Quaderni di Dipartimento

A Reassessment of Italian Regional Convergence through a Non-Parametric Approach

Silvia Dal Bianco
(Università di Pavia)

99 (06-09)

Dipartimento di economia politica
e metodi quantitativi
Università degli studi di Pavia
Via San Felice, 5
I-27100 Pavia

Giugno 2009

A Reassessment of Italian Regional Convergence through a Non-Parametric Approach

Silvia Dal Bianco
University of Pavia

June 2009

Abstract

This paper employs the distribution dynamics approach to investigate cross-regional convergence of GDP per worker in Italy, between 1980 and 2003. Two sets of competitive hypotheses are tested: absolute *versus* conditional and neoclassical *versus* technological. Supportive evidence of only technological conditional convergence is found. This means that, should the current dynamic persists, cross-regional convergence will take place only if the differences in technological initial conditions and structural characteristics will be evened out. Moreover, as the pervasiveness of organized crime has been considered as a structural factor, the analysis suggests that technical upgrading together with institutional strengthening should be policy makers' priorities.

*JEL Classification Code:*C14, O33, O47.
Keywords: Italian Regions; Neoclassical and Technological Convergence; Distribution Dynamics.

1 Introduction

Italy is characterized by labor productivity differences among regions, defined as NUTS2.[1] In particular, while the Northern part of the country exhibits labor productivity levels comparable to the most industrialized high-income economies, in the Southern regions labor productivity is similar to the one of upper-middle-income countries, such as Mauritius.[2] The object of this paper is testing whether Southern regions will improve their relative disadvantaged position so that to close the labor productivity gap in the long run, as according the convergence prediction.[3] In particular, employing an unified distribution dynamics framework, originally proposed by Quah (1996), two sets of competitive hypotheses are tested, for the period 1980-2003: first, absolute *versus* conditional and, second, neoclassical *versus* technological convergence.

The use of distribution dynamics is particularly convenient for convergence analysis. Since the seminal work of Quah (1993), the parametric approaches has been progressively substituted by non-parametric and semi-parametric estimation techniques, Durlauf*et al.*(2005). This is because, traditional linear regressions are unable to capture the relative performance of each economy with respect to the others. And this is exactly what matters for the analysis of convergence, meant to check whether *each country eventually becomes as rich as all the others*, Quah(1993).

My contributions to the field are different.

From the methodological perspective, the use I make of distribution dynamics in disentangling the relative strength of capital deepening and technological catch-up is completely new. In fact, the approaches through which neoclassical and technological convergence hypotheses have been compared, so far, are related either to linear regression analysis, such as in Dowrick and Rogers (2002) and Di Liberto *et al.*(2008), or to growth accounting exercises, Maffezzoli (2006) and Wong (2007), or both, Bianchi and Menegatti (2005).

Moreover, my work provides a threefold sensible contribution on the Italian case.

To begin, the paper offers a reassessment on classical and technological convergence in the country, comparing the results available in the recent literature, namely Maffezzoli (2006), Di Liberto *et al.*(2008) and Bianchi and Menegatti (2005).

Then, this is the first study that employs distribution dynamics to test for both absolute and conditional convergence. Magrini (2007), which is the only work I am aware of employing the aforementioned methodology to Italy, tests for absolute convergence only.[4]

[1]The Nomenclature of Territorial Units for Statistics (NUTS) is the classification employed by Eurostat and European Union to identify European regions.

[2]The definitions used are the ones of World Bank Development Indicators, while the figures are based upon author's calculation using Penn World Tables 6.1 in the latest year available, which is 2000.

[3]See, among other reviews, Barro and Sala-i-Martin (1995) and Islam (2003)

[4]Absolute and conditional convergence have been both tested, through distribution dynamics, across Spanish regions and Indian states, respectively in Lamo (2000) and Bandyopadhyay (2006), and among

Finally, one spin-off of my research is constituted by a completely new set of Total Factor Productivity (TFP) levels, obtained employing the superlative index number approach, introduced by Caves (1982a) and Caves *et al.*(1982b) and extensively used by the recent literature, Griffith *et al.*(2004), among others. Such estimates represent a solid ground for further research on Southern Italy technological backwardness.
Turning now to my results, supportive evidence for the technological conditional convergence hypothesis is found.
On the one hand, this means that Italian regions will reach the same labor productivity level, in the long run, if structural differences will be evened out.[5] In particular, as for the inclusion of organized crime among the conditioning factors, my analysis shows that Italian rackets, reducing competitiveness and fundamentally contributing to the misallocation of resources, Lavezzi (2008) and Caruso (2008), inhibit the convergence process, as in Tullio and Quarella (1999). So that, a general institutional strengthening is needed, as precondition, for closing the North-South gap. On the other, it is shown that technological transfer towards Southern regions is the key factor behind labor productivity convergence, in the long run.
The rest of the paper is organized as follows. The second paragraph illustrates the theoretical ground of the alternative convergence hypotheses, namely: absolute and conditional; neoclassical and technological. A brief review of the literature on the Italian case will be done through the text. The third paragraph presents some stylized evidence on the so-called Italian divide, together with the variables employed, data sources and some descriptive evidence on cross-regional convergence. The fourth illustrates the distribution dynamics approach and conditioning techniques. The fifth, the results obtained and their interpretation. Final comments and open lines for further research conclude. Details on the TFP estimation technique employed are reported in the appendix.

2 The convergence hypotheses and evidence from the Italian regions

The *neoclassical convergence hypothesis* is an implication of the Solovian growth framework with or without technological progress and its extensions, Solow (1956), (1957), Mankiw *et al.*(1992). The crucial assumption of diminishing returns to capital implies that, in reaching the long run equilibrium, the lower the initial capital stock per worker the higher the capital accumulation rate and, then, the output per worker growth rate.

developed and developing countries, Quah (1996) and Dal Bianco (2007).

[5]This is a robust result in the literature about Italian convergence. For a synthetic and up-to date review see Magrini (2007).

It must be noted that, although originally the Solow model was meant to explain the growth path of a single economy, in the 1980s cross-country analysis begun to arise, so that according to the neoclassical convergence prediction poorer countries, having an initial lower capital stock, would grow faster and, eventually, catch-up with their richer counterparts.[6]

The testable equation for checking the convergence hypothesis is derived from the transitional dynamics of the neoclassical model with Cobb-Douglas technology. In the case of cross-section analysis, this is:[7]

$$log\left(\frac{y_{i,t+T}}{y_{i,t}}\right) = \alpha + \beta log(y_{i,t}) + \varepsilon_{it} \tag{1}$$

where $y_{i,t}$ is output per effective unit of labor, in country -region- i at time t, $log\left(\frac{y_{i,t+T}}{y_{i,t}}\right)$ is its growth rate and ε_{it} is an idiosyncratic error term.

The key parameter for convergence analysis is β. If it turns out to be statistically significant and negative, *absolute* β *convergence* can be claimed.

As for the Italian case, strong supportive evidence for the absolute β convergence hypothesis is found for the period 1960-1975, then a clear pattern of divergence seems to dominate 1980s and 1990s. Such results are found, among others, by Aiello and Scoppa (2007), Carmeci and Mauro (2002), Paci and Saba (1998), Bianchi and Menegatti (1997), Mauro and Podrecca (1994) and Di Liberto (1994).

A theoretical issue, arising in a cross-section of countries, is that heterogenous economies might exhibit different long-run equilibria, due to differences in their structural characteristics (i.e. saving rate, population growth, development stage reached...). Equation (1) is then modified to take into account steady state differences into:

$$log\left(\frac{y_{i,t+T}}{y_{i,t}}\right) = \alpha + \beta log(y_{i,t}) + \sum_j \phi_j log(x_{j,i,t}) + \varepsilon_{it} \tag{2}$$

where x_j is the value of j-th structural variable. In this case, a negative β implies *conditional* β *convergence*.

Regarding the Italian case, the confirmation of such an hypothesis is a well established result. Aiello and Scoppa (2007), using panel data Arellano and Bover (1995) Generalized Method of Moments (i.e. GMM) system estimator and controlling for inter-regional technological differences, estimate a speed of β conditional convergence of almost 12%, in the period 1980-2002. Carmeci and Mauro(2002), employing Blundell and Bond (1998) linear system GMM estimator and augmenting the standard Solovian framework with a markup proxy and unemployment, find cross-regional conditional convergence in per capita

[6] See Islam (2003) for an historical treatment of this point.

