, however, generally concentrated on larger cities only or have a limited time depth. The recent availability of more complete datasets for all locations on very long

[1] A well known example is Detroit, whose decline in population coincided with a severe downturn in manufacturing activity. In Italy, similar dynamics were experienced in cities like Trieste or Naples.

time spans has radically changed the empirical picture. Michaels et al. (2012) and Desmet and Rappaport (2015) have shown that US location have never grown randomly over the last 200 years.

Our results on Italy are remarkably similar to those recently found for the US. Gibrat's law is rejected over the period 1951-2011 when we consider both administrative units (cities and town) and local labor markets (LLMs). As for the US, in the last 60 years there was a remarkable shift of population. Initially very small locations consistently lost population over time, while cities of intermediate size grew more than average; the growth rate of very large cities is instead similar to the average one, thus by presence of congestion forces that counterbalance the effects of agglomeration economies. This latter result is compatible with early tests on Gibrat's law, which mostly focused on larger urban zones.

This is the first work, to the best of our knowledge, that studies population growth for *all* Italian locations over a relatively *long time span*. Italy represents an interesting case because of its great variability on economic conditions and geographic economic imbalances over this period of time. Between 1951 and 2011 experienced distinct phases in its economic history, passing from decades of intense growth and industrialization (1951-1971) to marked economic and productivity slowdown (1991-2011). Italy has also the widest and most persistent economic imbalances of advanced economies, with marked economic differences between an economically advanced Centre-North and a lagging South (the so-called *Mezzogiorno*). Population patterns across cities over decades are quite similar, except for the fact that early industrialization (1951-1971) was associated with a high population growth even in very large cities. Population dynamics are also quite similar between the Centre-North and the South.

The rejection of the random growth leads to the second part of the paper. As suggested by the most recent literature (Michaels et al., 2012; Desmet and Rappaport, 2015), the violation of the Gibrat's law may be linked to the structural change away from agriculture that characterized all modern economies in the last 60 years. While agricultural activities are intensive in the use of land and have very limited returns from agglomeration, non-agricultural sectors rely much less on land and enjoy more from face-to-face interactions. This implies that structural change may have made relatively larger locations more economically attractive at the expenses of smaller ones.

We check this hypothesis on Italian data by considering the two main structural transformations observed in the last 60 years: the one away from agriculture from to sixties to the seventies and the one away from manufacturing in the last 30 years. As in Michaels et al. (2012), results show that structural change, especially the one away from agriculture, plays an important role in describing population growth over locations and is in part responsible for the observed population divergence. Structural change has instead a more limited role when we analyze the last two decades.

The paper is organized as follows. Section 2 describes the data. Section 3 presents the test for random growth. Section 4 analyzes the relationship between structural change and population growth. Section 5 concludes.

2. Data and descriptive statistics

This paper uses harmonized population and employment data at city level for all decennial censuses from 1951 to 2011.

The first challenge in this kind of study is the choice of the unit of observation. As explained by Cuberes (2011) both administrative and functional definitions of cities have advantages and disadvantages. For example, administrative boundaries are sometimes arbitrary and lack of economic content. Functional definitions of metropolitan areas, instead, have more economic meaning but they change over time; this makes them less suitable for long run comparisons.

For this reason, we use both administrative and functional boundaries. As for the administrative units, in Italy cities may change over time due to the merges and splits of administrative areas; for this reason we rely on the homogenization recently made available by the Italian Statistical Office (Istat) on the web portal http://ottomilacensus.istat.it/. These data are made comparable from 1951 to 2011 by keeping the start-of-the-period administrative boundaries. Table 1 presents the number of municipalities for each census year.

As for the functional boundaries we use the Istat definition of LLMs. Starting from 1981, Istat started surveying the commuting patterns across municipalities by Italian workers. This allowed constructing commuting matrixes among municipalities. The Istat

LLM is a set of at least two contiguous municipalities characterized by self-contained commuting patterns (at least 75% of local population lives and works in the LLM). For the purposes of this paper, we use the LLMs map based on 1981 commuting patterns for all analyses between 1971 and 2011.[2]

Table 2 presents a number of descriptive statistics for both municipalities and LLMs. Average population growth across cities between 1951 and 2011 is negative (-8.8% in log points); this indicates the prevalence of a number of locations that lost population since in the same period total population in Italy grew by 22% (in log points). The great variability in growth rate is apparent by the standard deviation that is seven times the mean. Figures 1a and 1b show different population growth patterns according to different geographical zones: growth process has been weaker for municipalities and LLMs located in internal mountainous areas, especially in the Apennines, than for those placed in coastal areas. The prevalence of small rural locations in 1951 can be also detected by the average high share of employees in agriculture (57%), that is larger than national figure at that time (41%).

3. Random growth for Italian cities

3.1 Testing Random Growth

Gibrat's law for population growth predicts the orthogonality between growth rates and initial conditions. We test this implication by using both parametric and non-parametric estimates.

