Discussion Paper

n. 105

Davide Fiaschi - Lisa Gianmoena - Angela Parenti

The Dynamics of Labour Productivity across Italian Provinces: Convergence and Polarization

Abstract

This paper analyses the dynamics of labour productivity across Italian Provinces in the period 1995-2006. Inequality decreased but a clear pattern of polarization emerged, with the formation of a cluster of high-productive provinces in the North and Center-West of Italy and a cluster of low-productive provinces in the South and in the Center-East. A core of provinces belonging to five regions (Lombardy, Veneto, Emilia-Romagna, Tuscany and Lazio) appears to benefit of a higher growth of productivity. This regional component favoured both inequality and polarization, while the initial level of productivity decreased inequality but increased polarization.

Classificazione JEL: C21; R11; O47; O52

Keywords: distribution dynamics, spatial dependence, output composition, entrepreneurial fabric, human capital

Contents

I. Introduction

Italian Provinces present a high heterogeneity in terms of industrialization, economic development, productive and entrepreneurial fabrics.[1] Moreover, there exists a clear geographical pattern, being the most of high-productive provinces located in the North and in the Center of Italy. This paper analyzes the dynamics of labour productivity across Italian Provinces in the period 1995-2006. In particular, we try to answer the following questions: Are the Italian Provinces converging in terms of productivities? Which is the evolution of the distribution of productivity? How the observed spatial pattern is persistent over time? Which factors affect the growth rate of productivity and the evolution of distribution?

We find evidence of a decrease in inequality and a concomitant emergence of a pattern of polarization, with the formation of a cluster of high-productive provinces in the North and Center-West of Italy and a cluster of low-productive provinces in the South and Center-East. These two clusters are mainly characterized by differences in export and import (as share of GVA), economic density, composition of output (mainly in favour of industry in high-productive provinces and services in low-productive provinces), entrepreneurial fabric, and firm size distribution.

Growth regressions highlight how the growth rate of employment exerted a negative impact on the growth of productivity. The same holds for the number of firms with only one employee, i.e. self-employer, per inhabitant in 1996 (a proxy for entrepreneurial fabric), suggesting possible misallocations in labour market, and for the share of export on GVA in 1995, a probable result of the impossibility to use devaluation to increase competitiveness after Italy's entrance into Monetary Union.

On the contrary, the share of resident population six years old or older with at least a tertiary education in 2001[2], a proxy for

[1] The classification criterion of Italian Provinces corresponds to the Eurostat NUTS 3 classification.

[2] Data are taken 2001 Census

the human capital of labour force, had a positive impact (but its coefficient is statistically significant only at 11% significance level). Lastly, regional dummies point out the presence of a core of provinces belonging to five regions (Lombardy, Veneto, Emilia-Romagna, Tuscany and Lazio) which appear particularly benefited by a regional component.

We do not find any evidence of spatial dependence on the growth rate of productivity. We also test the possibility (and rule out it) that estimates are biased by endogeneity arising from the potential reverse causality effect of growth rate of productivity on growth rate of employment.

We find evidence of σ and conditional convergence.[3] However, the analysis of counterfactual distribution for 2006 calculated "factoring out" any difference in the initial level of productivity (i.e. the distribution there would have been if all provinces had had the same level of productivity in 1995) points out that convergence happened towards the two clusters and within each cluster, but not between clusters. In this respect conditional convergence appears to increase polarization.

Regional component favoured inequality and polarization across provinces. Indeed, regional dummies account for almost the half of inequality across provinces as measured by Gini index, and the counterfactual distribution calculated "factoring out" any regional difference appears single peaked in contrast with the twin-peaked actual distribution. The growth rate of employment, the share of export on GVA, self-employers per inhabitants and tertiary education did not have any significant distributional impact.

The paper is organized as follows. Section 2 contains a review of literature. Section 3 discusses the estimate of the distribution dynamics of the productivity of Italian Provinces. Section 4 reports the estimate of growth regressions, and Section 5 the analysis of the determinants of distribution dynamics. Finally, Section 6 concludes. Appendix gathers the list of Italian Provinces used in the analysis,

[3]The estimated speed of convergence appears very fast between 7.1% and 9.7% (see for comparison Barro and Sala-i-Martin (2004)).

some descriptive statistics of sample and the analysis of endogeneity.

II. Related Literature

So far literature has neglected the dynamics of labour productivity of Italian Provinces focusing on the dynamics of per capita income. In this regard, Fabiani and Pellegrini (1997) and Arbia and Basile (2005) show that in the period 1970-2000 there is no evidence of absolute and σ-convergence across per capita income of Italian Provinces (while for the period 1950-1970 there is evidence of such convergence). Partially contrasting this evidence is Magrini (2007) who finds convergence for the period 1996-2002, even though he highlights the presence and the persistence of a twin-peaked distribution.[4] Such distribution dynamics is also found in Fabiani and Pellegrini (1997), whose estimation of transition matrix for the period 1970-1992 shows a high level of persistence and a bimodal ergodic distribution.

As regards conditional convergence, Fabiani and Pellegrini (1997) and Forni and Paba (2000) find convergence in 1970-1992 and 1971-1991 respectively. In particular, Fabiani and Pellegrini (1997) find a statistically significant impact of illiteracy rate (negative), geographical concentration of banks (positive, but only for the provinces of the Center and North of Italy) and the share of workers in agriculture (positive). Forni and Paba (2000) find the same negative impact of illiteracy, with the additional findings that the share of population with a technical education in 1991 has a positive impact, as well as the share of employment in industrial districts and the average size of firms in 1971. They also find evidence that some demographic, political and social variables, as the share of young population, political participation, the concentration of voting and

[4] Magrini (2007) also studies the convergence taking as unit of observation the Local Labour Systems (LLSs). The Local Labour Systems (Sistemi Locali del Lavoro) are defined as "small areas characterized by internal commuting patters that produce a self-contained labor market" (see ISTAT (1997)). Their number amounts to 784 in 1991. He finds that per capita income across LLSs are converging for the period 1996-2002; distribution of per capita income in 1996 appears to be not unimodal, while the estimated long-run distribution appears bimodal.

the investment subsidize by Government, have some explanatory power.

Literature on the dualistic development across Italian Regions (see, e.g., Paci and Pigliaru (1995)) suggests the presence of strong spatial effects also across Italian Provinces. Indeed, Fabiani and Pellegrini (1997) find evidence that the distance from Milan has a negative effect on the growth of per capita income for provinces in the South and positive for provinces in the Center and in the North of Italy; and Forni and Paba (2000) find that geographical dummies for Center, Northeast and South have negative coefficients. Finally, Arbia and Basile (2005) show that over the period 1971-2000 two spatial regimes exist, and convergence occurred only within these two subgroups of Italian Provinces.

At regional level, i.e. NUTS 2, the dynamics of labour productivity has been the objective of several studies. Cellini and Scorcu (1997) find that the absolute convergence across Italian Regions stopped at the begin of 1980. Accordingly, Paci and Saba (1998) find no convergence in the period 1975-1993, as well as Di Liberto et al. (2008) for the period 1981-1993. As for the explanatory variables of growth rates of productivity, Cellini and Scorcu (1997) find that investment rates, secondary school enrolment and the growth rate of employment do not have a statistically significant impact, while public expenditure has a significant negative impact in the period 1970-1991. In a fixed-effect panel Aiello and Scoppa (2007) find the same result of no impact of investment rates in the period 1980-2002, but differently a significant and positive impact of human capital. Paci and Saba (1998) find that dummies of Southern and Adriatic regions are very significant and interpret this finding as evidence in favour of a persistent dualism across Italian regions. Aiello and Scoppa (2007) find evidence of a strong (unexplained) heterogeneity across Italian Regions.

Finally, Paci and Pigliaru (1995) find a very robust impact of the structural change (more precisely, the change in manufacturing share) on the growth rate of per capita income.

III. Convergence and Polarization in the Distribution of Productivity

In this section we analyse the distribution dynamics of labour productivity across Italian Provinces in the period 1995-2006. Data used in the analysis come from the Italian official statistics provided by ISTAT (National Institute of Statistics) and Bank of Italy, and refer to 103 Italian Provinces.[5]

III.A. Convergence

As discussed in Durlauf *et al.* (2005) the usual methodology used to study convergence, i.e. the estimate of a parametric model with average growth rate explained by the initial level of productivity, could be misleading in order to detect the effective distribution dynamics of productivity. Nonparametric method, instead, appears a more appropriate approach (see Fiaschi and Lavezzi (2007)). Eqq. (1) and (2) are respectively the parametric and nonparametric specifications of the estimate of convergence:

$$AV.PROD.GR_i = \alpha + \beta \log\left(PROD.REL.1995_i\right) + \varepsilon_i \qquad (1)$$

and

$$AV.PROD.GR_i = \alpha + s\left(\log\left(PROD.REL.1995_i\right)\right) + \varepsilon_i, \qquad (2)$$

where $AV.PROD.GR_i$ is the average growth rate of province i in 1995-2006, $PROD.REL.1995_i$ is relative (to the sample mean) productivity of province i in 1995, ε_i is a random component and $s(.)$ in Eq. (2) is a unknown function, i.e. the smooth term, to be estimated (see Wood (2006)).

The estimate of the parametric specification reported in Table 1 indicates the presence of (absolute) convergence, being $\hat{\beta}$ negative and statistically different from zero (see Magrini (2007) for a similar

[5]The list of variables are reported in Appendix A, wile the list of provinces is reported in Appendix B. Sardinia includes only four provinces: Sassari, Nuoro, Cagliari and Oristano.

Estimate of Eq. (1)		Estimate of Eq. (2)	
Param. coeff.		Param. coeff.	
$\hat{\alpha}$	0.005***	$\hat{\alpha}$	0.005***
$\hat{\beta}$	-0.026***	-	-
		Smooth term	*EDF*
		$log(PROD.REL.1995)$	7.9***
	Obs. = 103	GCV(x10^3) = 0.014	
	$\bar{R}^2 = 0.36$	Dev. exp = 0.56	
		Obs. = 103	
		$\bar{R}^2 = 0.52$	

Table 1: Parametric and nonparametric estimates of convergence. Dependent variable: AV.PROD.GR.,*** indicates significance at 1%. For the smooth term the estimated degrees of freedom (EDF) are reported.

result). The implied rate of convergence is equal to 0.032, while the half-life for filling the gap between the productivity of the relatively poorer regions and the relatively richer ones is about 26.3 years.[6] However, the nonparametric estimate reported in the second column of Table 1 shows that the smooth term $s(\cdot)$ is highly significant and nonlinear, given the high value of EDF (Estimated Degrees of Freedom) equal to 7.9.[7]

Figure 1 reports the estimated relationship for parametric (dashed line) and nonparametric specification (solid line). While the parametric estimate suggests the convergence to a globally stable equilibrium around 1, the nonparametric estimation shows the existence of a strong nonlinear relationship, with the emergence of two possible stable equilibria around 0.9 and 1.05 (i.e. the points where the nonparametric estimate crosses the sample average from above).

III.B. Distribution Dynamics

A first information on the distribution dynamics of productivity is obtained by the dynamics of the variance of distribution, the so-

[6]The estimate rate of convergence is equal to $-log[1 + \hat{\beta}T]/T$, where T is the number of periods, while the half-life is equal to $\log(2)/\hat{\beta}$ (see Barro and Sala-i-Martin (2004)).

[7]All nonparametric regressions are estimated following Wood (2006), and implemented by the package *mgcv* in R. For more details see Fiaschi and Lavezzi (2007).

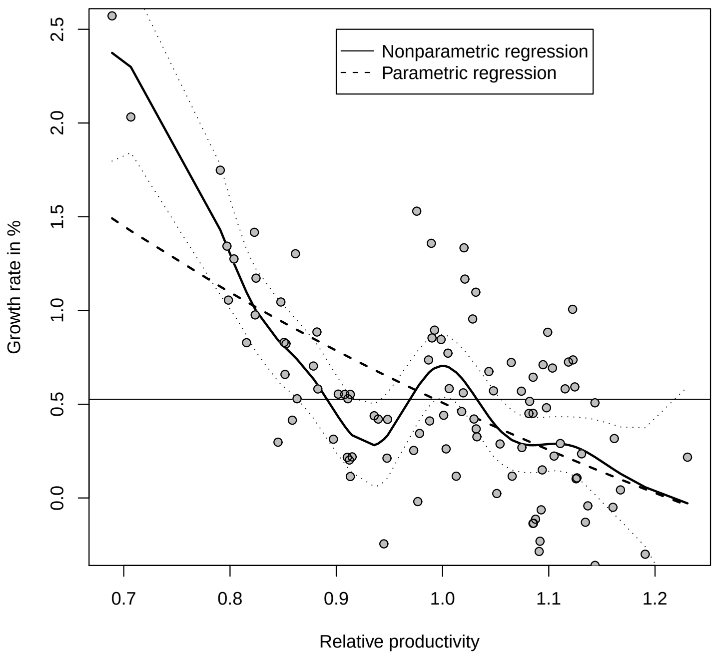

Figure 1: Parametric and nonparametric estimates of convergence across Italian Provinces' productivity. The horizontal line indicates the sample average of growth rate, while dotted lines are the confidence bands of the nonparametric estimation at 95%. Points represent the observed growth rates of provinces against their initial level of productivity.

called σ-convergence analysis (see Barro and Sala-i-Martin (2004)).