[7] Note that Equation (1) can be applied to panel and time series data. For more details see Durlauf *et al.*(2005)

regional output between 1965-1995. Paci and Saba (1998) find conditional convergence in 1960-1975, controlling for Macro-regional differences (i.e North and South) and using standard OLS. The same technique is employed by Bianchi and Menegatti (1997), which, controlling for Research and Development expenditures, productivity and human capital differences, confirm the conditional convergence hypothesis between 1970 and 1994.

Turning now to the *technological convergence hypothesis*, it originates from the endogenous growth theory of technological catch-up.[8] Especially designed to take into consideration the different stages of technological development reached by different countries,[9] its building block is represented by the innovation-imitation dynamics. In this model, countries are divided into the leader (i.e. the country having the highest technological level), in which technical change is brought about by innovation, and all the others, called followers, for which technical advances are directly linked to the possibility and the ability to imitate leader's technology.[10] To proxy followers' potential of imitiation, it is employed the so called technological gap (i.e. the log difference between leader's and follower's Total Factor Productivity, i.e. TFP)[11]. In this context, convergence tendencies arise because, although innovation tends to increase labor productivity and technological differences between countries, technological diffusion tends to decrease them, Fagerberg (1988).

Traditionally, the empirical test of the two alternative convergence mechanisms (i.e. capital accumulation and technological catch-up) has been carried in the literature either using growth accounting techniques or employing a modified version of Equation (2), where a proxy for the technological gap is explicitly considered.[12]

Concerning the Italian case, only three studies, so far, have attempted to disentangle convergence inner drivers and their results are mixed. Such an evidence will be discussed in more detail in section 5. For now, it is sufficient to briefly mention that Maffezzoli (2006), adopting the Data Envelope Analysis, supports the technological convergence prediction, in the period 1980-2004. Di Liberto *et al.*(2008), using a fixed-effect panel methodology, show that both capital accumulation and technological catch-up were at work between 1963 and 1993. And, finally, Bianchi and Menegatti (2005), employing Ordinary Least Squares, find only neoclassical conditional convergence between 1970 and 2002.

To conclude this overview, it is worth looking at the meaning of testing the β convergence hypothesis, through standard cross-section or panel data regression techniques. Such an exercise consists in verifying the convergence behavior of the representative (i.e. average)

[8]See Rogers (2003) for an excellent review on technological catch-up literature.

[9]As known, the Solovian framework assumes that the production function is the same for all the cross-section of economies considered, so that all countries share the same technological knowledge.

[10]The seminal contributions to this field are constituted by Nelson and Phelps (1966) and Abramovitz (1986).

[11]Total Factor Productivity, also known as Solow Residual, accounts for any output change not accounted by inputs or economies of scale

[12]See Di Liberto *et al.*(2008) for a good review of this literature.

economy.[13]

Two types of problems arise from this observation:

1. as demonstrated by Friedman (1992) and Quah (1993), among others, a negative relationship between growth rates and initial values is a necessary, but not a sufficient, condition for a reduction in the cross-sectional dispersion of per capita income over time;[14]

2. the impossibility of capturing the intra-distributional changes, such as clustering, churning or leapfrogging dynamics, that is how different economies perform with respect to the others along time.

To solve the first of the aforementioned problems, the analysis of β convergence has been complemented, in standard parametric studies, with the one of the so-called σ convergence. According to this hypothesis, a group of economies are converging if the dispersion of their real GDP per worker is decreasing over time, Sala-i-Martin (1996). In the Italian case, σ convergence has been found in most of the studies previously mentioned .

The second and most fundamental issue, instead, has been tackled departing from standard regression techniques.[15] In particular, after the seminal works of Quah (1993) an increasing number of studies has been using the distribution dynamics approach, such as Lamo (2000), Epstein et al. (2003), Bandyopadhyay (2006), Desmet and Fafchamps (2006) and Dal Bianco (2007). This methodology, in fact, allowing the estimation of the law of motion of the entire labor productivity cross-sectional distribution, sheds light on countries' relative performance over time.

3 Data, Stylized facts and a flavor of convergence

This paragraph describes the variables employed and data sources. Moreover, it provides evidence of the Italian macro-regional divide between 1980 and 2003. Finally, it gives a

[13]It must be noted that the convergence prediction, either absolute or conditional, has been investigated also through time-series techniques, namely: unit root tests, as Evans and Karras (1996), or cointegration analysis, as in Bernard and Durlauf (1995). A detailed treatment of those techniques goes beyond the scope of this paper, given that, employing time series tests, the steady state variation is generally limited to time-invariant differences and trend breaks. For more details see Islam (2003) and Durlauf *et al.*(2005).

[14]To show this point, Quah (1993) advocates the so-called Galton's fallacy:

> Galton, in aristocratic manner, was concerned about the sons of tall fathers regressing into a pool of mediocrity along with the sons of everyone else (...) He could not, however, reconcile this with the population of male heights continuing to display significant cross-section dispersion.

[15]Quah (1996a) demonstrates, in fact, that a constant standard deviation is consistent with very different dynamics ranging from poverty traps to leapfrogging.

flavor of cross-regional convergence tendencies, considering some standard indicators. Two features clearly emerge from such a descriptive analysis: first, Figure 1 shows that the gap between Northern and Southern regions is still well open and, second, the evidence regarding convergence tendencies is mixed. In particular, while Figure 2 seems to support both absolute β and σ convergence, Figure 3, showing the persistence of regions' relative position and the increasing bimodality of labor productivity distribution, does not.

3.1 Variables and Data sources

My empirical analysis focuses on the evolution over time of log relative GDP per worker distribution, that is labor productivity in region i, at time t, relative to the one of Lombardy, the leader among Italian regions. Although not having the highest labor productivity in the whole period, Lombardy can be considered the Italian leader because it is the most innovative among the Italian regions, as reported by LabMiM (2006).[16]

Normalizing with respect to leader is a very convenient way of removing (some of) the trend from the cross-section, Quah (1996). As noted in Desmet and Fafchamps (2006), working with de-trended data is of particular importance to avoid degenerate long-run distributions. Further, it must be observed that this normalization leaves unaltered how regions differ from each other but, obviously, requires to take Lombardy out from the cross-sectional units under study. So, employing distribution dynamics, the behavior of 19 regions' relative productivity will be analyzed.

Labor productivity, in each region, is measured as real GDP per worker, where 1995 is the base year and workers are measured in standard units of labor. Both series are obtained from ISTAT, Economic Regional Accounts.[17]

To test for conditional convergence, standard steady state proxies as the investment rates in both physical and human capital and three macro-regional dummies (i.e. one for North, for Center and South Italy) were used together with the number of homicides per 100.000 inhabitants, to proxy for organized crime, as in Tullio and Quarella (1999).

Investment rates in physical capital refer to Gross Fixed Capital Formation (GFCF) share to GDP. GFCF series comes from ISTAT, Economic Regional Accounts.

To proxy human capital accumulation rate, I use the average years of schooling in the workforce, taking the data from Ascari and Di Cosmo (2005).

[16] LabMiM, which pertains to Milan Chamber of Commerce, combines five indicators of Eurostat (2005) to compute the innovativeness of European and Italian regions. The indicators are: number of innovative firms in R&D and IT sectors; number of patents registered in the European Patent Office; number of patents in knowledge intense industries registered in the European Patent Office; knowledge intense manufacturing sectors' employment and knowledge intense manufacturing and services sectors' employment. Lombardy turns out to be the Italian technological leader and it is ranked 14th among European regions.

[17] ISTAT is the Italian National Institute for Statistics.

Among other indicators, such as robberies, frauds, blazes, extorsion, the number of voluntary homicides proxies the presence of criminal organizations rather well. Firstly, it is less likely to be underreported and, secondly, *mafia* interests almost always lead killing decisions in Italy, Tullio and Quarella (1999).[18] Finally, such a series is available in the whole period considered in ISTAT, Justice and Security Statistics. The use of composite measures, such as the so-called organized crime index, employed by Caruso (2008) and published by ISTAT, Breaking Variables-AsseVI, is inhibited because the series is only available from 1995 to 2003.

In the empirical implementation, such variables, dummies apart, are taken in natural logarithms and normalized with respect to Lombardy values.

I turn now to present the variables employed for discriminating between neoclassical and technological convergence.

In checking whether labor productivity convergence dynamics can be eventually ascribed to physical capital accumulation, I employ the regional capital stock series estimated by Maffezzoli (2006), aggregating private and public investment flows in both machinery and buildings, according to the Perpetual Inventory Method with random service life. As before, such variable is taken with respect to Lombardy and is expressed in natural logarithms.