Non-parametric estimates are obtained by the following regression:

$$(1) \qquad g_i = m(S_i) + \varepsilon_i$$

where *i* indicates the city or the LLM; g_i is the standardized log growth rate of city *i*; standardized growth rate is the difference between the growth rate and the sample mean,

[2] There are two main sources of arbitrariness. The first is the use of 1971 as a staring year, despite the fact that commuting patterns were registered in 1981. The second is the use of 1981-LLMs for the following years, despite the fact that we know that LLMs changed over time. It should be noted, however, that our results are quite robust to the change of the initial year or the change of the definition of LLMs.

divided by standard deviation. S_i is the log population of location *i* at the start of the period (Ioannides and Overman, 2003).

The objective of this regression is to provide an approximation of the unknown relationship between growth and size using smoothing, without making parametric assumptions about the functional form of $m(S_i)$. We denote the estimate of $m(S_i)$ with $\widehat{m}(S_i)$ that is the local average of the dependent variable around S_i. This local average smooths the value around S_i S_i by using a kernel, that is a continuous weight function symmetric around S_i S_i. The kernel *K* used in the remainder of the paper will be an Epanechnikov kernel, with optimal bandwidth h as computed by Silverman (1986). In formula, the estimator is equal to:

(2) $$\widehat{m}(s) = \frac{n^{-1}\sum_{i=1}^{n} K_h(s-S_i)g_i}{n^{-1}\sum_{i=1}^{n} K_h(s-S_i)}$$

Despite the presence of a large number of observations, non-parametric tests are generally quite sensitive to the presence of outliers. For this reason we also perform a parametric test by running the following regression:

(3) $$g_i = \sum_{j=1}^{10} \beta_j D_{ij} + \xi_i$$

where D_{ij} is a dummy variable equal to one if location *i* belongs of the j-th decile of the log size distribution in the initial period.

3.2 Non parametric results

We start first with non-parametric estimates.

Figure 2, panel (a) presents the kernel estimation for the standardized growth rates of Italian cities between 1951 and 2011; panel (b) shows the same estimation for LLMs over the period 1971 and 2011. Both figures are remarkably similar and decisively reject the random growth hypothesis. Over both time spans middle-size locations registered a positive standardized growth, while smaller location grew less than the average. The turning points are around 5,000 inhabitants for cities and 20,000 for LLMs. Results for bigger municipalities are instead inconclusive due to larger standard errors; it should be noted, however, that point estimates are lower than those registered by middle-sized locations.

Heterogeneity over time. – Panels (c) and (d) present a breakdown of the relationship between growth and size by splitting growth rates into three (for cities) or two (for LLMs) 20 years subperiods. This periodization is particularly meaningful for the Italian case.

From 1951 to 1971 Italy experienced the fastest economic growth in its history (the so-called economic miracle); this period is characterized by a dramatic shift from a mostly agricultural to a manufacturing economy. In the same period Italy registered vast internal migrations from the countryside to urban centers and, in particular, from the North to the South.

In the second period (1971 to 1991) Italy experienced lower but still vibrant economic growth; industrialization continued at fast pace, especially in some areas of the North East and the Center, with the rise of industrial districts.[3] Other cities of ancient industrialization like Milan and Genoa started instead to deindustrialize.

The third period (1991-2011) is characterized by a marked economic slowdown and (especially in the second decade) and an intense deindustrialization.

Panels (c) and (d) clearly show that random growth is rejected in all subperiods. During all decades, small locations have (relatively) lost population in favor of the bigger ones. City level data (panel (c)) also show that between 1951 and 1971 the largest municipalities grew relatively more than all other municipalities; this is not surprising since the ’50s and the ‘60s were characterized by a very strong rural-urban migration.

The following four decades (from 1971 to 2011) observed instead a relevant shift of population from smaller to intermediate locations. This is confirmed both at municipality and LLM level. Quite interestingly, population reshuffling from small to intermediate locations continued in both high (1971-1991) and low (1991-2011) growth periods, without relevant differences between cities and LLMs.

[3] This area was subsequently called “Third Italy” with the aim to distinguish it from the already industrialized North-West and the still lagging South.

Heterogeneity over space. – We finally investigate the existence of heterogeneity across space. Italy is well-known for its huge and persistent economic imbalances between an economically developed Centre-North and a lagging South (known also as *Mezzogiorno*). This may reflect in possible heterogeneities in the test for random growth. Panels (e) and (f) still reject the hypothesis that population growth is random for Italian location: patterns of geographical concentration from small to intermediate locations are confirmed within the Centre-North and the South, for both municipalities and LLMs. However, some differences across macro-regions emerge: turning points from negative to positive growth are markedly larger in the South. In order to achieve a larger-than-average population growth, municipalities should be greater than 2,000 inhabitants in the Centre-North and 8,000 in the South; for LLMs these figures are, respectively, 14,000 and 35,000. This seems to suggest the existence of negative externalities hampering population growth in smaller southern locations.

3.3 Parametric results

We now turn to the estimates of equation (3).

Results for municipalities are displayed in table 3 for the entire period (1951-2011, first column) e for each 20-years subperiods (second to fourth columns). Parametric tests reject, once again, random growth: estimates for the first decile of the initial log population is always negative and significant. Very small locations grew 41 (log) percentage points less than the average over the period 1951-2011; this is confirmed for all subperiods. Dummies for other coefficients are bigger the larger the decile of initial size.