Figure 2 shows that the dispersion of (log of) productivity across the Italian Provinces has a clear downward trend.[8] This downward trend is also confirmed in the sub-periods 1995-2000 and 2001-2006, even if the decrease in dispersion is faster in the first subperiod than in the second (dotted lines in Figure 2 report the fitted values of the estimates in the two subperiods).[9]

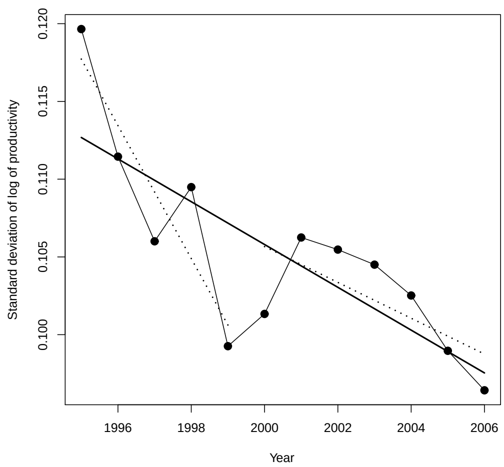

Figure 2: σ-convergence in productivity. Estimated trend for the whole period (solid line) and for sub-periods 1995-2001 and 2002-2006 (dotted lines)

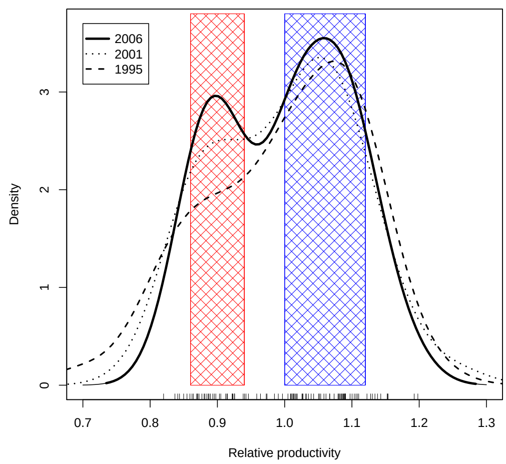

Figure 3: The estimated distribution of productivity in 1995, 2001 and 2006, and the identification of two clusters of provinces in 2006 according to the estimated peaks of distribution (see Section III.C.)

Figure 3 shows how in 1995 the distribution of productivity is unimodal with the mode around 1.06, although there already exists evidence of a cluster of low-productive regions around 0.9.[10] In 2006 the estimated density is drastically changed with the emergence of two peaks around 0.9 and 1.06, a polarization pattern already displayed by the distribution of productivity in 2001.[11]

[8]From the estimation of $\sigma_t = \alpha + \beta t + \varepsilon$ for the period 1995-2006 we obtain: $\hat{\sigma_t} = 2.86 - 0.0014t$. Both coefficients α and β are significant at 1% level.

[9]In particular, from the estimations of $\sigma_t = \alpha + \beta t + \varepsilon$ for the subperiods 1995-2000 and 2001-2006, we respectively obtain $\hat{\sigma_t} = 7.22 - 0.0036t$ and $\hat{\sigma_t} = 4.15 - 0.002t$ respectively. All coefficients are significant at 5% level.

[10]The estimate of density is made using package *sm* in R (see Bowman and Azzalini (1997)).

[11]The null hypothesis of unimodality cannot be rejected for both the distributions in 1995

Following Quah (1997), the analysis of the intra-distribution dynamics of productivity is based on the estimate of stochastic kernel by nonparametric methods.[12] Figure 4 reports the estimated stochastic kernel for Italian Provinces for the period 1995-2006 using a lag equal to 10. We also report a bold line representing the estimated median value of productivity at $t + 10$ conditional on its value at time t, and confidence bands of the estimated median value calculated by bootstrap procedure. The median line in Figure 4 crosses the bisector from below in two points, around 0.9 and 1.06, suggesting the emergence of two clusters of provinces; moreover, the median line is far below the bisector for values of productivity at time t lower than 0.9 and far above for values greater than 1.06. Therefore, on average, provinces whose relative productivity is out of the range $(0.9 - 1.06)$ are converging towards this range.

From the estimate of stochastic kernel we calculate the corresponding ergodic distribution following the procedure in Johnson (2005), adjusted for the use in the estimate of normalized variables (with respect to the average).[13] The ergodic distribution should indicate if the estimated distribution dynamics over the observed sample period has completely exhausted its effect on the distribution in the last year or, otherwise, significant distributional changes are expected in the future.[14]

and 2001, while it can be rejected at 5% level of significance for the distribution in 2006. Tests of multimodality use the bootstrap procedure proposed in Silverman (1986), p. 146, and are performed using 1000 bootstraps.

[12] Assuming that the process governing the evolution of the distribution $f(x)$ is time-invariant and Markovian, then:

$$f_{t+\tau}(z) = \int_0^\infty g_\tau(z|x) f_t(x) dx, \qquad (3)$$

where $g_\tau(z|x)$ is the τ-period ahead density of z conditional on x; $g_\tau(z|x)$ is also called stochastic kernel (see Durlauf *et al.* (2005) for more technical details). In the estimate we use a methodology known as *adaptive kernel* proposed by Silverman (1986), p. 100, with a Gaussian kernel.

[13] See Fiaschi and Romanelli (2009) for more details.

[14] Specifically, the ergodic distribution solves $f_\infty(z) = \int_0^\infty g_\tau(z|x) f_\infty(x) dx$.

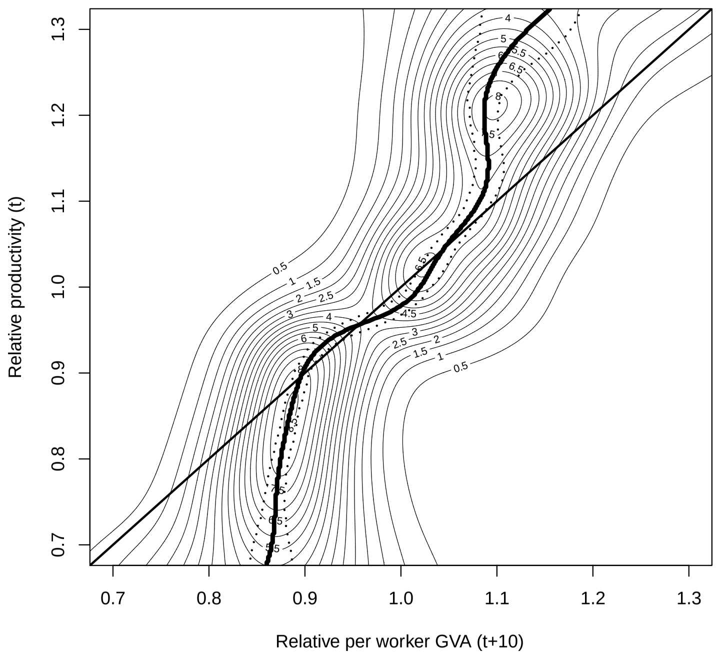

Figure 4: Contour plot of the estimated stochastic kernel of Italian Provinces in 1995-2006. Solid bold line represents the median value of productivity at time $t + 10$ conditioned at the level of productivity at time t, and solid line is the bisector. Dotted lines are the confidence bands for the estimate of the median value of productivity.

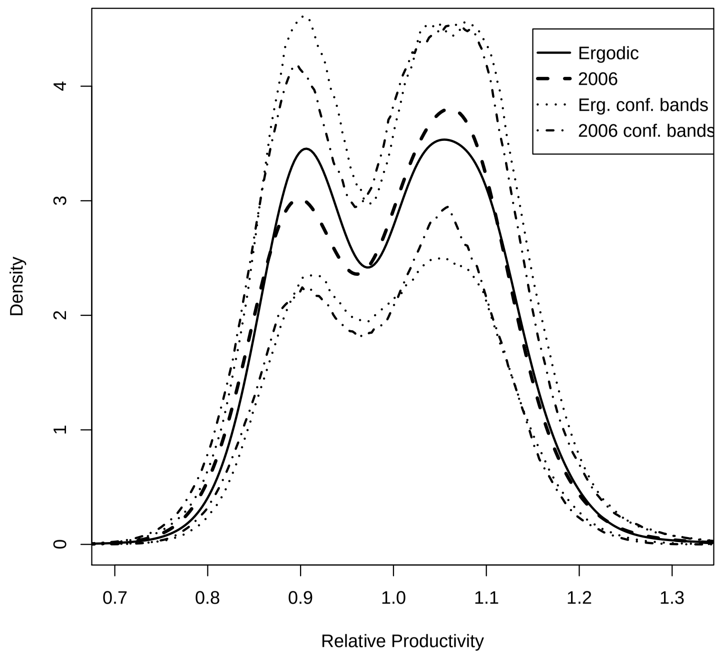

Figure 5: The estimated ergodic distribution of relative productivity (bold line), the estimated distribution of productivity in 2006 (dashed line), and their confidence bands (dotted lines).

Figure 5 shows that the estimated ergodic distribution displays two peaks and such distribution does not statistically differ from the one estimated in 2006; therefore, polarization should be a persistent phenomenon also in the long run.

III.C. The Characteristics of the Provinces in the Two Clusters

Below we investigate the characteristics of the two emerging clusters of provinces in 2006 resulting from the dynamics of polarization. In particular, Figure 5 indicates the centers of two clusters around 0.9 and 1.06. Accordingly, we define the provinces with productivity between 0.86 and 0.94 (i.e. within an interval of ± 0.04 around the center of the low-productive cluster) as belonging to the low-productive cluster (labelled Cluster L), while the provinces with productivity between 1 and 1.12 (i.e. within an interval of ± 0.06 around the center of the high-productive cluster) as belonging to the high-productive cluster (labelled Cluster H) (see Figure 3).[15]

Table 2(a) shows that 77% of provinces belongs to one of the two clusters in 2006 against 57% in 1995. Moreover, Table 2(b) reveals a high persistence into the two clusters, with the probability to remain in the same cluster from 1995 to 2006 respectively equal to 1 for Cluster L and 0.77 for Cluster H. The overall evidence therefore confirms that provinces are polarizing into two clusters of provinces.

(a)

	1995	2006
Cluster L (0.86 - 0.94)	16	31
Cluster H (1 - 1.12)	43	48
Total	59 (57%)	79 (77%)

(b)

	2006		
Cluster L (0.86 - 0.94)	1	0	1995
Cluster H (1 - 1.12)	0	0.77	

Table 2: (a) Number of provinces in Cluster L and Cluster H in 1995 and 2006, (b) Transition probabilities between Cluster L and Cluster H in the period 1995-2006

[15]The difference in the length of intervals between the two clusters reflects the different mass of the two peaks. We checked that small modifications of these intervals do not affect the results of the analysis.

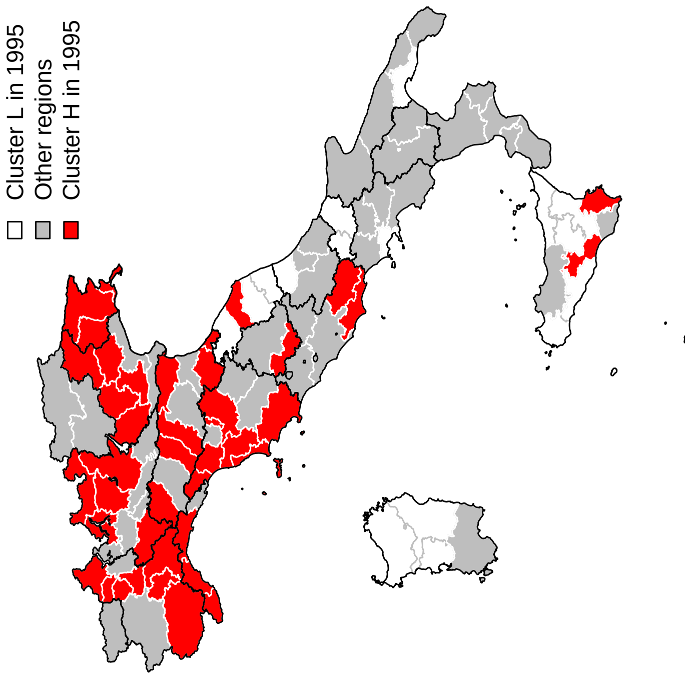

Figure 6: Productivity clusters in 1995

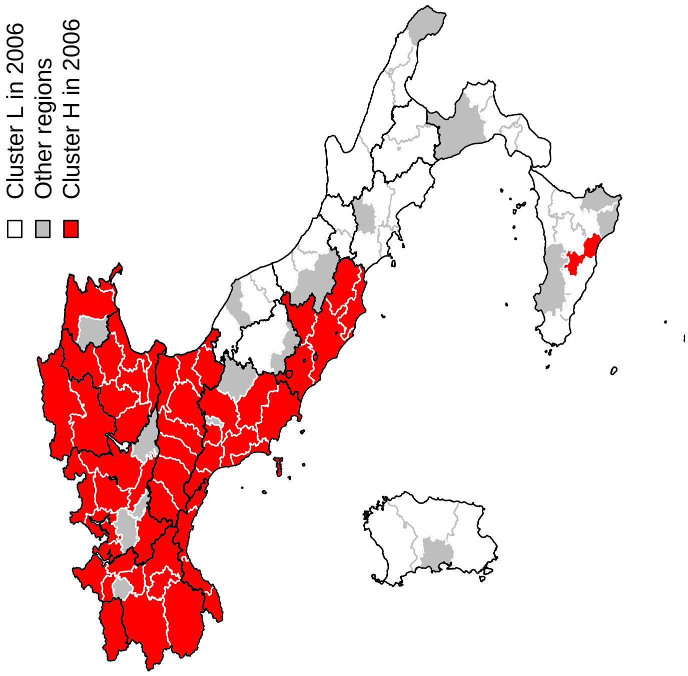

Figure 7: Productivity clusters in 2006

Figures 6 and 7 map the productivity clusters in 1995 and 2006.

In 1995 Cluster H includes some provinces in the North and Center-West (Tyrrhenian Coast in particular). Cluster L instead includes some provinces located in the Center-East (Adriatic Coast), South, Sicily and Sardinia. In 2006 Cluster L includes the most of provinces located in the Center-East (Adriatic Coast), South, Sicily and Sardinia, while Cluster H includes almost each province in the North and Center-West. Polarization therefore presents a clear geographical pattern (see Arbia and Basile (2005) for a similar result for per capita GDP).

Table 3 shows that provinces of Cluster H and Cluster L experience similar growth rates of productivity (PROD.GR), with a sharp decline of about 1% from 1995-1997 to 2004-2006.[16] The growth rate of employment (EMP.GR) is higher in 1995-1997 in Cluster L while the opposite holds in the 2004-2006, even though difference in 2004-2006 appears very small. The economic density (ECO.DEN) and population density (POP.DEN) are always higher in Cluster H.