To retrieve, instead, the relative strength of technological transfer for overall convergence, I employed the Total Factor Productivity Gap (TFPgap), that is the difference in technological levels (i.e. TFP levels) between the leader and any other region.[19] In particular, TFP levels were originally estimated employing the superlative index number approach introduced by Caves (1982a) and Caves *et al.*(1982b), using the previously mentioned data sources. By its very construction, it is already expressed in natural logarithms and no normalization is needed.[20]

[18]An exception to this general rule might be represented by homicides perpetrated by political extremists. Although diffused in the 1970s, this sort of practice turned down in the period considered.

[19]More precisely, according to the theory of technological catch-up, technological transfer depends on both the TFPgap, taken to proxy the potential for technological imitation, Griffith *et al.*(2004), and to the so-called absorption capability (i.e. recipient economies' ability to assimilate and fruitfully exploit new knowledge). Gerschenkron (1954) and Baumol (1986) provide the seminal contributions of the so-called 'capabilities approach'. Unfortunately, Italian regional data on capabilities' proxies, such as for example Research and Development expenses, registered patents, schooling attainment rates, number of articles published in scientific journals, are recorded for very short periods or do not exist at all. So that, only TFPgap has been considered.

[20]More details on the estimation methodology employed are given in the Appendix.

3.2 Stylized Facts: Italian Macro-regional divide

A flavor of the Italian divide, between 1980 and 2003, is given by Figure 1, where the 20 Italian regions (i.e. NUTS2) are aggregated into three categories: North, Center and South. [21] Interestingly, Southern regions out-perform the others with respect just to the number of homicides (panel f) and the physical capital stock per worker (panel d). This is hardly surprising. First, according to CENSIS (2007), the 77% of resident population in Southern Italian regions of Apulia, Campania, Calabria and Sicily lives in towns in which organized crime is recorded. Second, capital stock data were estimated by Maffezzoli (2006) aggregating private and public investment flows in both machinery and buildings, where buildings represent the 80% of total capital stock in Southern regions and almost the 70% in Northern ones. Adding to that the figures provided by Bonaglia and Picci (2000), according to which public capital stock in Southern regions is around 40% while in the Northern is 27%, it is quite clear that the majority of capital stock in South Italy consists in publicly financed infrastructure. Finally, it is worth noting that organized crime pervasiveness and high public investment can be considered as mutually enforcing. Gambetta and Reuter (1995) and Del Monte and Papagni (1998) document the role of criminal organizations in public procurements, describing how collusive agreements are enforced and guaranteed by such organizations. Moreover, Caruso (2008) found a statistically significant correlation between public expenditure and organized crime across Italian regions while, according to Lavezzi (2008), the pervasiveness of *mafia* in Sicily is raised by a huge public sector.

Turning to the other charts, panel (e) shows the general improvements in human capital accumulation, although cross-regional differences are still quite marked. In fact, 11 years of schooling in Italy imply the completion of a technical high school. Being below that level, as South Italy, means that, on average, people has formally completed just the secondary school.

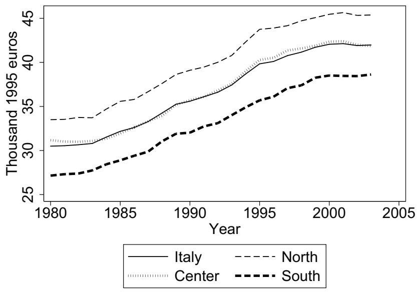
Panels (b) and (c) report the pattern of the technological proxies. Three things could be noted. First, the poor but improving performance of Southern regions. In fact, they are getting closer to the average TFP and the TFPgap seems to diminish with time (in both cases the relevant value is zero, for how the proxies were constructed). Second, Central region's catch-up, in terms of both variables, and, third, the quite poor performance of North from 1995 onwards. The loss of Northern regions innovativeness is usually indicated as one of the causes of the relative decline of Italy with respect to other European countries, Alesina and Giavazzi (2006).

Finally, panel (a) shows that labor productivity has evolved along parallel patterns among

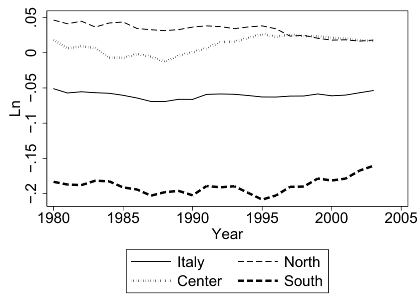
[21] The Northern regions are: Lombardy, Piedimont, Valle d'Aosta, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Liguria and Emilia Romagna; the Central: Tuscany, Umbria, Marche and Latium; the Southern: Abruzzo, Molise, Campania, Apulia, Basilicata, Calabria, Sicily and Sardinia.

Italian macro-regions. It seems, then, that there was no convergence in the period under analysis. This is not a conclusive statement, some further evidence will be provided in the next section, where each region is considered as a single observational unit.

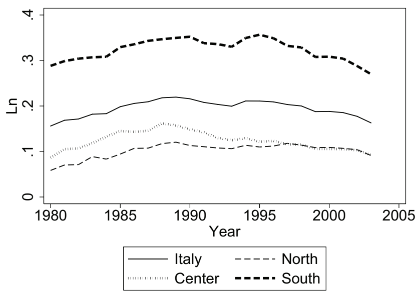
Figure 1: Italian Macro-Regional Division, averaged series



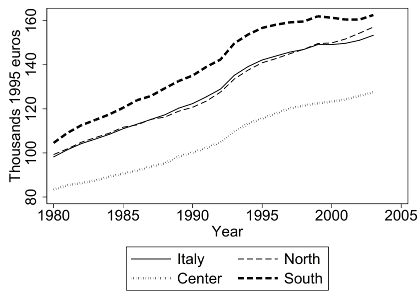
(a) GDP per worker



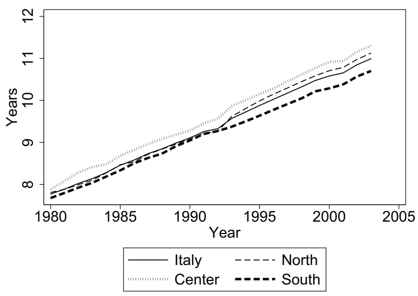
(b) TFP Level



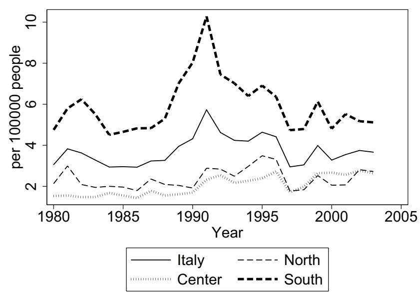
(c) TFP gap



(d) Capital stock per worker



(e) Education

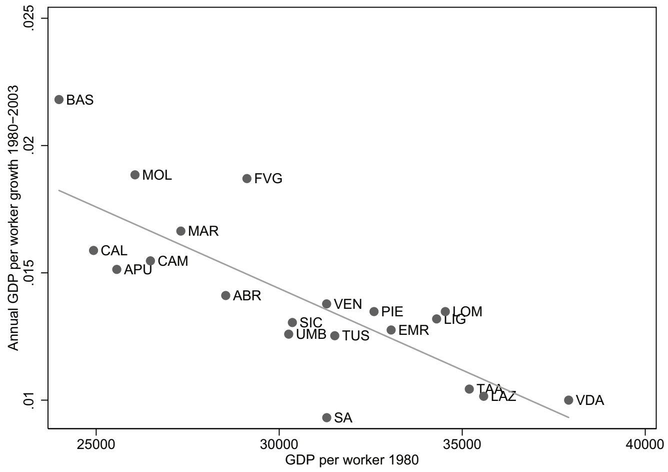


(f) Homicides

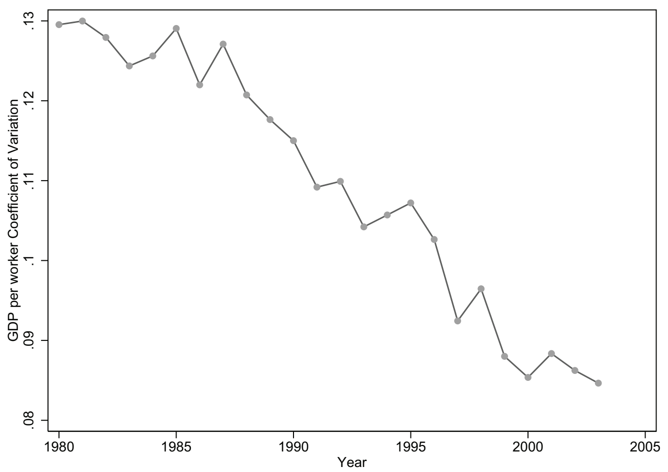
3.3 Cross-regional convergence

To tackle the convergence hypothesis more closely, I provide some 'graphic intuition' of β and σ convergence, showing both the correlation between GDP per worker at the beginning of the period and whole period GDP per worker growth rate (Panel a) and the cross-sectional coefficient of variation (Panel b).[22]

Figure 2: GDP per worker convergence signs. GDP in 1995 euros.