Heterogeneity in the behavior of the largest cities according to the subperiods is also apparent in the econometric test. The second column (for the period 1951-1971) shows that the dummy for the 10th decile is bigger than the one for the 9th. This relationship is reversed in the third and fourth columns, thus indicating a lower growth for very large locations from 1971 on.[4]

---

[4] F-tests show that these differences are all statistically significant.

A similar picture emerges once we focus on LLMs (table 4), for the period 1971-2011 (first column) and for the two subsequent subperiods (columns two and three).

4. Population growth and structural change

4.1 Detecting the effects of structural change

Gibrat's law is consistently rejected for Italian locations over a long time span. The aim of this section is to analyze this result through the lens of the effects of structural change on the location of population.

Michaels et al. (2012) highlight the role of the shift from agricultural to non-agricultural activities in explaining the rejection of random growth from 1880 to 2000 in the US. They present a model in which agricultural activities are intensive in the use of land and have very limited returns from agglomeration; non-agricultural sectors are instead more labor intensive and enjoy more from face-to-face interactions. Structural change (i.e. the shift of economic specialization away from agriculture) dramatically modifies the relative aggregate productivity of each location. Smaller locations (with a comparative advantage in agriculture, due to their land abundance) are likely to lose, while more populated areas are more able to exploit agglomeration economies and, hence, are more likely to register economic (and population) growth.

In this section we present some regression results on the relationship between population growth and structural change. The characteristics of Italian structural change are described in table 4. In 1951 40% of Italian workers were employed in agriculture; the corresponding share for manufacturing was 32%. From 1951 to 1971 agricultural share more than halved, while manufacturing share rose by 12 percentage points reaching its historical peak (44% in 1971). Starting from 1971, the share of industrial workers started to decrease, with a more intense fall in the period 2001-2011.

From an empirical point of view, as in Michaels et al. (2012), we measure the impact of structural change with the share of agriculture and manufacturing at the start of the period; this is aimed at reducing the impact of simultaneity biases in first differenced regressions (Baum-Snow and Ferreira, 2015).

The correlation between structural change and population growth is detected by estimating the following equation:

$$(4) \qquad g_i = \alpha + \beta S_i + \gamma_1 sh - agr_i + \gamma_2 sh - man_i + \lambda D - south_i + \varepsilon_i$$

where g_i and S_i denote the same variables explained in the previous section. $sh - agr_i$ and $sh - man_i$ are, respectively, the share of agriculture and the share of manufacturing in location i at the start of the period. $D - south$ is a dummy variable for southern municipalities.

Estimation of equation (4) is made by OLS; as it solely relies on the endogenous variability within the data, causality is not warranted and estimated coefficients should be considered as conditional correlations between the variables.

Parameter β represents the correlation between initial size and growth; it is equal to zero when population growth is random, while it is positive (negative) in case of divergence (convergence).

Parameters γ_1 and $\gamma_2$describe the relation between the initial industrial structure (and, thus, structural change) and population growth. Statistically significant coefficients imply that initial structure is associated with population dynamics due, for example, to cross-area differences in productivity and wages.

4.2 Results

Estimation results of equation (4) for municipalities are reported in table 5.

Panel (a) shows the estimates for the entire period (1951-2011). The first column reports the simple correlation between growth and initial population without controls for structural change; the coefficient is positive confirming the population divergence we observed in the previous section. A 10% increase in local population is associated with a rise in the relative population growth of 2.3 percentage points. The second column adds controls for agriculture: initial share is negatively correlated with population growth. This fact confirms the effects of structural change on population dynamics (Michaels et al., 2012). The third column includes controls for manufacturing: the positive effect of the initial industrial share is quite small: this is not surprising considering that the period between 1951 and 2011 is time characterized by opposing waves of industrialization and deindustrialization. The last column adds the dummy for the South. In this case population growth has a negative correlation with the South dummy, also reflecting the South-North migration patterns over the years, while coefficients for structural change remain basically unchanged.

The signs of the correlations between initial industrial structure and population growth are roughly confirmed across subperiods. The initial agriculture share impacts negatively during all the 20-years subperiods and its effect is stronger in the first two decades (1951-1971, panel (b)); the effect of the initial manufacturing share is not significant for the first 20 years, while it becomes strongly positive and significant in the subsequent periods, especially over the period 1971-1991 (panel (c)).

We calculate the impact of structural change in explaining population divergence by looking at the sensitivity of the coefficient of initial population across specifications. For the entire period (1951-2011) structural change impacts for almost 28% (=(0.2303-0.1656)/0.2303) of the observed population divergence. The importance of structural change for population divergence changes over time. From 1951 to 1971 45% (=(0.2777-0.1505)/0.2777) of observed divergence is explained by structural change; this share reduces to 22% (=(0.2664-0.2060)/0.2664) from 1971 to 1991 and to 6% (=(0.1654-0.1555)/0.1654) in the period 1991-2011.

Results for LLMs (table 6) basically confirms these results. From 1971 to 2011 (panel (a)) initial economic structure was an important driver for population dynamics, with a negative correlation for the initial specialization in agriculture and a positive correlation for the initial manufacture structure. The initial agriculture share becomes less relevant (but still significant) when we add the dummy South. Similarly to municipalities, the effect of structural change explains away part of the divergence across areas since the estimated coefficient for the initial population passes from 0.336 of column 1 to 0.264 of column 4. Back of the envelope calculations show that, contrary to the estimates for municipalities, the effect of initial industry composition on population growth does not fade away in the final 20 years.