SPATIAL.IND, an index of spatial autocorrelation,[17] points out that provinces of Cluster H are generally surrounded by high-productive provinces, instead provinces of Cluster L by low-productive provinces. This agrees with the spatial pattern of polarization reported in Figures 6 and 7.

Cluster H is more opened to trade: the export as share of GVA (EXP) and the import as share of GVA (IMP) are higher in both subperiods. It may be noticed that the slight decrease in EXP for Cluster H and the slight increase for Cluster L is contrasted by the sharp increase of IMP, notably in Cluster H. This suggests that openness could be a source of decrease in productivity in the considered period. The share of resident population 6 years old or older with tertiary education is about equal between the two

[16]In Table 3 for some variables we report time-average in order to remove the possible business cycle component.

[17]Spatial dependence is measured by the statistics G^* proposed by Ord and Getis (1995). The G^* statistics is computed by defining a set of neighbours for each province according to second-order-contiguity spatial matrix (see Anselin (1988)).

	Cluster L		Cluster H	
Variable	1995-97	2004-06	1995-97	2004-06
PROD.GR	0.013	**0.005**	**0.015**	0.004
EMP.GR	**0.012**	0.006	0.004	**0.009**
ECO.DEN	3.09	2.77	**4.14**	**4.13**
POP.DEN	5.03	5.01	**5.21**	**5.12**
SPATIAL.IND	-1.07	-1.71	**1.49**	**1.41**
EXP	0.11	0.12	**0.27**	**0.25**
IMP	0.07	0.13	**0.13**	**0.21**
LONGITUDE	**13.21**	**14.34**	10.88	10.78
LATITUDE	40.37	40.49	**44.35**	**44.49**
	2001		2001	
TERTIARY.EDU.2001	0.07	-	0.07	-
	1996	2001	1996	2001
FIRM.SIZE.1.on.POP. (x 1000)	35.75	40.79	**37.64**	**47.65**
FIRM.SIZE.10_15.on.POP (x 1000)	1.14	1.43	**2.12**	**2.11**
FIRM.SIZE.16_49.on.POP(x 1000)	1.04	1.2	**1.83**	**1.98**
FIRM.SIZE.50_250.on.POP(x 1000)	0.24	0.29	**0.44**	**0.53**
SHARE.FIRM.SIZE.1_9 (in %)	**96.15**	**95.81**	94.1	94.38
SHARE.FIRM.SIZE.10_49 (in %)	3.43	3.74	**5.24**	**4.91**
SHARE.FIRM.SIZE.50_249 (in %)	0.38	0.41	**0.59**	**0.65**
SHARE.FIRM.SIZE.250.and.more (in %)	0.03	0.03	**0.06**	**0.06**
	2001		2001	
SHARE.EMP.DISTRICTS	0.2	-	**0.31**	-
	1997		1997	
CREDIT.on.GVA	468.2	-	**587.22**	-
CREDIT.to.PRIVATE.FIRMS (x 1000)	37.74	-	**69.49**	-
	1999-2003		1999-2003	
EXTORTIONS.on.POP (x 1000)	**0.77**	-	0.51	-

Table 3: Characteristics of provinces in Cluster L and Cluster H. Bold characters indicate the highest values between the two clusters in the same year.

clusters.[18]

Italy presents a high level of entrepreneurship with respect to other developed countries, as measured by the number of firms per inhabitants (see ISTAT (2002)). However, many researchers argue that the huge number of firms, especially the firms with just one employee, i.e. self-employers, is not an index of entrepreneurship, but the joint result of i) a distortion of Italian fiscal system and labour market institutions which favour contractual intra-business relationship to hide what is otherwise a simple employer-employee relationship (i.e. many self-employers are actually workers); and ii) firms with one (or very few) employee, especially in the construction and service sectors, are the result of not a true entrepreneurial choice but of a lack of valuable alternative opportunities (see, e.g., Altieri and Oteri (2001) and Mandrone (2008)). The entrepreneurial fabric appears more developed in Cluster H. Indeed, the number of firms per (1000) inhabitants is always greater in Cluster H than in Cluster L in both periods independently of firm size. However, the difference between the two clusters is small for FIRM.SIZE.1.on.POP (firms with just one employee per (1000) inhabitants), and increasing (in relative terms) with the firm size; eventually, FIRM.SIZE.50_250.on.POP (the number of firms with 50-250 employees per (1000) inhabitants) is almost double in Cluster H.

The very high number of firms is also reflected in the size distribution of firms, i.e. the productive fabric, with a strong prevalence of very small firms. Some authors argue that firm size is inversely related to growth and competitiveness (see, e.g, Onida (2002) and ISTAT (2002)). Cluster L is indeed characterized by a higher share of firms with 1-9 employees on the total number of firms (SHARE.FIRM.SIZE.1_9), while Cluster H has a higher share of firms with 10-49, 50-249 and more than 250 employees (SHARE.FIRM.SIZE.10_49, SHARE.FIRM.SIZE.50_249 and

[18]Unfortunately, available data on education refer to resident population and not to labour force; moreover, there is not breakdown in different types of university degrees (e.g. scientific versus humanities degrees).

SHARE.FIRM.SIZE.250.and.more, respectively). As for entrepreneurial fabric, the difference between the two clusters is increasing with the firm size.

Cluster H appears more populated by industrial districts, measured by the share of employment in the industrial districts in 2001 (SHARE.EMP.DISTRICTS). This provides a further qualification on the kind of firms present in the two clusters. In 1997 the availability of credit is much higher in Cluster H, both in terms of amount of credit per unit of GVA (CREDIT.on.GVA) and in terms of amount of credit to private firms (in thousands of euros, CREDIT.to.PRIVATE.FIRMS).[19] Finally, Cluster L seems more afflicted by criminal activities hurting business as extortions, being the number of extortions per (1000) inhabitants (EXTORTIONS.on.POP) strongly higher.

The output composition and the employment share of the provinces in Cluster L and Cluster H reported in Table 4 highlights very significant differences.

Both in 1995 and 2006 the output of Cluster L is composed by a relative higher share in services (SERV), construction (CONSTR) and agriculture (AGRI), while Cluster H presents a higher share of industry (IND). This evidence also holds for the employment shares and with respect to sample average.[20]

IV. Growth Regressions of Italian Provinces

Taking as reference the Solow growth model augumented with human capital, Durlauf *et al.* (2005), pp. 577-579, show that around the steady-state the average annual growth rate of productivity of province i in the period 1995-2006, $AV.PROD.GR_i$, can be expressed as:

[19]The number of firms in 1997 is proxied by the number of firms in 1996.
[20]Suffix REL indicates that the value is relative to sample average.

	Output composition				Employment share			
	Cluster L		Cluster H		Cluster L		Cluster H	
Year	1995	2006	1995	2006	1995	2006	1995	2006
AGRI	**0.05**	**0.04**	0.04	0.03	**0.11**	**0.09**	0.05	0.04
AGRI.REL	**1.1**	**1.29**	0.77	0.83	**1.3**	**1.45**	0.63	0.74
IND	0.18	0.16	**0.28**	**0.22**	0.18	0.17	**0.28**	**0.22**
IND.REL	0.77	0.78	**1.21**	**1.07**	0.79	0.8	**1.22**	**1.05**
CONSTR	**0.07**	0.07	0.05	0.07	**0.08**	**0.09**	0.07	0.08
CONSTR.REL	**1.14**	**1.06**	0.89	0.99	**1.05**	**1.07**	0.93	0.99
SERV	**0.7**	**0.73**	0.63	0.69	**0.63**	**0.66**	0.6	0.66
SERV.REL	**1.06**	**1.04**	0.95	0.99	**1.03**	**1.01**	0.98	1.01

Table 4: Output composition and employment shares in Cluster L and Cluster H. Bold characters indicate the highest between the values of the two clusters related to the year.

$$\begin{aligned} AV.PROD.GR_i &= g + \beta_0 \log(PROD.REL.1995_i) + \\ &+ \beta_1 \log(AV.INV.RATE_i) + \\ &+ \beta_2 \log(\delta + g + AV.EMP.GR_i) + \\ &+ \beta_3 \log(TERTIARY.EDU.2001_i) + \\ &+ \beta_4 \log(A.1995_i) + \varepsilon_i, \end{aligned} \tag{4}$$

where g is the exogenous growth rate of technological progress, $AV.INV.RATE_i$ is the investment rate of province i and δ is the depreciation rate of capital; $A.1995_i$ should be interpreted as reflecting provincial-specific influences on productivity growth as technology, economic structure, institutions, etc., and ε_i a province-specific shock distributed independently of all the other variables. Eq. (4) is the baseline econometric model of all growth regressions presented below. Solow model predicts that β_0 and $\beta_2 < 0$, while β_1 and $\beta_3 > 0$.

Descriptive statistics on variables used in regressions can be found in Tables 6-9 in Appendix C. The log of PROD.REL.1995

(LOG.PROD.1995) is positively correlated with SPATIAL.IND.1995 ($\rho = 0.81$), LOG.ECO.DEN.1995 ($\rho = 0.31$), IND.1995 ($\rho = 0.56$), EXP.1995 ($\rho = 0.57$) and IMP.1995 ($\rho = 0.48$). Therefore, provinces with a high initial level of productivity are characterized by neighbouring provinces with high level of productivity, high share in industry, high trade openness, and high level of economic activity. Not surprisingly most of these provinces belong to Cluster H. As expected, IND.1995 is high correlated with SERV.1995 ($\rho = -0.90$), as well as with EXP.1995 ($\rho = 0.84$), SHARE.FIRM.SIZE.1_9.1996 ($\rho = -0.85$) and CREDITS.to.PRIVATE.FIRMS.1997 ($\rho = 0.69$). In turn, SHARE.FIRM.SIZE.1_9.1996 is highly correlated with all the other variables related to size distribution of firms SHARE.FIRM.SIZE.10_49.1996, SHARE.FIRM.SIZE.50_249.1996 (ρ is always above 0.80), except with SHARE.FIRM.SIZE.250_more.1996 ($\rho = -0.52$). Finally, FIRM.SIZE.10_15.on.POP, FIRM.SIZE.16_49.on.POP and FIRM.SIZE.50_250.on.POP are highly correlated between themselves, but not with FIRM.SIZE.1.on.POP. As we will discuss below, such high correlations crucially affect the selection of variables to be included in growth regressions.

Model (1) in Table 5 represents the baseline specification of Eq. (4). AV.INV.RATE is not included being not available at provincial level. According to Eq. (4) AV.EMP.GR is augmented by the rate of depreciation of capital δ and the growth rate of technological progress g. Given that we do not have any data at provincial level, we use the value of 0.05 proposed by Mankiw *et al.* (1992). The stock of human capital is proxied by TERTIARY.EDU.2001, being 2001 the last year for which we have data.

Model (1) includes as proxies for $A.1995_i$ the most of variables available at provincial level in 1995 or 1996, and others (see Appendix A for their complete list). In particular, in order to avoid perfect collinearity with the other sectoral shares AGR.1995 is excluded. Moreover, due to their high collinearity with other variables we do not include SHARE.FIRM.SIZE.10_49.1996, SHARE.FIRM.SIZE.50_249.1996, SHARE.FIRM.SIZE.250_more.1996, FIRM.SIZE.16_49.on.POP and FIRM.SIZE.50_250.on.POP and

Variable-Model	(1)	(2)	(3)
REGIONAL DUMMIES	YES	NO	YES
Intercept	-0.0501	-0.1243	0.0146*
LOG.PROD.1995	-0.0597***	-0.0507***	-0.0490***
LOG.AV.GR.EMP	-0.0046***	-0.0050**	-0.0057***
LOG.TERTIARY.EDU.2001	0.0006	0.0000	0.0025
LOG.ECO.DEN.1995	0.0000	0.0000	
SPATIAL.IND.1995	0.0001	0.0005	
CONSTR.1995	0.0363	0.0403	
SER.1995	0.0170	0.0172	
IND.1995	0.0258*	0.0106	
EXP.1995	-0.0060*	-0.0111***	-0.0040*
FIRMS.SIZE.1.on.POP.1996	-0.2537	-0.4640***	-0.2891***
FIRMS.SIZE.10_15.on.POP.1996	-0.5281	3.1948**	
SHARE.FIRM.SIZE.1_9.1996	0.0004	0.0006	
SHARE.FIRMS.SIZE.250_more.1996	0.0092	0.0084	
CREDITS.to.PRIVATE.FIRMS.1997	0.0195	0.0391*	
AV.EXTORSION.on.POP	0.0000	-0.0002	
SHARE.EMP.DISTRICTS.2001	-0.0014	-0.0007	
LONGITUDE	0.0002	0.0002	
LATITUDE	0.0001	0.0001	
INFRASTRUCTURES.1995		0.0001	
INTENSITY.PATENT.1995		0.0000	
LOG.AV.INV.RATE		-0.0160***	
AV.SHARE.IRREGULAR.WORKERS		0.0094	
AV.SERVICES.to.FIRMS		0.0006***	
$\bar{R}^2$	0.831	0.633	0.842
Breusch-Pagan test	33.71	42.55	25.92
	(0.624)	(0.008)	(0.357)
I Moran (2° contiguity)	0.601	2.057	1.201
	(0.545)	(0.040)	(0.230)
I Moran (all-linked)	0.387	5.622	-0.948
	(0.699)	(0.000)	(0.357)

Table 5: Growth regressions. Significance codes: 0.01"***" 0.05"**" 0.1"*". Estimation method: OLS. White robust standard errors for Model (2). P-values in parenthesis.

CREDIT.on.GVA.