(a) β Convergence



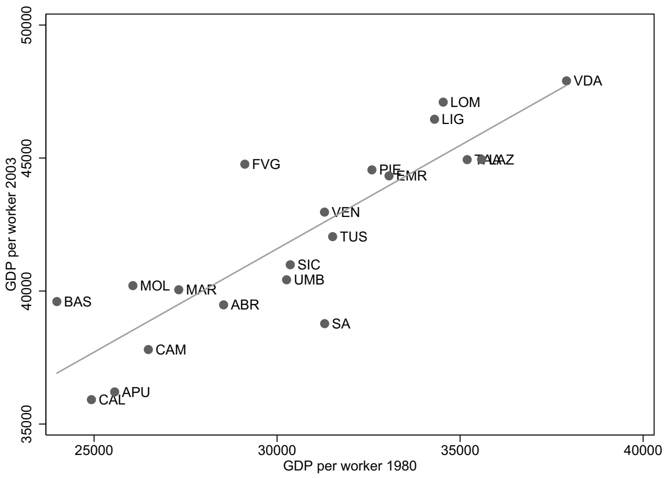
(b) σ Convergence

At first sight, the graphs above seem to provide supportive evidence to the convergence

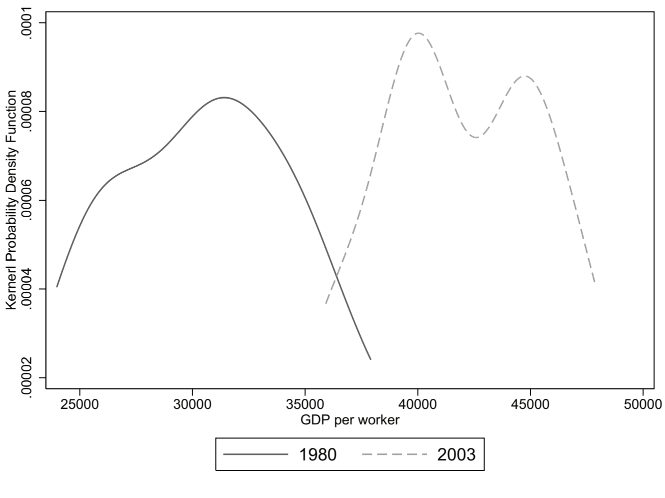
[22]The coefficient of variation (i.e. the standard deviation divided by the mean) is preferable to standard deviation because, being a normalized measure of dispersion, it eliminates the comparisons problems related to a changing mean.

hypothesis. Poorer regions grow relatively faster and cross-regional dispersion diminishes over time. However, such an analysis overlooks the relative position of the regions along the period and the shape of the overall GDP per worker distribution, which might well show churning dynamics among poor and rich regions. These facts are considered in the next graphs, where the correlation between labor productivity in 1980 and 2003 is mapped together with these variables' distributions.

Figure 3: GDP per worker divergence signs. GDP in 1995 euros.



(a) Persistence



(b) Initial and final distributions

Figure 3, panel (a), highlights two features. First, in absolute terms, each region has experienced a labor productivity improvement in the period considered. Second, the high degree of persistence in regions' relative position. With a correlation coefficient of 0.864

between initial and final GDP per worker, it could be easily seen that poor regions stay poor and rich regions stay rich.
Panel (b) reports the distributions of labor productivity in the first and in the last year considered. This panel, showing the shift towards higher GDP per worker levels, is consistent with the labor productivity improvement already shown by panel (a). However, the most interesting feature is the clustering dynamics which makes the 2003 distribution clearly bimodal.
Such a descriptive evidence provides a prolific ground for distribution dynamics analysis, which is discussed in the next section.

4 Methodology

4.1 Distribution dynamics and conditioning: a brief non-technical summary

When distribution dynamics is employed, convergence tendencies among observational units, which are the Italian regions in this specific case, can be retrieved analyzing the evolution along time of cross-regional relative labor productivity distribution.[23] In particular, the main question to be answered is whether all economies considered will converge to the same level of labor productivity, such that the cross-regional distribution is single peaked, or whether the economies converge only within small clubs, such that the distribution exhibits more than one peak.
Operatively, the changes along time of cross-regional labor productivity distribution are retrieved using the stochastic kernel density estimator. In fact, this estimator allows to measure the probabilities of dynamic transitions from one labor productivity class to another, for each region.
Intuitively, the stochastic kernel can be thought as a refinement of the histogram. In particular, while in histogram the frequency distribution is calculated for disjoint states, with kernel density estimator the frequency distribution is estimated for a large number of overlapping class intervals, which gives a much smoother appearance, resembling a probability density function.
Two are the types of kernels employed in this paper:

1. unconditioned kernels
2. conditioned kernels

The unconditioned kernels give information on the likelihood that an economy, starting from a given relative position in the initial period t, will end up improving or worsening its

[23] Please note that in this section 'relative labor productivity' and 'labor productivity' are used interchangeably.

relative position in the final period $t+s$. In other words, it can be said that unconditioned kernels measure the transition probabilities from t to $t+s$.
Unconditioned kernels are used here to test the absolute convergence hypothesis.
Conditioned kernels are an extension of unconditioned ones. In particular, they allow to identify the factors that eventually lead to intra-distributional changes. In fact, the effects of conditioning are identified by changes in shape and location of the kernel, with respect to the unconditioned case.
I will use conditioned kernels for testing conditional convergence hypothesis and for disentangling neoclassical from technological convergence. In particular, if the unconditioned kernel shows twin peaks feature and, after conditioning with respect to steady state proxies, it is found that the conditioned kernel is single peaked, then, it can be said that clustering dynamics is lead by structural differences and that conditional converge hypothesis can not be rejected. At this point, neoclassical *vs* technological convergence can be tested employed the same technique.

4.2 Unconditioned transition probability estimates

In this section I provide a technical illustration of the methodology employed to estimate unconditioned transition probabilities, which are used to test the absolute convergence hypothesis.
With y_{it}, I indicate the logarithm of relative labor productivity, that is individual region i GDP per worker relative to the one of Lombardy, at time t (i.e. $y_{it} = log(Y_{it}/Y_{Lt})$) and with $f_{Y_t}(y_t)$, the cross-regional labor productivity distribution at time t, where Y_t indicates the corresponding random variable.
I assume that year-to-year changes in the distribution of labor productivity can be represented by an homogeneous Markow process, in such a way that, $\forall t$:

1. $f_{Y_{t+1}|Y_t}(y_{t+1}|y_t) = f_{Y_{t+1}|Y_t}(y_{t+1}|y_t, y_{t-1}, y_{t-2}, ...)$

2. $f_{Y_{t+1}|Y_t}(y_{t+1}|y_t) = f_{Y_t|Y_{t-1}}(y_t|y_{t-1})$

The first property guarantees that only previous period income distribution impacts on next period one (i.e. history does not matter). The homogeneity assumption in 2 ensures that the transition probabilities do not vary with the time. Although quite restrictive, both hypotheses are necessary for estimating long run transition probabilities given the available data.
Conditional density functions, $f_{Y_{t+1}|Y_t}(y_{t+1}|y_t)$, represent the cornerstone of distribution dynamics convergence analysis. This kind of distribution, in fact, encodes information about individual economies' passages over time. Thus, it sheds light on both intra-

distribution dynamics and external shapes, making inference about convergence tendencies possible. For example, observing conditional density mappings, is it possible to know whether poor economies are catching-up with their richer counterparts, whether rich are still enriching, whether regions are converging overall or are clustering within clubs.
The empirical estimation of conditional densities is handled by non-parametric techniques. By definition, the conditional distribution is the joint distribution divided by the marginal distribution. Formally:

$$f_{Y_{t+1}|Y_t}(y_{t+1}|y_t) = \frac{f_{Y_{t+1},Y_t}(y_{t+1}, y_t)}{f_{Y_t}(y_t)} \tag{3}$$