Summing up the results, tables 5 and 6 show that structural transformations are important determinants for local population growth and partially explain the observed divergence in population growth across locations. As in Michaels et al. (2012), a major role is played by structural change away from agriculture. City-level estimates show that the effect of structural change in explaining divergence fades away over time, while LLMs estimations show that the impact is relevant throughout the period. A possible explanation for this different result is that, over time, municipalities within a LLM have become more homogeneous due to both suburbanization (i.e. the relocation of population in the suburbs) and a relocation of industrial activities in other areas within the same LLM. Conversely, LLMs are still characterized by marked differences in terms of sectoral specialization, with relevant effects on population trends.

Table 7 reports the estimates for a subsample of southern municipalities; main results are confirmed, highlighting the role of structural change away from agriculture in explaining population divergence. Differently from the nation-wide estimates, the share of industrial activities has a negative sign in the 1951-2011 and 1951-1971 regressions: this is not surprising, considering that the intense process of market integration in the ‘50s and the ‘60s (driven by the opening of major highways on the North-South axis) generated the disappearance of several, inefficient Southern industrial districts. Quantitatively, compared with nation-wide estimates, structural change in the South explains a larger share of the violation of the Gibrat’s law: over the period 1951-2011 initial industrial shares impact for 42% of total population divergence, while, for the latest two decades (1991-2011) the share is 13%.

Table 8 shows the estimates for southern LLMs, broadly confirming previous results.

5. Concluding remarks

In this paper we present some evidence on the dynamics of population across all Italian locations (municipalities or LLMs) over the period 1951-2011. Our aim is to detect whether Gibrat’s law holds for Italy over a relatively long (and quite diversified) period of time.

We find that random growth is consistently rejected. From 1951 to 2011, population shifted from initially smaller to relatively larger locations; while the growth rate of very large urban areas was not different from the average one. This result is confirmed for all subperiods and holds when we consider the Centre-North and the South separately.

We interpret this result through the lens of structural change, as recently proposed by the literature. We find that structural change away from agriculture and, to a lesser extent, industrialization had a big role in explaining population divergence across locations. The effect was particularly strong until 1971, while it weakened in the last two decades, and it is robust to the inclusion of controls for human capital or geographical dummies.

In interpreting these results two things should be kept in mind.

First, we just show conditional correlations, in which problems of omitted variables or reverse causality might be pervasive. While we think that a good description of data is important for a primer in analyzing economic issues, we are aware that detecting causality is important when we derive policy implications. We leave this problem to future research.

Second, the structural change interpretation for population dynamics has recently become quite popular. In our estimates we find that it is able to explain an important part of the observed divergence (45% in the ‘50s and the ‘60s); we also find that its role is decreasing over time, probably due to the fact that the transition away from agriculture has terminated. However, observed divergence is still robust across locations, even in the most recent decades. A possible explanation, that we leave for future research, is linked to the path dependence in the location of population. Individuals may react quite slowly (and with several lags) to local economic shocks; this implies that “old” systems of cities may perpetuate over time despite the fact that some locations have completely lost their locational advantages. Michaels and Rauch (2015), for example, find that this is a feature that characterizes French cities in comparison with and British urban centers with large detrimental effects on aggregate growth.

References

Baum-Snow N. and F. Ferreira (2015): "Causal inference in urban economics," in G. Duranton, J. V. Henderson, and W. Strange (eds.) *Handbook of Regional and Urban Economics*, vol. 5.

Cuberes D. (2011): "Sequential City Growth: Empirical Evidence," Journal of Urban Economics, vol. 69.

Desmet K., J. Rappoport (2015): "The Settlement of the United States, 1800-2000: the Long Transition Toward Gibrat's Law," Forthcoming Journal of Urban Economics.

Eaton J., Z. Eckstein (1997): "Cities and Growth: Theory and Evidence from France and Japan," Regional Science and Urban Economics, 1997, 27(4 –5), pp. 443– 74.

Eeckhout J. (2004): "Gibrat's Law for (All) Cities," American Economic Review, 94, 1429–1451.

Glaeser E., J. Scheinkman, A. Shleifer (2005): "Economic Growth in a Cross-Section of Cities," Journal of Monetary Economics, 36(1), pp. 117– 43.

Ioannides Y., and H. Overman (2003): "Zipf's Law for Cities: An Empirical Exam- ination." Regional Science and Urban Economics, 33(2), pp. 127–37.

Michaels G., F. Rauch, S. Redding (2012): "Urbanization and Structural Transformation," The Quarterly Journal of Economics, vol. 127.