In the regression we also include regional dummies (dummy for Lombardy is excluded being the region with the highest initial productivity) in order to catch the possible impact of common characteristics of provinces belonging to the same region, as the quality of institutions and regional policy (for which we don't have any information at provincial level). In this regard we remark that five regions out of twenty benefit of a broad autonomy, especially for their financial revenues and expeditures.[21] In order to control for possible geographical effects, we also include the longitude and the latitude of provinces (respectively LONGITUDE and LATITUDE).[22]

Model (2) includes all variables of Model (1) but regional dummies, and other variables available only at regional level (NUTS 2) (see Appendix A for their complete list). Finally, Model (3) represents our preferred specification and it is obtained starting from Model (1) and sequentially eliminating the least significant variable, in order to obtain the highest goodness of fit measured by the adjusted R^2, $\bar{R}^2$ (see Wooldridge (2003), pp. 192-196 for a discussion of model selection based on $\bar{R}^2$).

In Table 5 we also report two diagnostics on the the goodness of fit of estimates: the Breusch-Pagan test on the presence of heteroskedasticity (see Wooldridge (2003)), and the global Moran's I test on the presence of spatial dependence in the residuals (see Anselin (1988)). Distance beyond which spatial dependence is washing out is hardly detectable in practice, making somehow arbitrary the definition of the spatial weight matrix (on which Moran's I is based on). We consider two different definitions: i) a *second-order-contiguity* spatial weight matrix (the same used for the calculation of SPATIAL.IND);[23] and ii) a *all-linked* spatial weight matrix, i.e.

[21] This five regions keep between 60% (Friuli-Venezia Giulia) and 100% (Sicily) of all their levied taxes.

[22] We do not introduce an interaction term between longitude and latitude because of the high correlation among this interaction term and the longitude (0.96).

[23] In particular, the second-order-contiguity spatial weight matrix assumes that spatial effects come in from all contiguous provinces and from all provinces sharing a border with the contiguous provinces.

a matrix with a distance cut-off such that each province has at least one neighbour (in our case within a distance of 75 km).[24]

In the estimates of Models (1) and (3) we can reject the hypothesis of the presence of both heteroskedasticity and spatial dependence. On the contrary, in Model (2) without regional dummies such hypotheses cannot be rejected, suggesting that spatial dependence could be the result of the omission of a control for regional effects; accordingly, we report White-heteroskedasticity robust standard errors for Model (2).

The estimates provide evidence of conditional convergence across Italian Provinces, being the coefficient of the (log of) initial level of productivity, LOG.PROD.1995, always negative and statistically different from zero in all three models (see Table 5).[25] As is well-known, the presence of conditional convergence does not exclude the presence of club convergence (see, e.g., Durlauf *et al.* (2005)). Indeed, we argue that the detected convergence happened towards the two clusters (their masses significantly increased in the period, see Table 2) and within each cluster (the concentration around the two peaks increased from 1995 to 2006, see Figure 3). On the contrary, the two peaks seem not to have converged among them (again, see Figure 3). Section V.B. provides further discussion of this point.

Regional dummies appear to play a very crucial role, as witnessed by the increase in $\bar{R}^2$ from 0.633 in Model (2) to 0.842 in Model (3), although the inclusion in Model (2) of additional variables at regional level. Paci and Saba (1998) and Aiello and Scoppa (2007) present a similar finding.

The (log of) average growth rate of employment, LOG.AV.EMP.GR, has the expected negative impact; it results statistically significant in all three models with a fairly stable coefficient. Appendix D shows that such result is robust to potential endogeneity of the growth rate of employment.

The (log of) share of resident population with tertiary educa-

[24]Global Moran's I tests are implemented by the package *spdep* in R developed by Bivand (2009).

[25]The implied rate of convergence of the three models ranges from 7.1% to 9.7% per year.

tion in 2001, LOG.TERTIARY.EDU.2001, has the expected positive sign, but the coefficient is statistically significant only at 11% level.

Economic density (LOG.ECO.DEN.1995) and spatial spillovers (SPATIAL.IND.1995) are never statistically significant. The same applies for output composition, with the exception of the share of industry which is significant only in Model (1). This result partially contrasts with Paci and Pigliaru (1995), even if they consider the structural change at regional level. Therefore, we do not find evidence that some particular sectors, e.g. the industrial sector, are more conductive to productivity growth (at least sectors at very aggregate level).

The negative and statistically significant coefficient for the share of exports on GVA in 1995 (EXP.1995) in all three models means that openness to trade exerted a negative impact on productivity growth. A possible explanation of this finding is the return of Italy into the European Monetary System in 1996 and, later, its entrance into the Monetary Union; indeed, the related impossibility to use competitive devaluation to increase competitiveness, jointly with a some rigidity in factor reallocation, could have hurt the productivity growth of export-oriented provinces. A support to this explanation is the strong increase of the share of import both in Cluster H and Cluster L, and the decrease of export in Cluster H from 1995 to 2006 (see Table 3). Given this finding, the claim made by Onida (2002) that the high share of very small firms is one of the cause of the decreasing competitiveness of Italy in the recent year is to question; indeed, such fall appears to be an independent phenomenon (we also control for the distribution firm size, see below).

We find evidence that entrepreneurial fabric has a relevant impact on productivity growth. In particular, FIRM.SIZE.1.on.POP.1996, i.e. self-employers, has a negative and strongly significant coefficient in all three models. This supports the claim that high levels of self-employment are not signal of strong entrepreneurship but, instead, of distortions and misallocations in the labour market (see Altieri and Oteri (2001) and Mandrone (2008)).

The number of firms with 10-15 employees (FIRM.SIZE.10_15.on.POP.1996), a more suited proxy for entrepreneurship, has the expected positive and significant impact only in Model (2) without regional dummies, suggesting that entrepreneurship has a strong geographical pattern. Baldwin and Chowhan (2003) find a similar negative impact of self-employment on productivity growth for Canada.

In all models the variables related to firm size distribution are never statistically significant at the usual significance levels. This is in contrast with the findings of Forni and Paba (2000), but for the growth of per capita GVA, and of ISTAT (2002) and Onida (2002) which find a positive relationship between firm size and productivity growth, but directly considering the growth rate of productivity of firms (Pagano and Schivardi (2003) show that such relationship also holds in a cross-section of European countries).

The amount of credit to private firms (CREDITS.to.PRIVATE.FIRMS.1997) has a positive impact, but it is statistically significant only in model without regional dummies, again suggesting a geographical pattern in the amount of available credit to firms.

Neither extortions (AV.EXTORTION.on.POP), a proxy for a (negative) social environment for business, nor the employment in industrial districts (SHARE.EMP.DISTRICTS.2001) are statistically significant. It may be noticed that, since the exact definition of industrial districts is still under scrutiny (see, e.g., Becchetti and Rossi (2000)), the latter result is to be considered not conclusive.

Finally, neither latitude nor longitude turn to be significant even in the regression without regional dummies, suggesting the absence of spatial heterogeneity.[26]

[26] The geographical pattern of two clusters would suggest to consider also a term of interaction between longitude and latitude; however, this interaction term results highly correlated with longitude (about 0.96).

V. *The Determinants of Distribution Dynamics of Productivity*

In this section we discuss the contribution of variables included in Model (3) of Table 5 to the observed pattern of inequality and polarization. In particular, for each variable i) we map its impact on the growth rate of productivity at provincial level; ii) we calculate its counterfactual distribution of productivity in 2006, i.e. the distribution would have prevailed in 2006 if all provinces had had the same level of that variable, in particular equal to its sample average;[27] finally, iii) we report the relationship between its estimated impact and the initial level of productivity.

We report the analysis only for regional dummies, the initial level of productivity and the share of firms with more than 250 employees, being the variables in Model (3) with a significant distributional impact. For the sake of completeness the analysis of the growth rate of employment, export and self-employer are reported in Appendix E.

V.A. Regional Dummies

Figure 8 shows that the impact of regional dummies is very sizable ranging from -0.27% to 1.46% in term of annual growth rate. The unexplained regional component points out the existence of three main geographical clusters of provinces: i) provinces of Lom-

[27] More precisely, the counterfactual level of productivity in 2006 of province i with respect to variable Z^k, $PROD.2006.CF_i^k$, is given by:

$$PROD.2006.CF_i^k = PROD.1995_i \left(1 + AV.PROD.GR.CF_i^k\right)^{11}, \tag{5}$$

where $PROD.1995_i$ is the level of productivity in 1995 of provinces i and $AV.PROD.GR.CF_i^k$ is the counterfactual average growth rate of provinces i relative to variable k (11 is the total number of years in the sample). $AV.PROD.GR.CF_i^k$ is calculated in the following way:

$$AV.PROD.GR.CF_i^k = AV.\widehat{PROD}.GR_i - \hat{\gamma}^k \left(Z_i^k - \bar{Z}^k\right), \tag{6}$$

where $AV.\widehat{PROD}.GR_i$ is the fitted value of growth rate of productivity, and $\hat{\gamma}^k$ is the estimated coefficient for Z^k in Model (3) of Table 5; finally, $\bar{Z}^k$ is the sample average of Z^k. See Fiaschi et al. (2009) for more details on this methodology.

bardy, Veneto, Emilia-Romagna, Tuscany and Lazio, which had an impact in the range [0.80%; 1.46%]; ii) provinces of Aosta Valley, Piedmont, Trentino Alto-Adige, Friuli-Venezia-Giulia and Marche, which had an impact ranging from 0.3% to 0.8%; and, finally, iii) provinces of the remaining regions of the Center and of all the regions of the South of Italy, which had an impact between 0.3% and -0.27%. The overall picture therefore suggests the presence of a core of provinces, part in the North and part in the Center-West, benefiting of an advantage in terms of growth rate of productivity derived by economic, social and institutional factors with a crucial regional characterization. Taking into account more qualitative characteristics of regions - quality of governance, working of labour market, availability of credit, and economic performances - Nicolai and Benvenuti (2010) find a very similar picture of the best Italian regions in 2010.

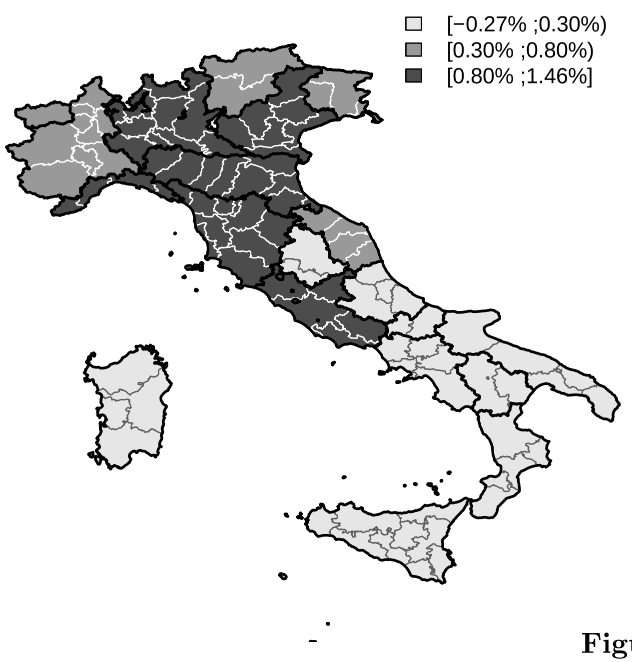

Figure 8: The estimated impact of regional dummies on the annual growth rate of productivity

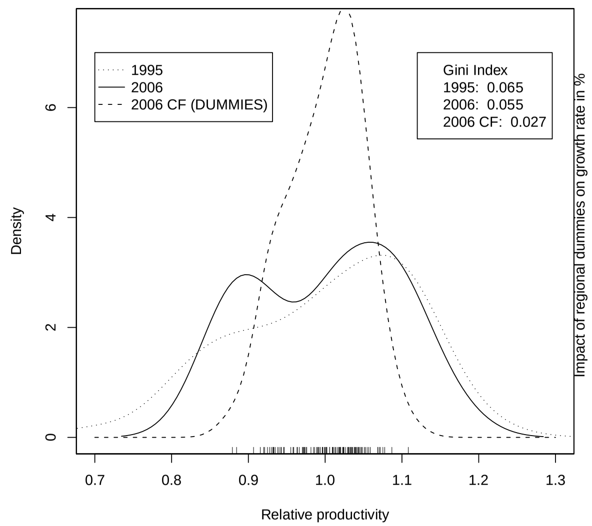

Figure 9: The initial, final and counterfactual final distributions of productivity

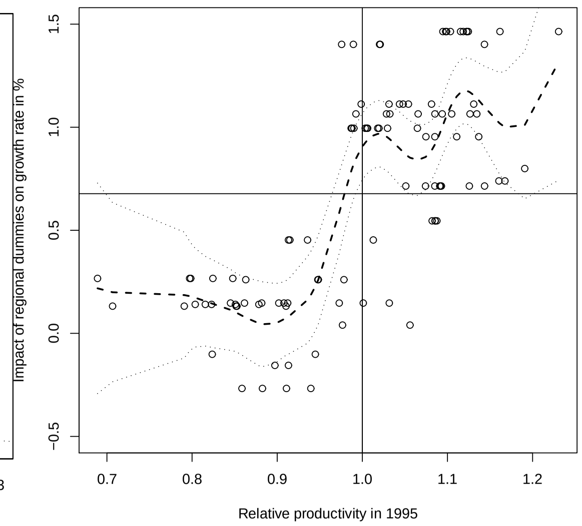

Figure 10: The relationship between the estimated impact of regional dummies on the average growth rate of productivity and the initial level of productivity

A comparison between Figures 7 and 8 highlights that this regional component mirrors the localization of Cluster L and Cluster H in 2006, with the exception of the peripheral regions of North (i.e. Aosta Valley, Friuli Venezia Giulia, Liguria, Piedmont and Trentino Alto-Adige). This finding should lead to reconsider the usual approach to the analysis of distribution dynamics across Italian regions, generally based on the common wisdom of a "dualistic" development between Northern and Southern regions of Italy (see, e.g., Paci and Pigliaru (1995), Cellini and Scorcu (1997), and Forni and Paba (2000)).

The counterfactual distribution of productivity in 2006 reported in Figure 9 shows the notable impact of the regional component on the distribution dynamics of productivity and, in particular, its strong positive impact on polarization and inequality. Indeed, we observe that the counterfactual distribution in 2006 is concentrated around 1, the twin peaks are completely disappeared, and the Gini index of counterfactual distribution is equal to 0.027 against 0.055 of the actual distribution in 2006. Figure 10 confirms the divergence impact of regional component, with the low-productive provinces in 1995 having on average a strong negative impact on their annual growth rates as opposed to an average non negative impact for high-productive provinces.