The joint distribution of (Y_{t+1}, Y_t) can be estimated non parametrically using a bivariate stochastic kernel, while the marginal distribution of Y_t is obtained by numerical integration of the joint distribution. Finally, the conditional distribution is simply obtained by dividing one to the other, after appropriate discretization of the joint support.[24]
Long run tendencies towards convergence are encoded by the ergodic distribution. This is the stationary distribution of labor productivity, which will be approached in the long run should certain technical conditions hold.[25] In particular, if the ergodic distribution is unimodal and has a low variance, then long run cross-country convergence can be claimed. Formally, the ergodic is the distribution f which solves the following functional equation:

$$f(y_{t+1}) = \int_{-\infty}^{+\infty} f_{Y_{t+1}|Y_t}(y_{t+1}|y_t)f(y_t)dy_t \tag{4}$$

In order to compute the ergodic distribution the support of y is discretized in a set of N equally large intervals, where interval h is denoted as Ω_h.[26]
Then, the probabilities of transition from one interval to another are calculated. Formally the probability of transition from the interval Ω_h to another, Ω_k, in one time period, is denoted as:

$$\alpha_{kh} = Pr(y_{t+1} \in \Omega_k | y_t \in \Omega_h)$$

At this point, it is useful to adopt a compact matrix notation. Hence, the ergodic distribution is the vector p that solves the following system of equations:

$$p = Ap$$

[24]Bivariate stochastic kernel estimation is performed using the command *kdens2* in STATA 8.2. Marginal, conditional and ergodic distributions are calculated in Matlab. All programs are available from the author upon request.

[25]See Stockey, Lucas and Prescott (1989); Luenberger (1979).

[26]To avoid crude ergodic calculations, it is necessary to work with a sufficiently high N. My calculations have been done for N=50. Using N=200 does not alter any conclusions but it has the disadvantage of slowing down computer's routines.

$$(I - A)p = 0$$

where each component of the vector p represents the probability of y assuming a value comprised in a given Ω and A is the matrix of transition probabilities α_{kh}.
Since each column of matrix A is a marginal density and, then, its elements sum to 1; A does not have full rank and, by consequence, the system does not have a unique solution. To find a unique solution it is standard to simply drop one row of A (to make its columns linearly independent) and then add the restriction that the entries of vector p sum to 1.[27]
Then, matrix A is rewritten as B:

$$\text{B=}\begin{pmatrix} 1-\alpha_{11} & \dots & -\alpha_{1N} \\ \dots & 1-\alpha_{ii} & \dots \\ -\alpha_{N-1,1} & ... & -\alpha_{N-1,N} \\ 1 & ... & 1 \end{pmatrix}$$

The modified system is then:

$$Bp = b$$

where the vector b, for the constraint added, has all entries equal to 0 except the last one, which is equal to 1.
At this point, the unique ergodic distribution, p, can be easily found inverting B:

$$p = B^{-1}b$$

4.3 Conditioning techniques

This part outlines the conditioning technique I used to test for conditional convergence and neoclassical *versus* technological convergence.
Under the conditional convergence hypothesis, cross-country productivity equalization can not be found in the original relative labor productivity distribution, f_Y, but in the conditioned one, $f_{Y|X}$, where X denotes steady state proxies. Then, the object of interest are the transition probabilities of the part of labor productivity not explained by the auxiliary variables (i.e. steady state proxies). Employing the former notation, such transition probabilities are formally written as:

$$f_{Y_{t+1}|Y_t,X_t}(y_{t+1}|y_t, x_t) \tag{5}$$

Exploiting Chamberlain (1984) results, the part of labor productivity orthogonal to auxiliary variables is computed as Ordinary Least Squares (OLS) residuals of the projection of

[27] This constraint must hold for the definition of probability.

labor productivity growth on each of the steady state proxies.[28] Such calculation involves three steps:

1. estimating the part of countries' relative productivity *growth rate explained* by conditioning steady state variables;

2. finding the *initial level* of relative labor productivity *explained* by conditioning steady state variables ;

3. combining the previous results to find the *level* of relative labor productivity *unexplained* by the auxiliary variables (i.e. orthogonal to steady state proxies).

Call g_{it} the growth rate of y_{it} (i.e. log relative productivity in region i at time t). Name $\widehat{g_{it}}$ the part of g_{it} explained by steady state proxies, which are: investment rate in both physical and human capital, indicated as r_{it} and h_{it}, the macro-regional dummies, dN, dC and dS, and organized crime proxy, oc_{it}. Finally, the part of labor productivity orthogonal to steady state proxies, which is the object of interest, is called $\widehat{\epsilon_{it}}$.
Step 1. is implemented regressing g_{it} on a two sided distributed lag of conditioning variables and saving the fitted values. For each steady state proxies one of such regressions is run. Then, cumulating the fitted values by region, the part of regions' relative productivity growth rate explained by conditioning steady state variables, $\widehat{g_{it}}$, is obtained.
Note that in empirical work, multi-sided regressions are employed to handle endogeneity issues, which are represented in this specific case by the likely bidirectional causality between labor productivity growth rate and steady state proxies. This technique, introduced by Sims (1972), has been extensively used by Quah, who noticed that just 2 leads and 2 lags are sufficient to clear the estimated growth rate from feedback effects, Quah(1996).
Step 2. is taken running a pooled OLS regression of y_{it} on time averages of steady state proxies (i.e. $\overline{r_{it}}$, $\overline{h_{it}}$ and $\overline{oc_{it}}$) and the estimated growth rate (i.e.$\widehat{g_{it}}$). For each sector, the coefficients that solves the following minimization problem are used to pin down the initial level of labor productivity explained by steady state variables, $\widehat{y_{i0}}$:[29]

$$min_{\beta_1,\beta_2,\beta_3,\beta_4,\beta_5,\beta_6} \sum_i \sum_t [y_{it} - (\beta_1 \overline{r_{it}} + \beta_2 \overline{h_{it}} + \beta_3 dN + \beta_4 dC + \beta_5 dS + \beta_6 \overline{oc_{it}} + \widehat{g_{it}})]^2$$

[28]Quoting Quah(1996), Chamberlain(1984) finds that:

> the projection of growth on investment, not allowing for individual effects, is precisely the best linear predictor and, thus, correctly gives residuals that are the components unexplained by (or, more correctly, orthogonal to) investment.

[29]As Quah(1996) explains, this technique exploits the cross section variation of conditioning variables to compute the initial value of productivity explained steady state proxies.

In fact, thanks to the estimated coefficients, $\widehat{\beta}s$, the initial level of log relative labor productivity explained by conditioning variables can be expressed as:

$$\widehat{y_{i0}} = \widehat{\beta}_1\overline{r_{it}} + \widehat{\beta}_2\overline{h_{it}} + \widehat{\beta}_3 dN + \widehat{\beta}_4 dC + \widehat{\beta}_5 dS + \widehat{\beta}_6\overline{oc_{it}}$$

Then, adding the growth rates of step 1, the level of relative labor productivity explained by steady state variable is calculated as:

$$\widehat{y_{it}} = \widehat{y_{i0}} + \widehat{g_{it}}$$

Finally, $\widehat{\epsilon_{it}}$, which represents the productivity level not accounted for (or conditional to) steady state proxies is simply found subtracting from actual the estimated relative labor productivity:

$$\widehat{\epsilon_{it}} = y_{it} - \widehat{y_{it}}$$

Once region specific $\widehat{\epsilon_{it}}$ series have been calculated, the empirical implementation for testing conditional convergence is the same as absolute (or unconditional) convergence.
In particular, bivariate stochastic kernel densities fit cross-regional distribution of relative productivity orthogonal to steady state variables, which I denote as $f_{\widehat{E_{t+1}},\widehat{E_t}}(\widehat{\epsilon_{t+1}}, \widehat{\epsilon_t})$. By numerical integration of the joint distribution, the marginal density $f_{\widehat{E_t}}(\widehat{\epsilon_t})$ is obtained. Finally, the transition probabilities of Equation(3) are found dividing the joint distribution, $f_{\widehat{E_{t+1}},\widehat{E_t}}(\widehat{\epsilon_{t+1}}, \widehat{\epsilon_t})$, by the marginal distribution, $f_{\widehat{E_t}}(\widehat{\epsilon_t})$.
Long-run distribution of relative labor productivity conditioned to steady state variables is retrieved from the ergodic distribution of random variable $\widehat{\epsilon_t}$. Such a distribution is calculated as for the unconditional case described in the previous section.
Turning now to the analysis of convergence inner drivers, it should be intuitive that the conditioning scheme described so far can be easily extended to determine the relative strength of capital accumulation or technological catch-up.
In particular, if conditional convergence hypothesis holds, the object of interest becomes the dynamics of labor productivity distribution conditioned to both steady state proxies and capital or technological initial conditions. Formally, the following transition probabilities has to be computed:

$$f_{Y_{t+1}|Y_t,X_t,Z_t}(y_{t+1}|y_t, x_t, z_t) \tag{6}$$

where the variable Z represents either initial capital stock or technological level.
For example, to retrieve the relative strength of capital deepening as convergence determinant, relative labor productivity orthogonal to both steady state proxies and capital stock initial level must be calculated. This is done implementing the three steps previously described, taking capital stock as an extra conditioning variable.
By the same tokens as before, the density in Equation (4) and the ergodic distributions are computed.