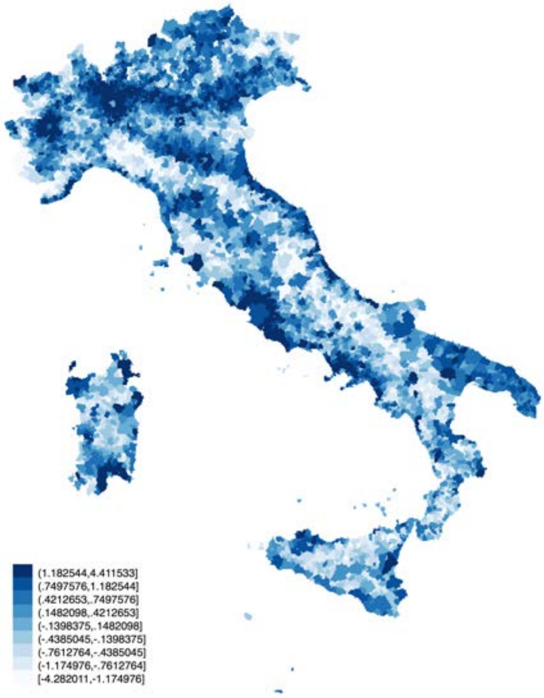
Michaels G., F. Rauch (2014): "Resetting the Urban Network: 117-2012," Forthcoming, The Economic Journal.

Silverman B. (1996): "Density estimation for statistics and data analysis," New York: Chapman and Hall.

Figure 1

POPULATION GROWTH

(a) Cities (1951-2011)



(b) LLM (1971-2011)

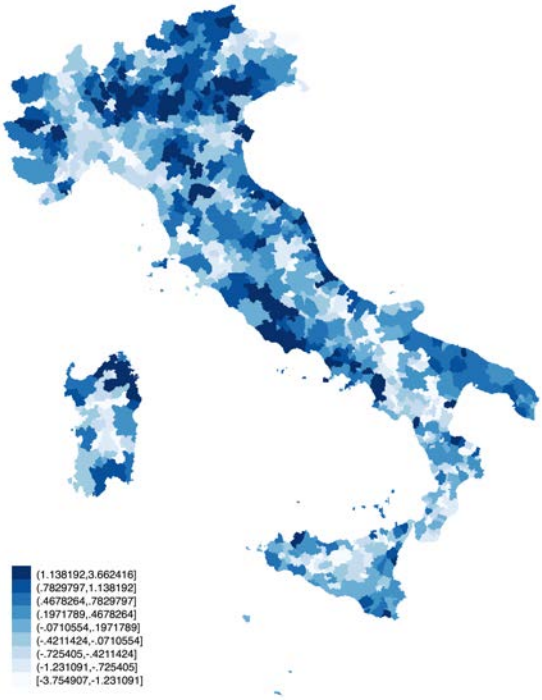
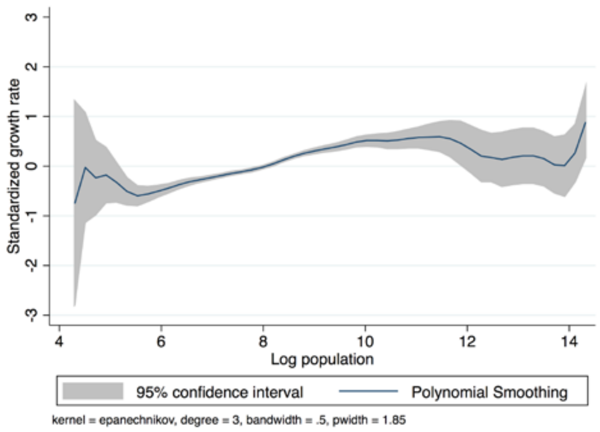


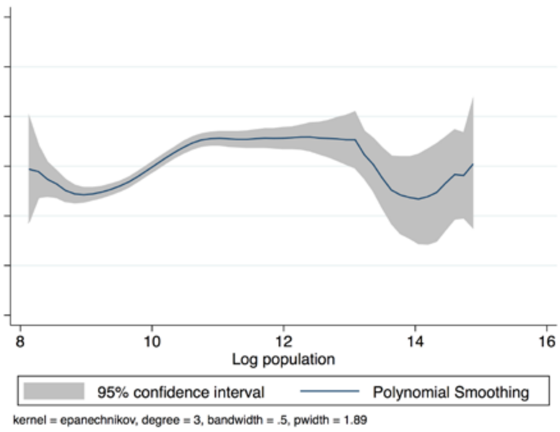
Figure 2

GROWTH AND CITY SIZE

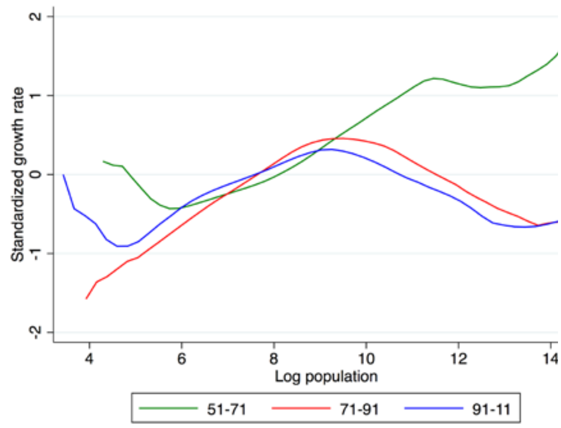
(a) Cities (1951-2011)