V.B. Initial Level of Productivity

Also the impact of the initial level of productivity looks very sizable across provinces, ranging from -1.02% to 1.83%. It also displays a clear geographical pattern, with the provinces in the South and in the Center having the higher impact (see Figure 11).

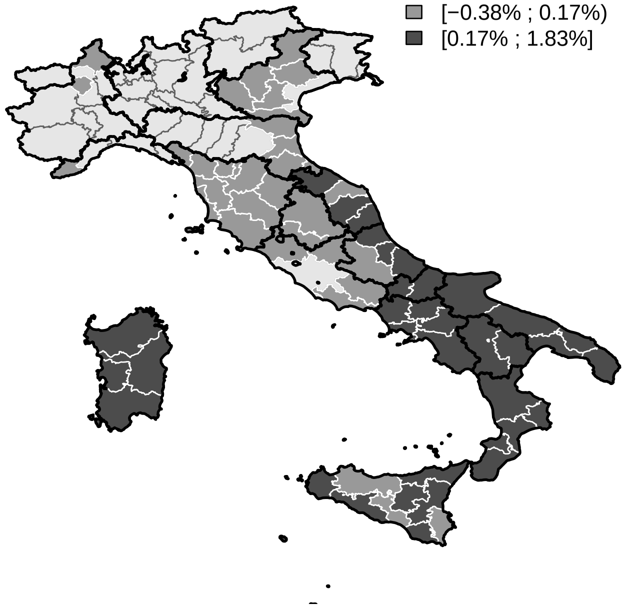

Figure 11: The estimated impact of the initial level of productivity on the annual growth rate of productivity

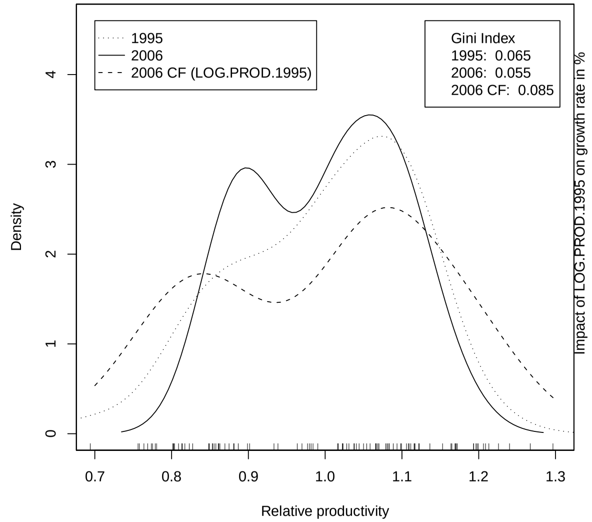

Figure 12: The initial, final and counterfactual final distributions of productivity

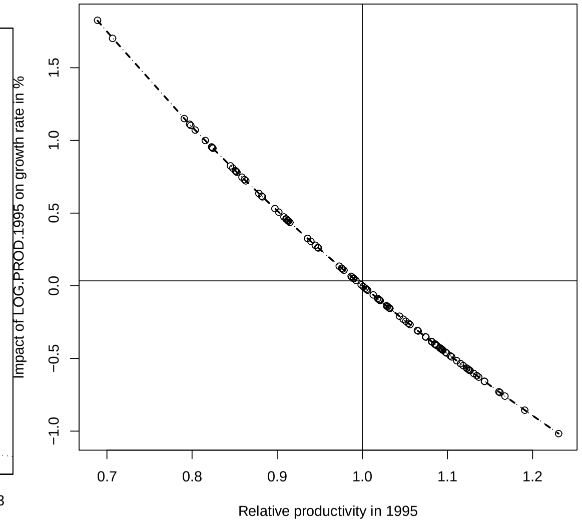

Figure 13: The relationship between the estimated impact of initial level of productivity on the average growth rate of productivity and the initial level of productivity

The comparison between the actual and the counterfactual distributions reported in Figure 12 highlights how the "catching-up" component represented by the initial level of productivity has the expected negative impact on the inequality of distribution, with the Gini index of the counterfactual distribution equal to 0.085 against a level of 0.055 of actual distribution. The two peaks of counterfactual distribution are about at the same distance of actual distribution, but the mass is notably less concentrated around the peaks. This evidence supports our previous claim that convergence happened towards the two clusters and within each cluster, but not between clusters. Therefore, the overall effect of the initial level of productivity is to increase polarization.

VI. Concluding Remarks

The analysis of distribution dynamics of Italian Provinces just presented leaves many open questions.

Firstly, the emergence of polarization in the distribution of productivity of Italian Provinces over the period 1995-2006 presents a clear geographical pattern, even though our findings partially challenge the standard view of the dualistic development of Italian regions (see, e.g., Paci and Pigliaru (1995)). Regional characteristics appear as the main determinants of this phenomenon, but they are left unexplained. Future research should aim at discovering the roots of this high cross-region heterogeneity.

Secondly, a large literature discussed the role of industrial district in the development of an economy (see, e.g., Becattini (2000)). In this regard our findings are mixed, not resulting statistically significant the share of employment in industrial districts and the composition of output, but detecting a significant impact of entrepreneurial fabrics. We reserve to future research a more detailed analysis of the effect of composition of output (e.g. considering sectoral data at 2-digit level) conditioned to firm size.

Finally, our analysis neglected the possible heterogeneity in the labour force. But, the differences in the type of human capital (e.g.

scientific versus humanities university degrees), and immigration across Italian Provinces could help to a better understanding of the observed dynamics of productivity.

Acknowledgements

Angela Parenti is grateful to Fondazione Cassa di Risparmio di Lucca for financial support under grants "Lo Sviluppo Economico delle Regioni Europee: Metodologie Empiriche, Politiche Regionali Europee e Dinamica Settoriale". A previous version of this paper was presented at International Conference in Honour of Salvatore Vinci in Naples. We thank the discussant, Antonio Acconcia, as well as the other participants to that meeting, and an anonymous referee for helpful comments.

Affiliation

Davide Fiaschi, University of Pisa
Lisa Gianmoena, IMT Lucca
Angela Parenti, University of Pisa

References

Altieri, G. and C. Oteri (2001). "Primo rapporto rapporto sul lavoro atipico in Italia: Analisi e tendenze," Roma: Ires, Working Paper, n. 5 (http://www.ires.it/node/187).

Aiello, F. and V. Scoppa (2007). "Convergence and Regional Productivities Divide in Italy: Evidence from Panel Data," mimeo.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*, Dordrecht, The Netherlands: Kluwer Academic Publishers.

Arbia, G. and R. Basile (2005). "Spatial Dependence and Nonlinearities in Regional Growth Behaviour in Italy," *Statistica*, 65(2), 145-167.

Baldwin, J. and J. Chowhan (2003). "The Impact of Self-Employment on Labour-Productivity Growth: A Canada and United States Comparison," *Economic Analysis Research Paper Series*, Statistics Canada.

Barro, R. J. and X. Sala-i-Martin (2004), *Economic Growth. Second Edition*, MIT Press.

Basile, R., A. Girardi and M. Mantuano (2010). "Interregional migration and unemployment dynamics: evidence from Italian provinces" mimeo.

Becattini, G. (2000), *Dal Distretto Industriale allo Sviluppo Locale, Svolgimento a Difesa di un'Idea*, Torino: Bollati Boringhieri.

Becchetti, L. and S. Rossi (2000). "The Positive Effect of Industrial Districts on the Export Performance of Italian Firms," *Review of Industrial Organization*, 16, 53-68.

Bivand, R. (2009). spdep: Spatial dependence: weighting schemes, statistics and models. R package version 0.4-36. http://CRAN.R-project.org/package=spdep

Bowman, A. W. and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis*, Oxford: Claredon Press.

Cahuc, P. and A. Zylbernberg (2004), *Labor Economics*, MIT Press, Cambridge.

Cellini, R. and A. Scorcu (1997). "How Many Italies? What Data Show about Growth and Convergence across Italian Regions, 1970-91," *Rassegna di Lavoro dell'ISCO*, 14, 93-124.

Di Liberto, A., F. Pigliaru and R. Mura (2008). "How to Measure the Unobservable: a Panel Tecnique for the Analysis of TFP Convergence," *Oxford Economic Papers*, 60, 343-368.

Durlauf, S.N., Johnson P. A. and J. R. W. Temple (2005), "Growth Econometrics", In Aghion, P. and S. N. Durlauf (Eds), *Handbook of Economic Growth*, 1(1), Amsterdam: Elsevier.

Fabiani, S. and G. Pellegrini (1997). "Education, Infrastructure, Geography and Growth: An Empirical Analysis of the Development of the Italian Provinces," Temi di discussione n. 323, Banca d'Italia.

Fiaschi D., A.M. Lavezzi and A. Parenti (2009). "Counterfactual Distribution Dynamics across European Regions," Discussion Papers, Dipartimento di Scienze Economiche, University of Pisa, n.85.

Fiaschi, D. and Lavezzi, A. M. (2007). "Productivity Polarization and Sectoral Dynamics in European Regions," *Journal of Macroeconomics*, 29, 612-637.

Fiaschi, D. and M. Romanelli (2009). "Nonlinear Dynamics in Welfare and the Evolution of World Inequality," Discussion Papers, Dipartimento di Scienze Economiche, University of Pisa, n.81.

Fingleton, B. and J. Le Gallo (2008). "Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: Finite sample properties," *Papers in Regional Science*, Blackwell Publishing, 87(3), 319-339.

Forni, M. and S. Paba (2000). “The Sources of Local Growth. Evidence from Italy,” *Giornale degli Economisti ed Annali di Economia*, 59 (1), 1-49.

“I Sistemi Locali del Lavoro,” (1997) ISTAT, Roma.

“Rapporto Annuale 2001,” ISTAT (2002), Roma.

Johnson, P. A. (2005). “A Continuous State Space Approach to Convergence by Parts,” *Economic Letters*, 86, 317-21.

Kennedy, P., (1992), *A Guide to Econometrics*, 3th edition, Blackwell, Oxford.

Magrini, S. (2007). “Analysing Convergence through the Distribution Dynamics Approach: Why and how?,” Working Papers 2007-13, University of Venice ”Ca’ Foscari”, Department of Economics.

Mandrone, E. (2008). “Quando la Flessibilitá Diviene Precarietá: una Stima Sezionale e Longitudinale,” collana Studi Isfol, n.6.

Mankiw, N. G., Romer, D. and D. Weil (1992). “A Contribution to the Empirics of Economic Growth,” *Quarterly Journal of Economics*, 107, 407-437.

Nicolai, D. and V. Benvenuti (2010). “Il Medagliere delle Regioni 2000-2010,” *Il Sole 24 Ore and Centro Studi Sintesi*, `http://www.ilsole24ore.com/pdf2010/SoleOnLine5/_Oggetti_Correlati/Documenti/Finanza\%20e\%20Mercati/2010/08/risultati-macroaree-telematiche.pdf?uuid=48e5fe0c-b3ff-11df-85f8-89f3d3b00320` .

Onida, F. (2002). “Growth, Competitiveness and Firm Size: Factors Shaping the Role of Italy’s Productive System in the World Arena,” *Review of Economic Conditions in Italy*, n. 3.

Ord, J. K. and A. Getis (1995). “Local Spatial Autocorrelation Statistics, Distributional Issue and an Application,” *Geographical Analysis*, 27, 286-305.

Paci, R. and F. Pigliaru (1995). "Differenziali di crescita tra le regioni italiane: un'analisi cross-section," *Rivista di politica economica* 85, 3-34.

Paci, R. and A. Saba (1998). "The Empirics of Regional Economic Growth in Italy. 1951-1993," *CRENOS Contributi di Ricerca* 97(1).

Pagano, P. and F. Schivardi (2003). "Firm Size Distribution and Growth," *The Scandinavian Journal of Economics*, 105(2), pp. 255-274.

Quah, D. T. (1997). "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs, *Journal of Economic Growth*, 2, 27-59.

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman & Hall: London.

Wood, S. N. (2006), *Generalized Additive Models. An Introduction with R*, London: Chapman and Hall.

Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT press.

Wooldridge, J. M. (2003), *Introductory Econometrics: A Modern Approach, 4^{th} Edition*, South-Western College Publishing.

Appendix

A List of Variables used in the Growth Regressions

Data used in the analysis come from the Italian official statistics provided by ISTAT (National Institute of Statistics) and Bank of Italy. In particualr, data on valued added, employment, population, territorial surface, import, export come from IPI PRINT web site (http://ipiprint.ipi.it/), while data on labour force come from Conti Economici Territoriali by ISTAT and data on industrial composition and industrial districts are from Censimento industria e servizi ISTAT 1996 and 2001 (http://www.istat.it). Data on tertiary education are from Censimento popolazione ISTAT 2001. The series of value added, import and export have been transformed at constant 1995 prices by using national value added deflator published by ISTAT. Series on the number of firms, its size distribution are taken from ISTAT, while data on credits from Bank of Italy. Codes and data are available at: `http://www.dse.ec.unipi.it/persone/docenti/fiaschi/`.