To conclude, it is worth noting that the conditioning scheme I employed allows not only to work out alternative convergence hypothesis within a unified framework but also to calculate the ergodic of distributions that have been conditioned to time varying (and likely endogenous) variables.[30]

4.4 Interpreting results

I now provide the fundamental tools for inferring convergence tendencies from the graphs that constitute the results of my analysis. Such diagrams, mapping the transition probabilities of different types of distribution (i.e. unconditional, conditional to steady state proxies, etc), allow to test for alternative hypothesis of convergence.

Panels (a) and (b) of Figures from 4 to 9 describe 5 year horizon distributions' evolution and are used to establish medium run tendencies to convergence.[31] More precisely, panel (a) shows a tridimentional plot of transition probabilities, estimated by stochastic kernels and panel (b), mapping the level curves, represents the stochastic kernels in just two dimensions. In both diagrams, the floor axis, marked as Period t and Period $t+5$, measure the log of relative productivity in different times.[32]

Observing the bidimensional plots, *convergence* tendencies, in the medium run, can be claimed if the kernel rotates clockwise and accumulates on a single ridge parallel to Period t axis. That is, relative productivity levels become equal across regions, regardless of economies' initial position. *Persistence* is found when the mass concentrates along the 45 degrees line. So, countries' initial and the final positions coincide. *Improvements*, with respect to the initial position, are detected if the mass piles above the 45 degrees line; by the same token, *worsening* occur when the mass lies below the diagonal. *Club convergence* is signalled by distinct peaks along the diagonal.

As explained in the methodological section, long run tendencies, should the current dynamics persist, are assessed through ergodic distributions, like panels (c). Ergodic's x-axis represents the support of labor productivity initial distribution (i.e. the 1980 one) and ergodic's shape can be anticipated by mobility analysis. Mobility analysis, in fact, values whether countries will change their relative position over time or not. An example can easily clarify this point.

[30]Such an improvement in ergodic distributions calculations was implemented for the first time in Dal Bianco (2007) and it represents a step forward with respect to long run convergence analysis based on both discrete transition probability matrices such as in Quah (1996), Quah (1997), Epstein et al. (2003), Bandyopadhyay (2006) and time invariant conditioning factors, like in Desmet and Fafchamps (2006).

[31]I also calculated transitions over one, four, six and eight years horizon. As the results do not change significantly, I choose 5 years periods because of its standard use in the literature. See Caselli *et al.*(1996) and references therein.

[32]To make graph interpretation easier, in Table 1 I explicitly express Period t and Period $t+5$ values, in % terms with respect to the leader (i.e. Lombardy).

Take the contour plot of Absolute convergence. It can be seen that there are four groups of regions, signalled by red and yellow circles. Call them: poorest, poor, rich and richest. Their starting positions in period t are, respectively: from -0.45 to -0.35; -0.32 to -0.18; -0.1 to 0 and, finally, 0.01 to 0.1. Pick the poorest and ask: where will these economies end up in the next five years? -0.35 is the answer. Then, ask again: what will happen to the same regions in the following five years, should the current tendencies persist? The most likely outcome is that they will improve their relative position, until a maximum of almost -0.2. Thus, the poorest will get relatively better. This kind of dynamics can be appreciated looking to clockwise rotation of the lowest part of the distribution, where most of the mass lies above the main diagonal. Take the poor now. Starting around -0.2 in period t, they will be trapped around that level for all the subsequent periods. Such a dynamics can be inferred looking at the peak, which is a 'convergence basin', centered in -0.18, along the 45 line. Interestingly, poorest and poor regions will converge, in the long run, to the same labor productivity level (i.e. -0.18). By the same tokens, it could be easily seen that rich and richest regions will either stay put around 0 or worsen their relative initial position (i.e. the distribution rotates anticlock-wise for the richest regions). The mobility analysis, then, predicts an ergodic with at least two peaks. Such intuition is confirmed by Panel (c).

I turn now to the interpretation of long run convergence tendencies.

In general terms, it can be said that any alternative convergence hypothesis is not rejected when the correspondent ergodic distribution is unimodal and has a low variance.

In the case of absolute convergence, a single peaked ergodic means that labor productivity will be equalized among all countries, no matter their difference in structural characteristics or initial conditions.

When absolute convergence is rejected, conditional convergence is tested. If its ergodic turns out to be unimodal, then regions' structural characteristics are responsible for the lack of absolute convergence. If conditional convergence is not rejected, the issue becomes disentangling convergence inner drivers. In particular, neoclassical conditional convergence hypothesis is not rejected when the ergodic distribution, conditioned with respect to steady state proxies and capital stock, is unimodal. The same reasoning applies for establishing technological conditional convergence.

5 Discussing Results

On the basis of Figure 4, absolute convergence hypothesis is safely rejected. As discussed in the previous section, the ergodic distribution exhibits two peaks, in correspondence of, respectively, 83% and 98% of Lombardy's initial labor productivity together with a fat

upper tail. Moreover, the majority of the regions will not catch-up with the leader in the long run (i.e. the median is below 0).
Conditional convergence is supported by the evidence provided in Figure 5. From panel (b), the clockwise rotation of the distribution shows that poor are becoming relatively richer and rich are becoming relatively poorer, so that they getting closer to each other. Moreover, the absence of peaks centered along the 45 line rules out the possibility of distinct convergence basins. Such intuitions are confirmed by the ergodic distribution which is singled peaked around 0.07 and has a median well above 0. This means that, in the long run, Italian regions will converge to labor productivity level equals to 37450 euros, at 1995 constant prices, which is 107% of the Lombardy's one in 1980. The precondition to be met, although, is smoothing out structural differences, that, in this specific case, are related to physical and human capital investment rates, organized crime and macro-regional differences. Then, a mix of economic policies and institutional strengthening has to be implemented to close the cross-regional labor productivity gap.
In Figures 6 and 7, neoclassical and technological conditional convergence hypotheses are tested. From their comparison, it is easy to see that supportive evidence for only technological conditional convergence is found. In fact, only the ergodic in Figure 7 is single-peaked. This means that, should the initial technological differences be smoothen, together with the structural characteristics, and should the current dynamics persist, regional labor productivity will be equalized in the long run. Moreover, it is interesting to note that capital deepening, although assumed to be even across regions, will act as a force enhancing inequality. It might be the case, in fact, that decreasing marginal returns to capital will cause slow economic growth in the Southern regions, which are relatively well equipped in terms of physical capital stock per worker,[33] enlarging the Italian gap.[34]
From a policy perspective, my analysis shows that fostering productive investment in South Italy is quite a myopic strategy if the aim is smoothing out cross-regional disparities. That is, the equalizing mechanism behind the conditional convergence result is technological catch-up, holding the *coeteris paribus* condition. So that, a wise economic policy would reinforce such dynamics, providing sound incentives for technological upgrading, structural production shifts from low tech to high tech products, higher education boosting and Research and Development activities. In fact, such actions will eventually ensure both GDP per capita growth and cross-regional convergence.[35]
Turning now to the assessment of the literature, my results confirm the ones of Maffezzoli (2006), they partially agree with Di Liberto *et al.*(2008), while they do not support the thesis of Bianchi and Menegatti (2005).

[33]See Figure 1 and Section 3.2 for further details.

[34]That a relatively slow growth rate acts as a disequalizing force has been clearly documented by Bourguignon and Morrisson (2002).

[35]See Lall (2001) for a complete treatment of these points.