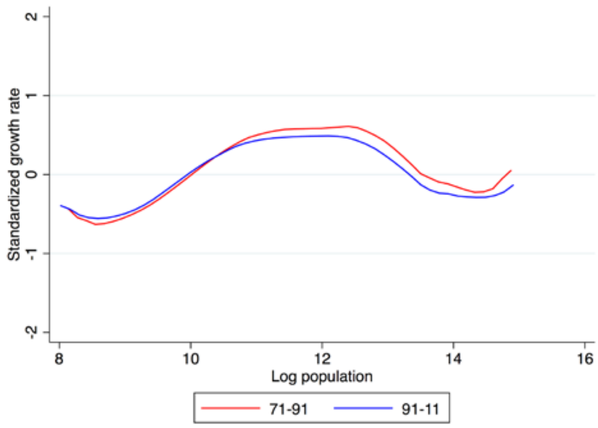
(b) LLM (1971-2011)



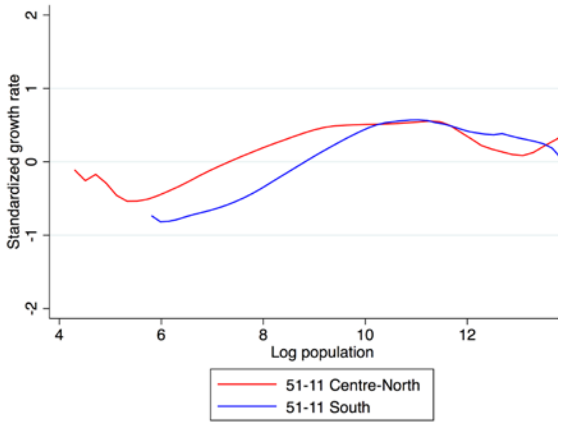
(c) Cities (20 years subperiods)



(d) LLM (20 years subperiods)



(e) Cities (51-11: CN-MZG)



(f) LLM (71-11: CN-MZG)

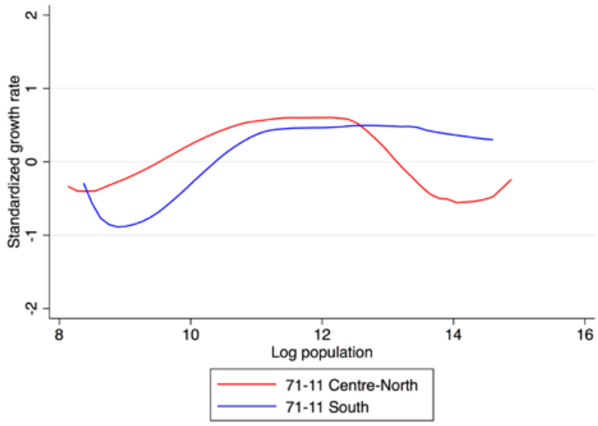


Table 1

NUMBER OF MUNICIPALITIES BY YEAR

1951	7,810
1961	8,035
1971	8,056
1981	8,086
1991	8,100
2001	8,101
2011	8,092

Table 2

DESCRIPTIVE STATISTICS

	No. Obs.	Mean	Standard Deviation
		Municipalities	
Log pop. Growth (1951-2011)	7,761	-0.088	0.630
Log pop. (1951)	7,761	7.945	1.040
Share Agr. 1951	7,761	57.396	24.070
Share Man. 1951	7,761	27.606	20.087
		LLMs	
Log pop. Growth (1971-2011)	955	0.023	0.238
Log pop. (1971)	955	10.159	1.052
Share Agr. 1971	955	29.000	16.280
Share Man. 1971	955	41.139	13.191

Table 3

RANDOM GROWTH FOR CITIES

Dependent variable: Log population growth	Cities (1951-2011)	Cities (1951-1971)	Cities (1971-1991)	Cities (1991-2011)
Constant (1st decile)	-0.4144***	-0.3848***	-0.7078***	-0.5390***
	[0.0401]	[0.0336]	[0.0431]	[0.0436]
Dummy 2nd decile	0.1888***	0.1462***	0.3358***	0.2882***
	[0.0554]	[0.0468]	[0.0555]	[0.0564]
Dummy 3rd decile	0.2497***	0.1705***	0.4863***	0.3855***
	[0.0547]	[0.0477]	[0.0565]	[0.0589]
Dummy 4th decile	0.2804***	0.2287***	0.5798***	0.4985***
	[0.0538]	[0.0471]	[0.0548]	[0.0570]
Dummy 5th decile	0.3603***	0.3120***	0.6584***	0.5317***
	[0.0543]	[0.0498]	[0.0526]	[0.0563]
Dummy 6th decile	0.3827***	0.3003***	0.7781***	0.6037***
	[0.0524]	[0.0480]	[0.0536]	[0.0561]
Dummy 7th decile	0.4918***	0.4197***	0.9125***	0.6585***
	[0.0526]	[0.0478]	[0.0531]	[0.0538]
Dummy 8th decile	0.6159***	0.5263***	1.0587***	0.8374***
	[0.0517]	[0.0486]	[0.0534]	[0.0524]
Dummy 9th decile	0.7263***	0.7005***	1.2023***	0.9021***
	[0.0513]	[0.0493]	[0.0526]	[0.0510]
Dummy 10th decile	0.8496***	1.0459***	1.0687***	0.6846***
	[0.0473]	[0.0474]	[0.0490]	[0.0487]
R^2	0.061	0.085	0.125	0.064
No. Obs.	7761	7784	8053	8076