In the following we report the list of variables used in growth regressions at provincial level (NUTS 3):

1. AV.PROD.GR: average growth rate of productivity in the period 1995-2006;

2. LOG.PROD.1995: productivity in 1995 (log);

3. LOG.AV.GR.EMP: the average growth rate of employment in the period 1995-2006 (log);

4. LOG.TERTIARY.EDU.2001: the share of resident population six years old or older with tertiary education in 2001 (log);

5. LOG.ECO.DEN.1995: the density of economic activity in 1995;

6. SPATIAL.IND.1995: the spatial autocorrelation index of productivities in 1995, measured by the statistics G^* proposed by Ord and Getis (1995);

7. AGRI.1995, CONSTR.1995, SER.1995 and IND.1995: the shares of agriculture, constructions, services and industries in 1995 respectively;
8. EXP.1995 and IMP.1995: import and export as share of GVA in 1995 respectively;
9. SHARE.FIRM.SIZE.1_9.1996, SHARE.FIRM.SIZE.10_49.1996, SHARE.FIRM.SIZE.50_249.1996, SHARE.FIRM.SIZE.250.and.more.1996: the share of firms with a number of employees between 1 and 9, 10 and 49, 50 and 249, more than 250, on the total number of firms in 1996 respectively;
10. FIRM.SIZE.1.on.POP.1996, FIRM.SIZE.10_15.on.POP.1996, FIRM.SIZE.16_49.on.POP.1996 and FIRM.SIZE.50_250.on.POP.1996: the number of firms with only 1 employee, with a number of employees between 10 and 15, 16 and 49, 50 and 249, on total population in 1996 respectively;
11. CREDIT.to.PRIVATE.FIRMS.1997: the amount of credit per firms in 1997 (the number of firms is in 1996);
12. AV.EXTORTION.on.POP: the average number of extortions per inhabitants in the period 1999-2003;
13. SHARE.EMP.DISTRICTS.2001: the share of employment in industrial districts on total employment in 2001;
14. LONGITUDE and LATITUDE: the longitude and latitude coordinates of provinces centroids.

In the following we report the list of variables used in growth regressions at regional level (NUTS 2):

1. INFRASTRUCTURES.1995: an index of the endowment of infrastructures in 1995;
2. INTENSITY.PATENT.1995: an index of the intensity of innovation activity based on the number of patents in 1995;

3. LOG.AV.INV.RATE: the average investment rate in 1995-2006 (log);
4. AV.SHARE.IRREGULAR.WORKERS: the average share of irregular workers on total workers in 2001-2005;
5. AV.SERVICES.to.FIRMS: the average of an index of advanced services to firms in 2000-2006.

B List of Italian Provinces

ITC1	**Piedmont**	ITD1+ITD2	**Trentino Alto Adige**	ITE12	Lucca	ITF2	**Molise**	ITG14	Agrigento
ITC11	Turin	ITD10	Bolzano	ITE13	Pistoia	ITF21	Isernia	ITG15	Caltanissetta
ITC12	Vercelli	ITD20	Trent	ITE14	Florence	ITF22	Campobasso	ITG16	Enna
ITC13	Biella	ITD3	**Veneto**	ITE15	Prato	ITF3	**Campania**	ITG17	Catania
ITC14	Verbano-Cusio-Ossola	ITD31	Verona	ITE16	Leghorn	ITF31	Caserta	ITG18	Racusa
ITC15	Novara	ITD32	Vicenza	ITE17	Pisa	ITF32	Benevento	ITG19	Syracusa
ITC16	Cuneo	ITD33	Belluno	ITE18	Arezzo	ITF33	Naples	ITG2	**Sardinia**
ITC17	Asti	ITD34	Treviso	ITE19	Siena	ITF34	Avellino	ITG21	Sassari
ITC18	Alexandria	ITD35	Venice	ITE1A	Grosseto	ITF35	Salerno	ITG22	Nuoro
ITC2	**Aosta Valley**	ITD36	Padua	ITE2	**Umbria**	ITF4	**Apulia**	ITG23	Oristano
ITC29	Aosta Valley	ITD37	Rovigo	ITE21	Perugia	ITF41	Foggia	ITG24	Cagliari
ITC3	**Liguria**	ITD4	**Fiuli-Venezia-Giulia**	ITE22	Terni	ITF42	Bari		
ITC31	Imperia	ITD41	Pordenone	ITE3	**Marches**	ITF43	Taranto		
ITC32	Savona	ITD42	Udine	ITE31	Pesaro-Urbino	ITF44	Brindisi		
ITC33	Genoa	ITD43	Gorizia	ITE32	Ancona	ITF45	Lecce		
ITC34	La Spezia	ITD44	Trieste	ITE33	Macerata	ITF5	**Basilicata**		
ITC4	**Lombardy**	ITD5	**Emilia-Romagna**	ITE34	Ascoli-Piceno	ITF51	Potenza		
ITC41	Varese	ITD51	Piacenza	ITE4	**Lazio**	ITF52	Matera		
ITC42	Como	ITD52	Parma	ITE41	Viterbo	ITF6	**Calabria**		
ITC43	Lecco	ITD53	Reggio Emilia	ITE42	Rieti	ITF61	Cosenza		
ITC44	Sondrio	ITD54	Modena	ITE43	Rome	ITF62	Crotone		
ITC45	Milan	ITD55	Bologna	ITE44	Latina	ITF63	Catanzaro		
ITC46	Bergamo	ITD56	Ferrara	ITE45	Frosinone	ITF64	Vibo Valenzia		
ITC47	Brescia	ITD57	Ravenna	ITF1	**Abruzzo**	ITF65	Reggio Calabria		
ITC48	Pavia	ITD58	Forlì-Cesena	ITF11	Aquila	ITG1	**Sicily**		
ITC49	Lodi	ITD59	Rimini	ITF12	Teramo	ITG11	Trapani		
ITC4A	Cremona	ITE1	**Tuscany**	ITF13	Pescara	ITG12	Palermo		
ITC4B	Mantova	ITE11	Massa-Carrara	ITF14	Chieti	ITG13	Messina		

C Descriptive Statistics

	AV.GR.PROD	LOG.PROD.1995	LOG.AV.GR.EMP
Mean	0.01	-0.01	-2.86
S.d.	0.01	0.12	0.22
	LOG.ECO.DEN.1995	SPATIAL.IND.1995	AGRI.1995
Mean	3.72	0.33	0.05
S.d.	5.42	2.10	0.03
	CONSTR.1995	SER.1995	IND.1995
Mean	0.06	0.66	0.23
S.d.	0.02	0.08	0.09
	EXP.1995	IMP.1995	SHARE.FIRM.SIZE.1_9.1996
Mean	0.19	0.15	94.96
S.d.	0.14	0.12	1.82
	SHARE.FIRM.SIZE.10_49.1996	SHARE.FIRM.SIZE.50_249.1996	SHARE.FIRMS.SIZE.250_more.1996
Mean	4.49	0.50	0.05
S.d.	1.62	0.21	0.03
	FIRMS.SIZE.1.on.POP.1996	FIRMS.SIZE.10_15.on.POP.1996	FIRMS.SIZE.16_49.on.POP.1996
Mean	0.04	0.0017	0.0015
S.d.	0.004	0.0008	0.0007
	FIRMS.SIZE.50_250.on.POP.1996	CREDITS.to.PRIVATE.FIRMS.1997	CREDITS.on.GVA
Mean	0.0003	0.06	556.60
S.d.	0.0002	0.03	173.17
	AV.EXTORSION.on.POP	SHARE.EMP.DISTRICTS.2001	LOG.TERTIARY.EDU.2001
Mean	0.63	0.22	-2.71
S.d.	0.35	0.29	0.17
	LONGITUDE	LATITUDE	
Mean	12.09	42.93	
S.d.	2.65	2.68	

Table 6: Descriptive statistics for the variables used in regressions

	AV.PROD.GR	LOG.PROD.1995	LOG.AV.GR.EMP	LOG.ECO.DEN.1995	SPATIAL.IND.1995
AV.GR.PROD	1.00	-0.61	-0.40	-0.13	-0.39
LOG.PROD.1995	-0.61	1.00	0.37	0.31	0.81
LOG.AV.GR.EMP	-0.40	0.37	1.00	0.16	0.22
LOG.ECO.DEN.1995	-0.13	0.31	0.16	1.00	0.19
SPATIAL.IND.1995	-0.39	0.81	0.22	0.19	1.00
AGRI.1995	0.32	-0.51	-0.18	-0.40	-0.37
CONSTR.1995	0.30	-0.40	-0.34	-0.41	-0.32
SER.1995	0.16	-0.37	-0.17	0.04	-0.48
IND.1995	-0.30	0.56	0.28	0.18	0.60
EXP.1995	-0.40	0.57	0.32	0.22	0.57
IMP.1995	-0.29	0.48	0.17	0.32	0.40
SHARE.FIRM.SIZE.1_9.1996	0.35	-0.67	-0.39	-0.25	-0.71
SHARE.FIRM.SIZE.10_49.1996	-0.34	0.66	0.39	0.23	0.70
SHARE.FIRM.SIZE.50_249.1996	-0.40	0.67	0.32	0.31	0.66
SHARE.FIRMS.SIZE.250_more.1996	-0.30	0.62	0.21	0.49	0.45
FIRMS.SIZE.1.on.POP.1996	-0.24	0.24	0.15	0.22	0.25
FIRMS.SIZE.10_15.on.POP.1996	-0.38	0.67	0.38	0.24	0.71
FIRMS.SIZE.16_49.on.POP.1996	-0.34	0.64	0.37	0.25	0.71
FIRMS.SIZE.50_250.on.POP.1996	-0.43	0.70	0.34	0.31	0.71
CREDITS.to.PRIVATE.FIRMS.1997	-0.36	0.70	0.35	0.47	0.58
CREDITS.on.GVA	-0.36	0.57	0.38	0.51	0.39
AV.EXTORSION.on.POP	0.32	-0.58	-0.27	-0.11	-0.52
SHARE.EMP.DISTRICTS.2001	-0.10	0.26	0.22	0.06	0.43
LOG.TERTIARY.EDU.2001	-0.16	0.18	0.23	0.39	-0.01
LONGITUDE	0.43	-0.736	-0.02	-0.14	-0.81
LATITUDE	-0.37	0.76	0.29	0.23	0.79

	AGRI.1995	CONSTR.1995	SER.1995	IND.1995	EXP.1995
AV.GR.PROD	0.32	0.3	0.16	-0.3	-0.4
LOG.PROD.1995	-0.51	-0.4	-0.37	0.56	0.57
LOG.AV.GR.EMP	-0.18	-0.34	-0.17	0.28	0.32
LOG.ECO.DEN.1995	-0.4	-0.41	0.04	0.18	0.22
SPATIAL.IND.1995	-0.37	-0.32	-0.48	0.6	0.57
AGRI.1995	1	0.39	0.09	-0.47	-0.4
CONSTR.1995	0.39	1	0.03	-0.36	-0.35
SER.1995	0.09	0.03	1	-0.9	-0.74
IND.1995	-0.47	-0.36	-0.9	1	0.84
EXP.1995	-0.4	-0.35	-0.74	0.84	1
IMP.1995	-0.21	-0.29	-0.38	0.46	0.63
SHARE.FIRM.SIZE.1_9.1996	0.49	0.36	0.72	-0.85	-0.82
SHARE.FIRM.SIZE.10_49.1996	-0.47	-0.35	-0.7	0.83	0.8
SHARE.FIRM.SIZE.50_249.1996	-0.49	-0.33	-0.72	0.85	0.78
SHARE.FIRMS.SIZE.250_more.1996	-0.47	-0.47	-0.34	0.54	0.51
FIRMS.SIZE.1.on.POP.1996	-0.3	-0.34	-0.07	0.23	0.23
FIRMS.SIZE.10_15.on.POP.1996	-0.51	-0.38	-0.59	0.74	0.75
FIRMS.SIZE.16_49.on.POP.1996	-0.48	-0.37	-0.7	0.83	0.81
FIRMS.SIZE.50_250.on.POP.1996	-0.53	-0.38	-0.71	0.86	0.8
CREDITS.to.PRIVATE.FIRMS.1997	-0.46	-0.4	-0.54	0.69	0.68
CREDITS.on.GVA	-0.45	-0.45	-0.19	0.4	0.43
AV.EXTORSION.on.POP	0.32	0.17	0.34	-0.43	-0.42
SHARE.EMP.DISTRICTS.2001	-0.27	-0.19	-0.71	0.74	0.63
LOG.TERTIARY.EDU.2001	-0.36	-0.42	0.36	-0.11	0.00
LONGITUDE	0.34	0.17	0.34	-0.44	-0.43
LATITUDE	-0.48	-0.35	-0.53	0.68	0.67

Table 7: Correlation between the variables used in regressions

	IMP.1995	SHARE.FIRM.SIZE.1_9.1996	SHARE.FIRM.SIZE.10_49.1996
AV.GR.PROD	-0.29	0.35	-0.34
LOG.PROD.1995	0.48	-0.67	0.66
LOG.AV.GR.EMP	0.17	-0.39	0.39
LOG.ECO.DEN.1995	0.32	-0.25	0.23
SPATIAL.IND.1995	0.40	-0.71	0.70
AGRI.1995	-0.21	0.49	-0.47
CONSTR.1995	-0.29	0.36	-0.35
SER.1995	-0.38	0.72	-0.70
IND.1995	0.46	-0.85	0.83
EXP.1995	0.63	-0.82	0.80
IMP.1995	1	-0.41	0.39
SHARE.FIRM.SIZE.1_9.1996	-0.41	1	-1
SHARE.FIRM.SIZE.10_49.1996	0.39	-1	1
SHARE.FIRM.SIZE.50_249.1996	0.48	-0.88	0.84
SHARE.FIRMS.SIZE.250_more.1996	0.49	-0.52	0.48
FIRMS.SIZE.1.on.POP.1996	0.02	-0.21	0.22
FIRMS.SIZE.10_15.on.POP.1996	0.32	-0.92	0.94
FIRMS.SIZE.16_49.on.POP.1996	0.37	-0.96	0.97
FIRMS.SIZE.50_250.on.POP.1996	0.45	-0.91	0.88
CREDITS.to.PRIVATE.FIRMS.1997	0.54	-0.75	0.72
CREDITS.on.GVA	0.34	-0.53	0.52
AV.EXTORSION.on.POP	-0.22	0.54	-0.54
SHARE.EMP.DISTRICTS.2001	0.19	-0.71	0.71
LOG.TERTIARY.EDU.2001	0.08	-0.04	0.03
LONGITUDE	-0.35	0.47	-0.46
LATITUDE	0.31	-0.83	0.83