In particular, the starting point of Maffezzoli (2006) is represented by a similar picture as my Figure 2, which is taken as evidence of cross-regional convergence tendencies, in the period considered, 1980-2004.[36] For disentangling convergence inner drivers, namely technical change and capital accumulation, he employs Data Envelope Analysis (DEA), which it might be thought as a refinement of traditional growth accounting.[37]
His analysis, like mine, confirms both technological β and σ convergence, in the sense that convergence in relative TFP (i.e. technological catch-up) seems to drive both the result of the negative correlation between initial labor productivity, and its subsequent growth rate, and the one of lower dispersion in cross-regional labor productivity distribution.
For the sake of completeness, however, it is important to underline that this study considers the absolute convergence prediction and not the conditional one. In fact, by the way in which the DEA decomposition has been constructed, structural factors have not been taken into account.[38] To check the potential of their relative importance for convergence analysis, I applied the distribution dynamics framework to labor productivity distributions conditioned only for capital stock and TFPgap, calling them respectively absolute neoclassical and absolute technological convergence. Interestingly, the results, reported in Figures 8 and 9, show club convergence dynamics, which is particularly strong in the case of neoclassical convergence. So, it could be said that for the convergence result, some pre-conditions have to be met. A part from technological development level, human capital deepening and crime reduction are of fundamental importance for cross-regional convergence. It is not surprising, in fact, that almost all of the studies about the Italian case did find supportive evidence for *conditional* and not *absolute* convergence. Although

[36]Figure 2 reports the negative correlation between initial labor productivity and its subsequent growth rate (i.e. absolute β convergence) and the decreasing dispersion in cross-regional labor productivity distribution (i.e. σ convergence).

[37]In particular, he estimates the production possibility frontiers, in 1980 and 2004, where each region pertains to the production possibility set. Moreover, to rule out the possibility of a technological regress, he construct the best practice frontier in year 2004 using all data points for both 1980 and 2004. He then decomposes regional labor productivity growth rate in the period, according to the Fisher ideal decomposition, in three parts: changes in the frontier position (technological change); changes in the distance from the frontier (efficiency changes) and, finally, movements along the frontier (capital deepening). His results are the following: first, technological change accounts for the greatest part of labor productivity improvements; second, cross-regional convergence in efficiency levels (i.e. backward regions experienced a relative faster pace into getting closer to the frontier, so that the technological gap is decreasing); third, a negative correlation between initial labor productivity and changes in the relative position with respect to the frontier and, finally, that convergence in TFP levels is responsible for the decrease in the dispersion of labor productivity levels. It is important to notice that a decreasing technological gap is found also by Leonida *et al.*(2004) employing the Malmquist productivity index to estimate TFP growth and DEA, in the 1970-1995 period.

[38]This is a general flaw of growth accounting exercises. For more details see the seminal contribution of Jorgenson and Griliches (1967).

this clarification is important, the bottom line does not change: technological catch-up will eventually lead to cross-regional convergence.

According to the work of Di Liberto *et al.*(2008), instead, both capital deepening and technological transfer have driven conditional convergence tendencies among Italian regions between 1963 and 1993. The proposed methodology for distinguishing between the two forces hinges upon the direct estimate of the unobserved heterogeneity, due to technological differences. More technically, they employ the Arellano and Bond (1991) Generalized Method of Moments estimator, as suggested by Caselli *et al.*(1996), refining the original work of Islam (1995) for direct TFP estimation.

One problem that this study might encounter is the so-called observational equivalence. That is, when employing parametric analysis, for 'isolating' the effects of technological improvements from the ones of capital deepening, it is necessary to control for capital stock initial conditions, even after having properly estimated -as they do indeed- regional TFP levels at different points of time. Put in another way, while their result on technological convergence is fully convincing, the one on neoclassical convergence is not, due to the fact that such mechanism is retrieved from the coefficient of lagged output, which might depend on either technical upgrading or capital deepening.

Finally, it is worth mentioning the study of Bianchi and Menegatti (2005). This is the only study which supports the neoclassical convergence hypothesis, instead of the technological one. The great merit of this work relies in its theoretical strength. In particular, the model employed allows for distinguishing the two convergence predictions. Although, the use of Ordinary Least Squares, dictated by scarce data availability, makes their estimates inconsistent, due to the well known problem of endogeneity of dynamic models.[39] So that, their conclusions are unquestionably weaken.

6 Conclusions

An unified distribution dynamics framework has been employed to test alternative hypotheses of convergence across Italian regions, between 1980 and 2003. It has been found supportive evidence for only technological conditional convergence. This, in turns, has two main implications. First: some pre-conditions have to be met for labor productivity equalization in the long run. In particular, as the pervasiveness of organized crime has been considered among the conditioning factors, institutional strengthening should have priority in policy makers' agenda. Second: the most important equalizing force is technological transfer from Northern to Southern regions. So that, providing sound incentives for technological upgrading, research and development and higher education seem to be

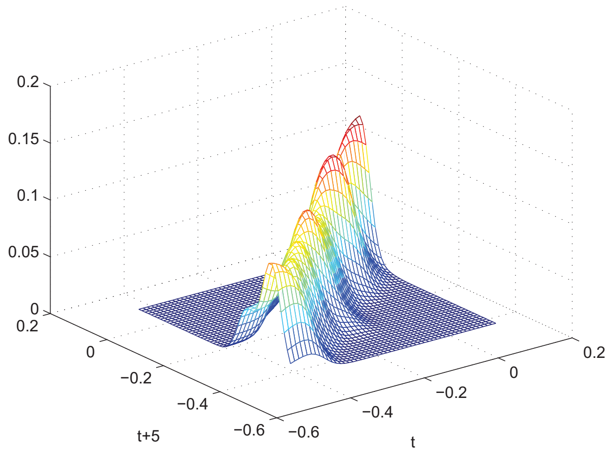
[39]See Caselli *et al.*(1996) for a clear explanation of this point in the context of growth regressions.

advisable.
To conclude, it is interesting to note that this investigation opens further lines of research. The most promising one seems to explain the persistence of cross-regional labor productivity differences in Italy and the increasing bimodality of its (unconditional) distribution. In other words, the next research questions will be: what will happen to Southern regions if the necessary convergence preconditions are not going to be met? What if technological upgrading will not succeed? Will the they remain relatively poor forever? Or, more technically, are Southern Italian regions stuck in a poverty trap?
Understanding whether the bimodality of the cross-regional labor productivity distribution is due to a self-reinforcing mechanism which cause poverty to persist (i.e. a poverty trap) or, instead, is due to the (non-unimodal) distribution of some exogenous factor is of fundamental importance for implementing sound economic policies. In fact, in the first case, the most advisable path to follow seems to be "a big push strategy", while, in the second, it is more appropriate to smooth regional differences.

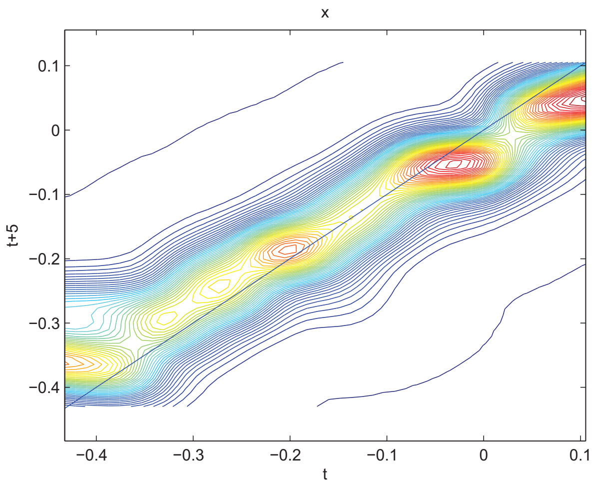
Logarithmic scale $y_{it} = log\left(\frac{Y_{it}}{Y_{Lt}}\right)$	**Relative Labour Productivity % with respect to Lombardy**
-0.6	54%
-0.4	67%
-0.3	74%
-0.2	81%
-0.1	90%
-0.05	95%
0	**100%**
0.05	105%
0.1	110%
0.15	116%
0.2	122%