Table 3

RANDOM GROWTH FOR LLMS

Dependent variable: Log population growth	LLMs (1971-2011)	LLMs (1971-1991)	LLMs (1991-2011)
Constant (1st decile)	-0.4735***	-0.5907***	-0.5606***
	[0.1103]	[0.1048]	[0.1154]
Dummy 2nd decile	-0.0454	0.0765	-0.0000
	[0.1588]	[0.1529]	[0.1645]
Dummy 3rd decile	-0.0588	0.1023	0.2428
	[0.1539]	[0.1477]	[0.1629]
Dummy 4th decile	0.1958	0.3347**	0.2043
	[0.1456]	[0.1418]	[0.1460]
Dummy 5th decile	0.2720*	0.4450***	0.4501***
	[0.1530]	[0.1455]	[0.1589]
Dummy 6th decile	0.5092***	0.5967***	0.7811***
	[0.1408]	[0.1369]	[0.1400]
Dummy 7th decile	0.8957***	0.9426***	0.9693***
	[0.1311]	[0.1282]	[0.1435]
Dummy 8th decile	0.9915***	1.1503***	0.9466***
	[0.1379]	[0.1342]	[0.1357]
Dummy 9th decile	1.0220***	1.1411***	1.0859***
	[0.1298]	[0.1275]	[0.1361]
Dummy 10th decile	0.9554***	1.1213***	0.9286***
	[0.1296]	[0.1296]	[0.1322]
R^2	0.188	0.196	0.165
No. Obs.	955	955	955

Table 4

STRUCTURAL CHANGE

	Share of Agriculture	Share of Manufacturing
1951	40.8	32.2
1961	28.0	40.5
1971	17.1	43.6
1981	12.8	40.4
1991	7.8	33.6
2001	6.0	32.3
2011	6.0	26.3

Table 5

STRUCTURAL CHANGE FOR CITIES

		(a) 1951-2011		
Log initial population	0.2303***	0.1418***	0.1484***	0.1656***
	[0.0100]	[0.0102]	[0.0103]	[0.0109]
Agricultural share		-0.0179***	-0.0149***	-0.0138***
		[0.0004]	[0.0012]	[0.0012]
Industrial share			0.0037***	0.0038***
			[0.0014]	[0.0014]
Dummy South				-0.1364***
				[0.0242]
Constant	-1.8301***	-0.1020	-0.4275***	-0.5830***
	[0.0823]	[0.0936]	[0.1516]	[0.1555]
R^2	0.057	0.234	0.234	0.238
No. Obs.	7761	7761	7761	7761
		(b) 1951-1971		
Log initial population	0.2777***	0.1772***	0.1734***	0.1505***
	[0.0097]	[0.0091]	[0.0096]	[0.0102]
Agricultural share		-0.0203***	-0.0221***	-0.0235***
		[0.0005]	[0.0012]	[0.0012]
Industrial share			-0.0022	-0.0023
			[0.0015]	[0.0015]
Dummy South				0.1811***
				[0.0219]
Constant	-2.2060***	-0.2421***	-0.0503	0.1570
	[0.0783]	[0.0864]	[0.1530]	[0.1566]
R^2	0.084	0.312	0.312	0.318
No. Obs.	7784	7784	7784	7784

Table 5 (cont.ed)

	(c) 1971-1991			
Log initial population	0.2664***	0.2004***	0.2152***	0.2060***
	[0.0099]	[0.0104]	[0.0101]	[0.0104]
Agricultural share		-0.0138***	-0.0050***	-0.0057***
		[0.0005]	[0.0010]	[0.0010]
Industrial share			0.0115***	0.0116***
			[0.0011]	[0.0011]
Dummy South				0.0799***
				[0.0259]
Constant	-2.0751***	-1.1427***	-2.0229***	-1.9603***
	[0.0798]	[0.0891]	[0.1186]	[0.1209]
R^2	0.100	0.169	0.181	0.182
No. Obs.	8053	8053	8053	8053
	(d) 1991-2011			
Log initial population	0.1654***	0.0881***	0.1124***	0.1555***
	[0.0088]	[0.0092]	[0.0091]	[0.0094]
Agricultural share		-0.0285***	-0.0191***	-0.0148***
		[0.0009]	[0.0011]	[0.0011]
Industrial share			0.0134***	0.0058***
			[0.0009]	[0.0009]
Dummy South				-0.5925***
				[0.0255]
Constant	-1.2897***	-0.2818***	-1.1258***	-1.0378***
	[0.0725]	[0.0795]	[0.0981]	[0.0954]
R^2	0.045	0.157	0.181	0.235
No. Obs.	8076	8076	8076	8076