	SHARE.FIRM.SIZE.50_249.1996	SHARE.FIRM.SIZE.250.and.more.1996	FIRMS.SIZE.1.on.POP.1996
AV.GR.PROD	-0.40	-0.30	-0.24
LOG.PROD.1995	0.67	0.62	0.24
LOG.AV.GR.EMP	0.32	0.21	0.15
LOG.ECO.DEN.1995	0.31	0.49	0.22
SPATIAL.IND.1995	0.66	0.45	0.25
AGRI.1995	-0.49	-0.47	-0.30
CONSTR.1995	-0.33	-0.47	-0.34
SER.1995	-0.72	-0.34	-0.07
IND.1995	0.85	0.54	0.23
EXP.1995	0.78	0.51	0.23
IMP.1995	0.48	0.49	0.02
SHARE.FIRM.SIZE.1_9.1996	-0.88	-0.52	-0.21
SHARE.FIRM.SIZE.10_49.1996	0.84	0.48	0.22
SHARE.FIRM.SIZE.50_249.1996	1	0.65	0.12
SHARE.FIRMS.SIZE.250_more.1996	0.65	1	0.03
FIRMS.SIZE.1.on.POP.1996	0.12	0.03	1
FIRMS.SIZE.10_15.on.POP.1996	0.71	0.37	0.47
FIRMS.SIZE.16_49.on.POP.1996	0.81	0.44	0.42
FIRMS.SIZE.50_250.on.POP.1996	0.96	0.61	0.34
CREDITS.to.PRIVATE.FIRMS.1997	0.77	0.69	0.28
CREDITS.on.GVA	0.51	0.51	0.42
AV.EXTORSION.on.POP	-0.46	-0.33	-0.16
SHARE.EMP.DISTRICTS.2001	0.65	0.17	0.29
LOG.TERTIARY.EDU.2001	0.08	0.32	0.36
LONGITUDE	-0.47	-0.36	-0.34
LATITUDE	0.74	0.52	0.33

Table 8: Correlation between the variables used in regressions

	FIRMS.SIZE.10_15.on.POP.1996	FIRMS.SIZE.16_49.on.POP.1996	FIRMS.SIZE.50_250.on.POP.1996
AV.GR.PROD	-0.38	-0.34	-0.43
LOG.PROD.1995	0.67	0.64	0.70
LOG.AV.GR.EMP	0.38	0.37	0.34
LOG.ECO.DEN.1995	0.24	0.25	0.31
SPATIAL.IND.1995	0.71	0.71	0.71
AGRI.1995	-0.51	-0.48	-0.53
CONSTR.1995	-0.38	-0.37	-0.38
SER.1995	-0.59	-0.70	-0.71
IND.1995	0.74	0.83	0.86
EXP.1995	0.75	0.81	0.80
IMP.1995	0.32	0.37	0.45
SHARE.FIRM.SIZE.1_9.1996	-0.92	-0.96	-0.91
SHARE.FIRM.SIZE.10_49.1996	0.94	0.97	0.88
SHARE.FIRM.SIZE.50_249.1996	0.71	0.81	0.96
SHARE.FIRMS.SIZE.250_more.1996	0.37	0.44	0.61
FIRMS.SIZE.1.on.POP.1996	0.47	0.42	0.34
FIRMS.SIZE.10_15.on.POP.1996	1	0.96	0.83
FIRMS.SIZE.16_49.on.POP.1996	0.96	1	0.90
FIRMS.SIZE.50_250.on.POP.1996	0.83	0.90	1
CREDITS.to.PRIVATE.FIRMS.1997	0.67	0.72	0.79
CREDITS.on.GVA	0.55	0.56	0.57
AV.EXTORSION.on.POP	-0.53	-0.53	-0.50
SHARE.EMP.DISTRICTS.2001	0.67	0.75	0.69
LOG.TERTIARY.EDU.2001	0.09	0.06	0.12
LONGITUDE	-0.53	-0.50	-0.55
LATITUDE	0.82	0.82	0.80

	CREDITS.to.PRIVATE.FIRMS.1997	CREDITS.on.GVA	AV.EXTORSION.on.POP
AV.GR.PROD	-0.36	-0.36	0.32
LOG.PROD.1995	0.70	0.57	-0.58
LOG.AV.GR.EMP	0.35	0.38	-0.27
LOG.ECO.DEN.1995	0.47	0.51	-0.11
SPATIAL.IND.1995	0.58	0.39	-0.52
AGRI.1995	-0.46	-0.45	0.32
CONSTR.1995	-0.40	-0.45	0.17
SER.1995	-0.54	-0.19	0.34
IND.1995	0.69	0.40	-0.43
EXP.1995	0.68	0.43	-0.42
IMP.1995	0.54	0.34	-0.22
SHARE.FIRM.SIZE.1_9.1996	-0.75	-0.53	0.54
SHARE.FIRM.SIZE.10_49.1996	0.72	0.52	-0.54
SHARE.FIRM.SIZE.50_249.1996	0.77	0.51	-0.46
SHARE.FIRMS.SIZE.250_more.1996	0.69	0.51	-0.33
FIRMS.SIZE.1.on.POP.1996	0.28	0.42	-0.16
FIRMS.SIZE.10_15.on.POP.1996	0.67	0.55	-0.53
FIRMS.SIZE.16_49.on.POP.1996	0.72	0.56	-0.53
FIRMS.SIZE.50_250.on.POP.1996	0.79	0.57	-0.50
CREDITS.to.PRIVATE.FIRMS.1997	1	0.84	-0.45
CREDITS.on.GVA	0.84	1	-0.35
AV.EXTORSION.on.POP	-0.45	-0.35	1
SHARE.EMP.DISTRICTS.2001	0.45	0.30	-0.31
LOG.TERTIARY.EDU.2001	0.32	0.46	-0.03
LONGITUDE	-0.46	-0.35	0.52
LATITUDE	0.61	0.45	-0.60

Table 9: Correlation between the variables used in regressions

	SHARE.EMP.DISTRICTS.2001	LOG.TERTIARY.EDU.2001	LONGITUDE	LATITUDE
AV.GR.PROD	-0.10	-0.18	0.43	-0.37
LOG.PROD.1995	0.26	0.21	-0.73	0.76
LOG.AV.GR.EMP	0.22	0.23	-0.02	0.29
LOG.ECO.DEN.1995	0.06	0.42	-0.14	0.23
SPATIAL.IND.1995	0.43	0.01	-0.81	0.79
AGRI.1995	-0.27	-0.37	0.34	-0.48
CONSTR.1995	-0.19	-0.44	0.17	-0.35
SER.1995	-0.71	0.37	0.34	-0.53
IND.1995	0.74	-0.11	-0.44	0.68
EXP.1995	0.63	-0.01	-0.43	0.67
IMP.1995	0.19	0.10	-0.35	0.31
SHARE.FIRM.SIZE.1_9.1996	-0.71	-0.05	0.47	-0.83
SHARE.FIRM.SIZE.10_49.1996	0.71	0.04	-0.46	0.83
SHARE.FIRM.SIZE.50_249.1996	0.65	0.10	-0.47	0.74
SHARE.FIRMS.SIZE.250_more.1996	0.17	0.36	-0.36	0.52
FIRMS.SIZE.1.on.POP.1996	0.29	0.36	-0.34	0.33
FIRMS.SIZE.10_15.on.POP.1996	0.67	0.10	-0.53	0.82
FIRMS.SIZE.16_49.on.POP.1996	0.75	0.07	-0.50	0.82
FIRMS.SIZE.50_250.on.POP.1996	0.69	0.13	-0.55	0.80
CREDITS.to.PRIVATE.FIRMS.1997	0.45	0.36	-0.46	0.61
CREDITS.on.GVA	0.30	0.52	-0.35	0.45
AV.EXTORSION.on.POP	-0.31	-0.05	0.52	-0.60
SHARE.EMP.DISTRICTS.2001	1	-0.21	-0.28	0.48
LOG.TERTIARY.EDU.2001	-0.21	1.00	0.05	0.08
LONGITUDE	-0.28	0.00	1	-0.64
LATITUDE	0.48	0.09	-0.64	1

Table 10: Correlation between the variables used in regressions

D Test of Endogeneity of the Growth Rate of Employment

A large literature suggests that growth rate of productivity affects the growth rate of employment (see, e.g., Cahuc and Zylbernberg (2004), Cap. 10). In this case the growth rate of employment LOG.AV.GR.EMP is potentially endogenous in our regressions.

The theoretical framework based on the standard Solow model (with the assumption of decreasing marginal returns to labour) implies that there exists a negative relationship between the growth of employment and productivity. In our estimates we indeed find this negative sign (see Table 5). However there exists a potential reverse casualty of the growth of productivity on the growth of employment but of positive sign. Indeed, a positive technology shock increases labour productivity but also, shifting upward the curve of labour demand, the equilibrium level of employment. Moreover, if wages follow productivity, provinces with higher growth rates of productivity could also show higher growth rates of employment thanks to immigration (the opposite holds for provinces with low growth rates of productivity).

The endogeneity of LOG.AV.GR.EMP is tested by the Durbin-Wu-Hausman test in its regression-based form, using as instruments all the exogenous explanatory variables of the model and some additional instruments.[28] We defined three different instruments for LOG.AV.GR.EMP. More precisely: i) INSTR.3G, derived by the three-group method described in Kennedy (1992), in which the instrumental variable takes values -1, 0 or 1 if the potentially endogenous variable is respectively in the top, middle or bottom third of its ranking. This type of instrument is usually utilized when variables are subject to measurement error; ii) INSTR.POP.FEM.ACTIVE, the share of female on the total active population in 1995, on the hypothesis that the demographic composition of labour force can affect the growth of employment; and, finally iii) INSTR.UNEM.RATE.1995, the unemployment rate in

[28] For more details see Wooldridge (2002), pp. 118-122.

1995, on the hypothesis that high level of unemployment favours the growth rate of employment.[29]

Table 11 reports the results of first-stage and second-stage regressions of Durbin-Wu-Hausman.

Dependent Variable	First-Stage Estimation LOG.AV.GR.EMP	Second-Stage Estimation AV.PROD.GR
REGIONAL DUMMIES	YES	YES
Intercept	-2.9346***	0.0070
LOG.PROD.1995	0.0611	-0.0461***
LOG.TERTIARY.EDU.2001	0.0321	0.0015
EXP.1995	0.1301	-0.0039*
FIRM.SIZE.1.on.POP.1996	-0.1832	-0.2830***
INSTR.3G	0.1442***	
INSTR.UNEM.RATE.1995	0.0122**	
INSTR.POP.FEM.ACTIVE	0.0024	
LOG.AV.GR.EMP		-0.0082***
LOG.AV.GR.EMP_RES		0.0042
	$\bar{R}^2$=0.589	$\bar{R}^2$=0.844
t-Test		H_0: LOG.AV.GR.EMP_RES=0 t=1.55, Pr(>t)= 0.12

Table 11: Endogeneity test for the growth rate of employment. Significance codes: 0.01"***" 0.05"**" 0.1"*".

The first-stage regression for LOG.AV.GR.EMP shows that the coefficients of INSTR.3G and INSTR.UNEM.RATE.1995 are both positive and statistically significant, while the other instrument SHARE.ACTIVE.POP.FEMALE is not statistically significant. Table 11 reports that the null hypothesis that LOG.AV.GR.EMP_RES (the residuals of the first-stage regression) is equal to zero cannot be rejected at usual levels of significance (i.e. with a p-value of 0.12). We therefore conclude that LOG.AV.GR.EMP is exogenous and the estimates are not biased.

[29]Endogeneity tests assume that all instruments used in the first-stage regressions are valid, i.e. not correlated with error term. However, this cannot be the case for the type of instrument like INSTR.3G as discussed by Fingleton and Le Gallo (2008). The Sargan test of overidentifying restrictions allows to check the hypothesis of validity of all instruments (for more details see Wooldridge (2002), pp. 122-124). The resulting statistics of the Sargan test is equal to 2.98 against a critical value of 56.92. We then conclude that all the instruments are valid.

E The Other Determinants of Distribution Dynamics of Productivity

EA. The Growth Rate of Employment

The geographical pattern is less clear with respect to regional dummies and initial level of productivity, but the overall picture suggests that the growth rate of employment mostly hurt the provinces of North and Center. On average its impact is small with respect to regional dummies and initial level of productivity, even though several provinces have an impact lower than -0.13% in terms of annual growth rate of productivity (see Figure 14). The migration of workers from the South to the Center and to the North of Italy, which was especially strong in the period 1995-2006, should help to explain the observed geographical pattern (see Basile et al. (2010)).

Figures 15 and 16 show that the growth rate of employment tends to decrease inequality, but the overall impact looks negligible both in terms of inequality (Gini index of counterfactual and actual distribution is indeed the same) and polarization.

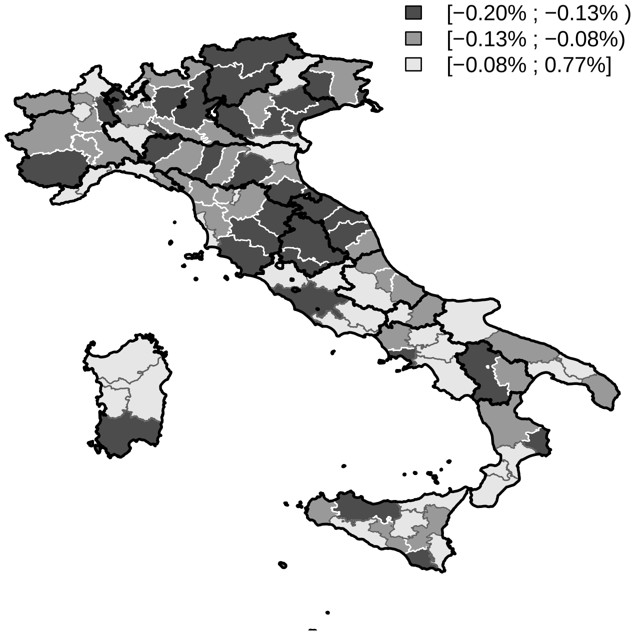

Figure 14: The estimated impact of employment on the annual growth rate of productivity

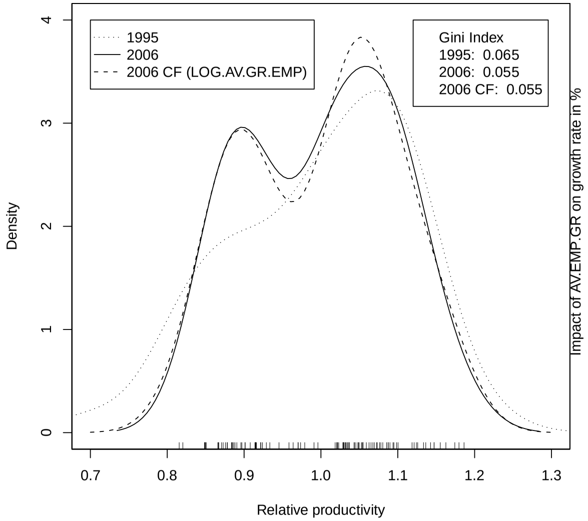

Figure 15: The initial, final and counterfactual final distributions of productivity

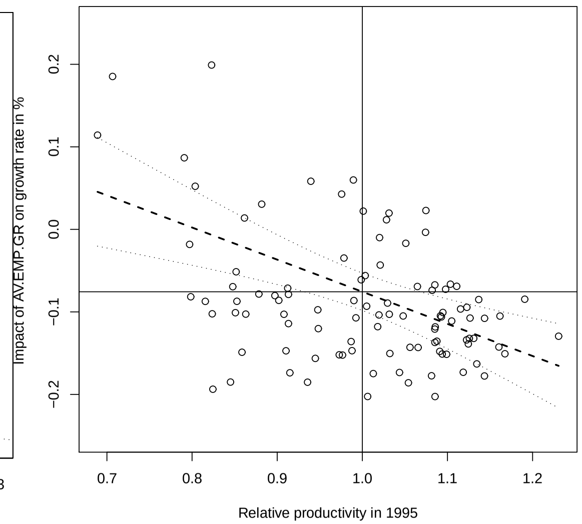

Figure 16: The relationship between the estimated impact of employment on the average growth rate of productivity and the initial level of productivity

EB. Export in 1995

The more export-oriented provinces (as measured by EXP.1995) appear mainly located in the North and in the Center of Italy (see Figure 17). The negative impact of openness on the growth rate of productivity is also low ranging from -0.22% to 0%.

Figures 18 and 19 show that the impact on inequality and polarization of export is negligible (compare Gini indices and polarization of counterfactual and actual distributions).

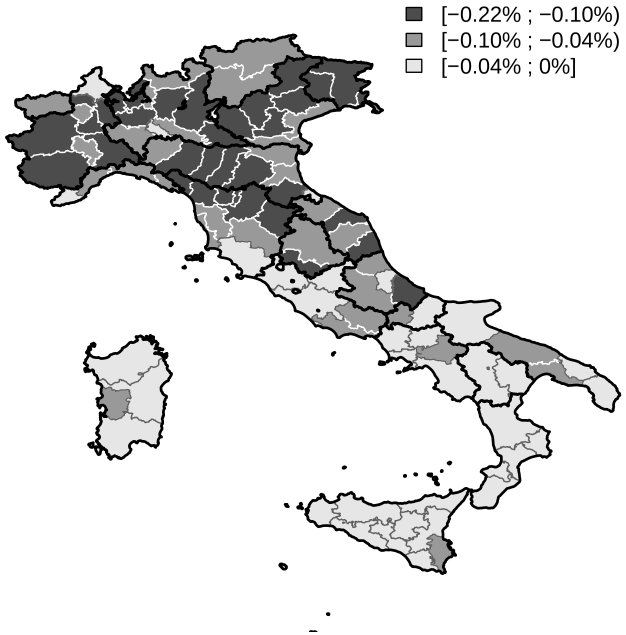

Figure 17: The estimated impact of the share of export in 1995 on the annual growth rate of productivity

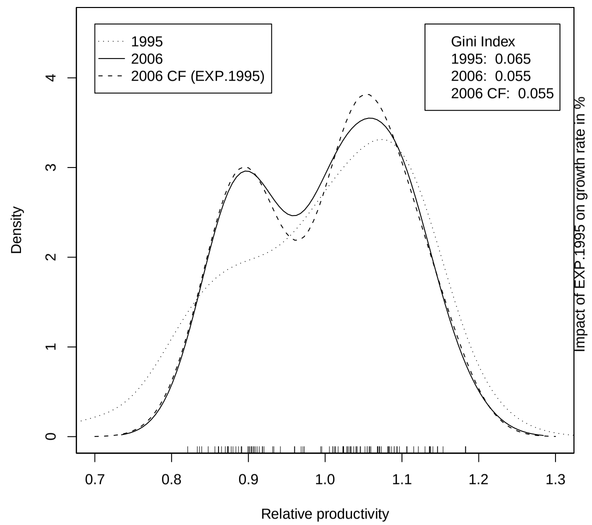

Figure 18: The initial, final and counterfactual final distributions of productivity

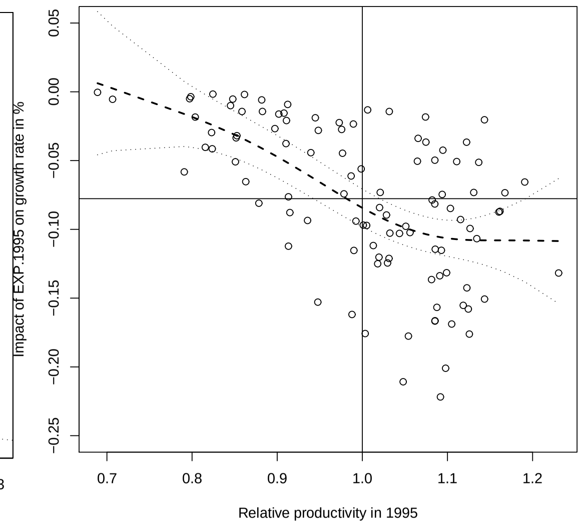

Figure 19: The relationship between the estimated impact of the share of export in 1995 on the average growth rate of productivity and the initial level of productivity

EC. Firms with one Employee on Population in 1996

The (negative) impact of FIRM.SIZE.1.on.POP.1996 on annual growth rate of productivity is high, ranging from -0.71% to -1.45%, with the highest effect for the provinces of Liguria, Emilia Romagna, Piedmont, Tuscany, Umbria, Marche e Lazio (see Figure 20).

However, FIRM.SIZE.1.on.POP.1996 does not appear to have relevant distributional impact (see Figures 21 and 22).

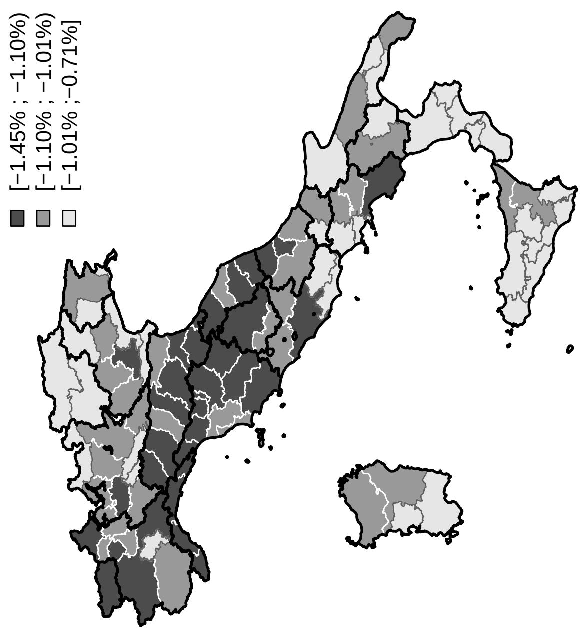

Figure 20: The estimated impact of FIRM.SIZE.1.on.POP.1996 on the annual growth rate of productivity

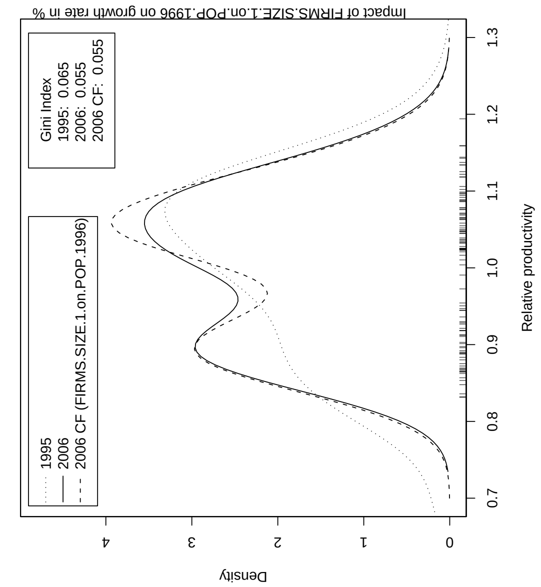

Figure 21: The initial, final and counterfactual final distributions of productivity

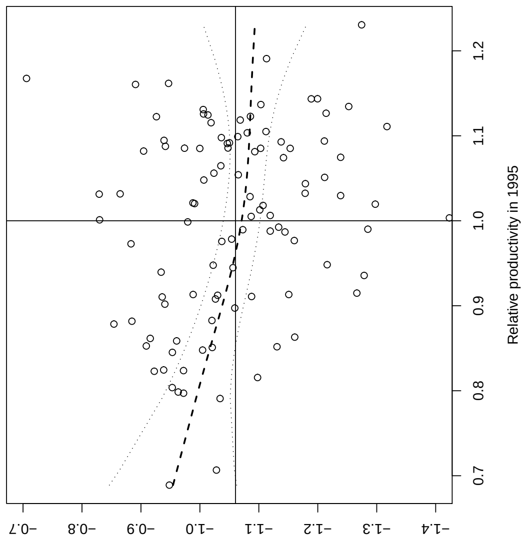

Figure 22: The relationship between the estimated impact of FIRM.SIZE.1.on.POP.1996 (self-employers) on the average growth rate of productivity and the initial level of productivity

ED. Share of Resident Population Six Years Old or Older with Tertiary Education

The provinces of Liguria, Emilia Romagna, Tuscany, Umbria, Marche e Lazio and some provinces in the South of Italy (see Figure 23) appear to have the highest level of TERTIARY.EDU.2001. However, the positive impact of TERTIARY.EDU.2001 on annual growth rate of productivity is low ranging from 0.008% to 0.024%.

Moreover, TERTIARY.EDU.2001 does not appear to have relevant distributional impact (see Figures 24 and 25).

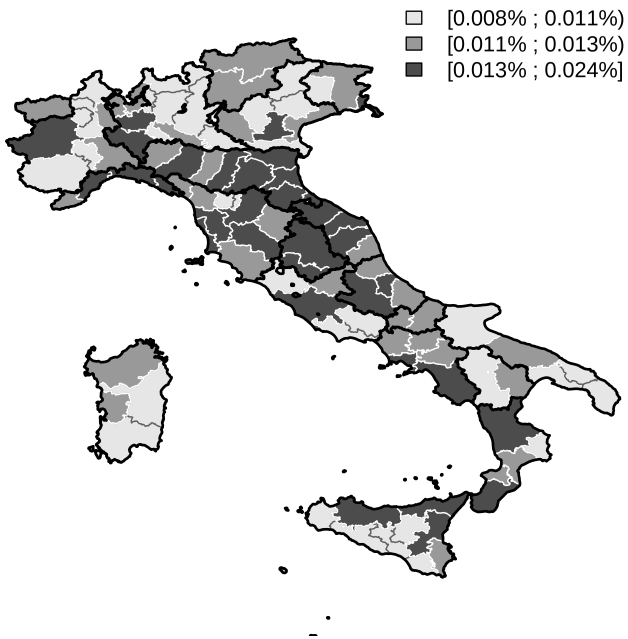

Figure 23: The estimated impact of share of resident population six years old or older with tertiary education on the annual growth rate of productivity

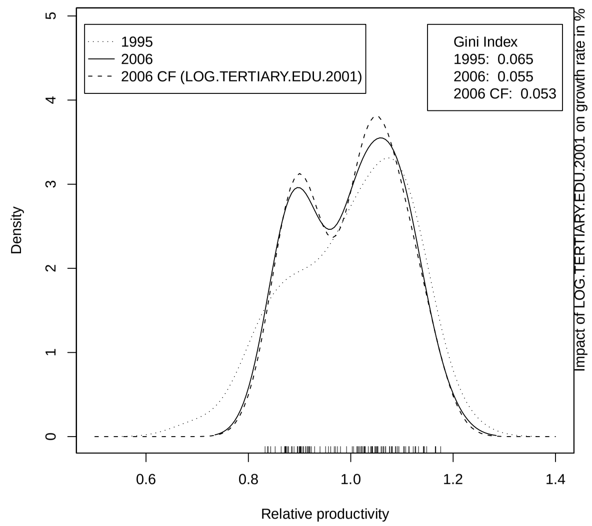

Figure 24: The initial, final and counterfactual final distributions of productivity

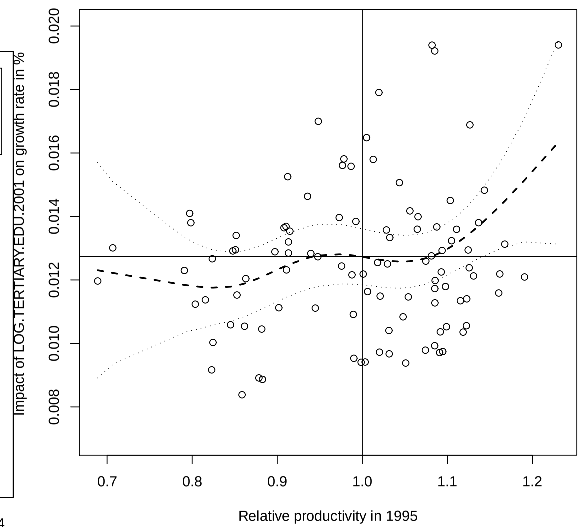

Figure 25: The relationship between the estimated impact of share of resident population six years old or older with tertiary education on the average growth rate of productivity and the initial level of productivity