Table 6

STRUCTURAL CHANGE FOR LLMS

	(a) 1971-2011			
Log initial population	0.3361***	0.2215***	0.2378***	0.2638***
	[0.0313]	[0.0339]	[0.0341]	[0.0362]
Agricultural share		-0.0172***	-0.0112***	-0.0065*
		[0.0020]	[0.0033]	[0.0035]
Industrial share			0.0089**	0.0078**
			[0.0037]	[0.0037]
Dummy South				-0.2900***
				[0.0713]
Constant	-3.4142***	-1.7509***	-2.4574***	-2.6963***
	[0.3247]	[0.3788]	[0.4756]	[0.4952]
R^2	0.125	0.189	0.195	0.208
No. Obs.	955	955	955	955
	(b) 1971-1991			
Log initial population	0.3632***	0.3138***	0.3271***	0.2992***
	[0.0321]	[0.0347]	[0.0354]	[0.0362]
Agricultural share		-0.0074***	-0.0025	-0.0077**
		[0.0021]	[0.0034]	[0.0035]
Industrial share			0.0073**	0.0085**
			[0.0037]	[0.0037]
Dummy South				0.3130***
				[0.0741]
Constant	-3.6905***	-2.9728***	-3.5505***	-3.2931***
	[0.3309]	[0.3829]	[0.4949]	[0.4924]
R^2	0.146	0.158	0.162	0.177
No. Obs.	956	956	956	956
	(c) 1991-2011			
Log initial population	0.2982***	0.1432***	0.2094***	0.2260***
	[0.0288]	[0.0295]	[0.0307]	[0.0310]
Agricultural share		-0.0415***	-0.0246***	-0.0153***
		[0.0032]	[0.0042]	[0.0042]
Industrial share			0.0207***	0.0099***
			[0.0028]	[0.0031]
Dummy South				-0.6297***
				[0.0695]
Constant	-3.0320***	-0.8709***	-2.5238***	-2.1780***
	[0.3020]	[0.3280]	[0.4096]	[0.4029]
R^2	0.109	0.255	0.298	0.354
No. Obs.	955	955	955	955

Table 7

STRUCTURAL CHANGE FOR CITIES (SOUTH)

		(a) 1951-2011	
Log initial population	0.3538***	0.2335***	0.2039***
	[0.0184]	[0.0231]	[0.0235]
Agricultural share		-0.0152***	-0.0275***
		[0.0013]	[0.0027]
Industrial share			-0.0192***
			[0.0037]
Constant	-3.1545***	-1.1188***	0.3188
	[0.1558]	[0.2598]	[0.3666]
R^2	0.120	0.184	0.196
No. Obs.	2452	2452	2452
		(b) 1951-1971	
Log initial population	0.3185***	0.2206***	0.1823***
	[0.0166]	[0.0202]	[0.0202]
Agricultural share		-0.0124***	-0.0283***
		[0.0012]	[0.0023]
Industrial share			-0.0248***
			[0.0031]
Constant	-2.7019***	-1.0447***	0.8155***
	[0.1395]	[0.2344]	[0.3163]
R^2	0.126	0.181	0.206
No. Obs.	2453	2453	2453
		(c) 1971-1991	
Log initial population	0.3612***	0.3093***	0.3144***
	[0.0185]	[0.0197]	[0.0195]
Agricultural share		-0.0111***	-0.0081***
		[0.0012]	[0.0020]
Industrial share			0.0055**
			[0.0025]
Constant	-2.9806***	-2.0991***	-2.4517***
	[0.1504]	[0.1824]	[0.2486]
R^2	0.147	0.179	0.181
No. Obs.	2511	2511	2511
		(d) 1991-2011	
Log initial population	0.3254***	0.2773***	0.2845***
	[0.0150]	[0.0158]	[0.0157]
Agricultural share		-0.0187***	-0.0165***
		[0.0014]	[0.0015]
Industrial share			0.0058***
			[0.0019]
Constant	-3.1430***	-2.3707***	-2.6355***
	[0.1239]	[0.1406]	[0.1609]
R^2	0.173	0.236	0.239
No. Obs.	2554	2554	2554

Table 8

STRUCTURAL CHANGE FOR LLMS (SOUTH)

		(a) 1971-2011	
Log initial population	0.4877***	0.3952***	0.3601***
	[0.0480]	[0.0555]	[0.0580]
Agricultural share		-0.0150***	-0.0234***
		[0.0038]	[0.0054]
Industrial share			-0.0182**
			[0.0079]
Constant	-5.2368***	-3.7110***	-2.4151***
	[0.5018]	[0.6602]	[0.8965]
R^2	0.211	0.243	0.254
No. Obs.	389	389	389
		(b) 1971-1991	
Log initial population	0.4904***	0.4079***	0.3853***
	[0.0493]	[0.0564]	[0.0596]
Agricultural share		-0.0134***	-0.0188***
		[0.0041]	[0.0058]
Industrial share			-0.0117
			[0.0083]
Constant	-4.9262***	-3.5668***	-2.7311***
	[0.5153]	[0.6748]	[0.9337]
R^2	0.190	0.212	0.216
No. Obs.	389	389	389
		(c) 1991-2011	
Log initial population	0.4427***	0.3307***	0.3336***
	[0.0420]	[0.0459]	[0.0475]
Agricultural share		-0.0268***	-0.0262***
		[0.0046]	[0.0052]
Industrial share			0.0016
			[0.0056]
Constant	-5.0441***	-3.3670***	-3.4489***
	[0.4365]	[0.5231]	[0.6181]
R^2	0.253	0.326	0.326
No. Obs.	389	389	389