Table 1: Graphs Scale

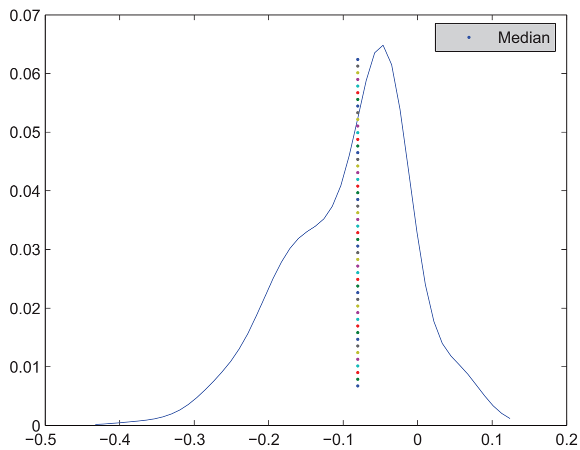
Figure 4: GDP per worker Absolute Convergence



(a) Kernel density

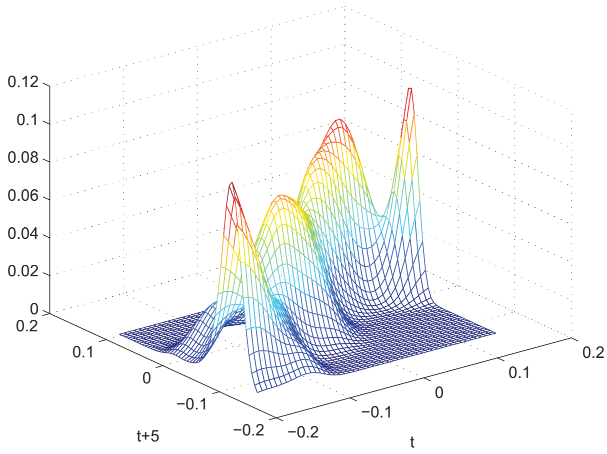


(b) Contour plot

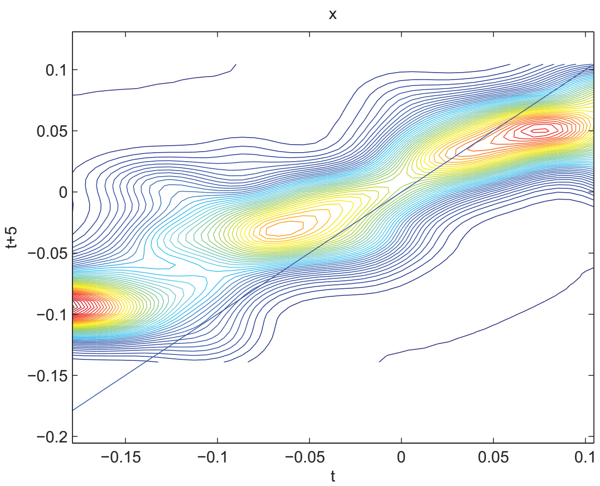


(c) Ergodic distribution

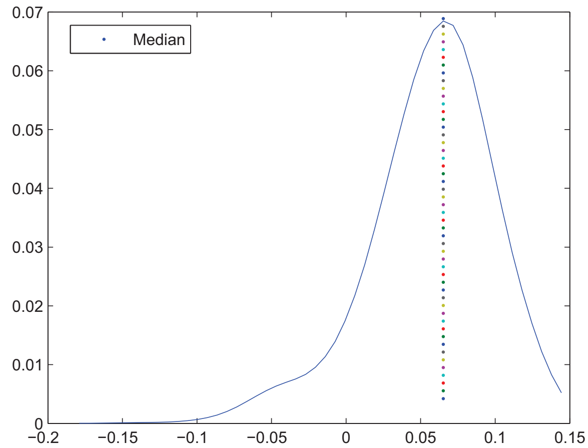
Figure 5: GDP per worker Conditional Convergence



(a) Kernel density

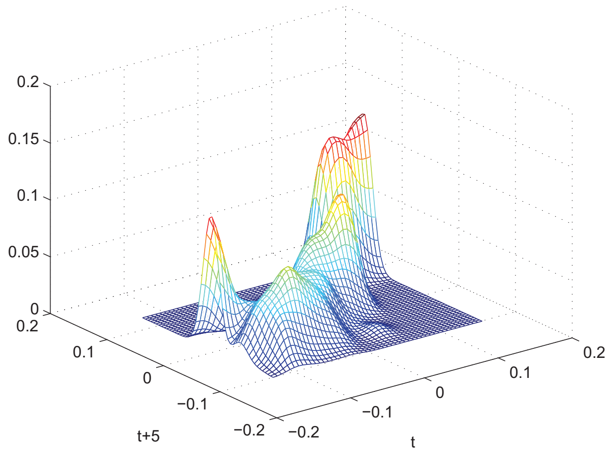


(b) Contour plot

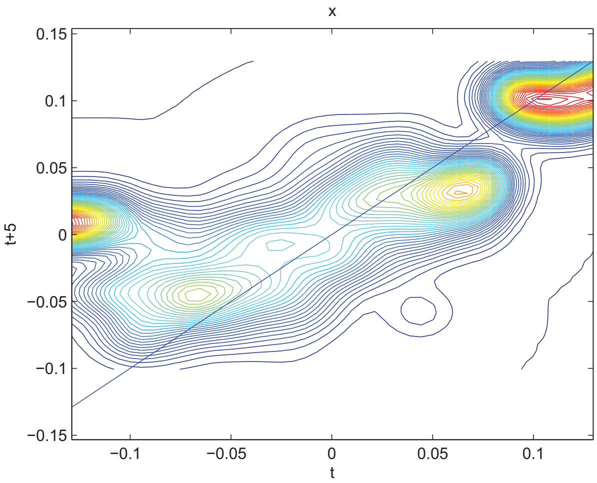


(c) Ergodic distribution

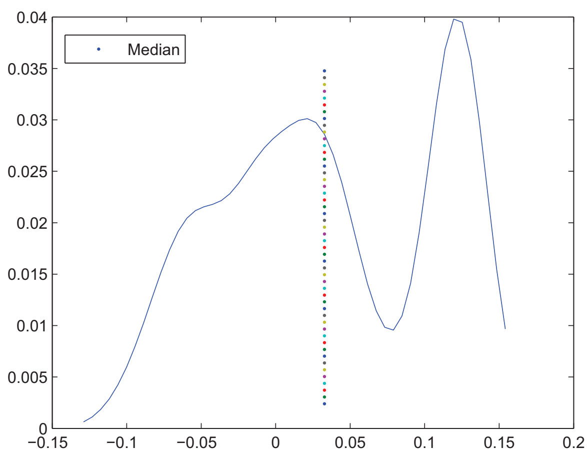
Figure 6: GDP per worker Neoclassical Conditional Convergence



(a) Kernel density

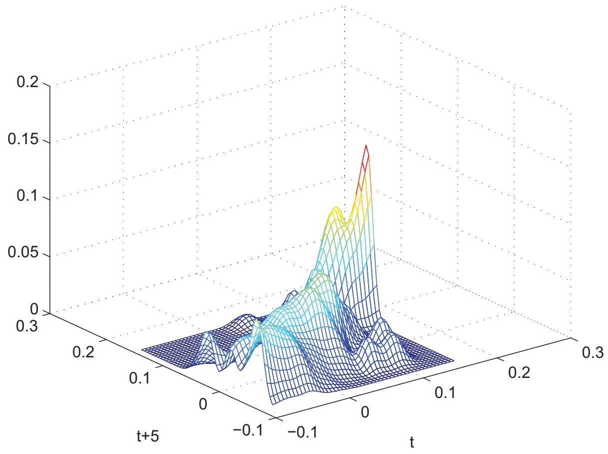


(b) Contour plot

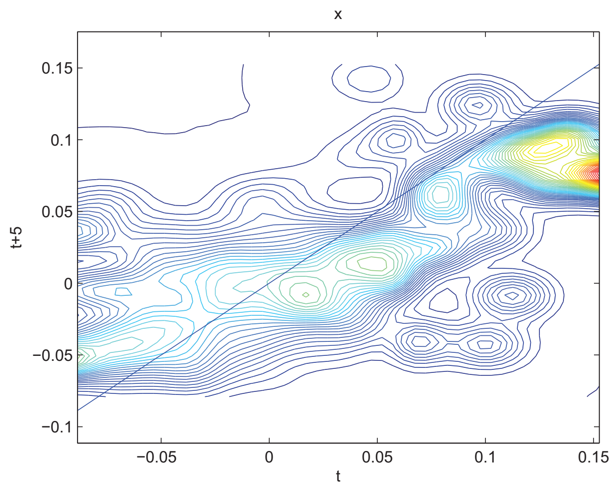


(c) Ergodic distribution

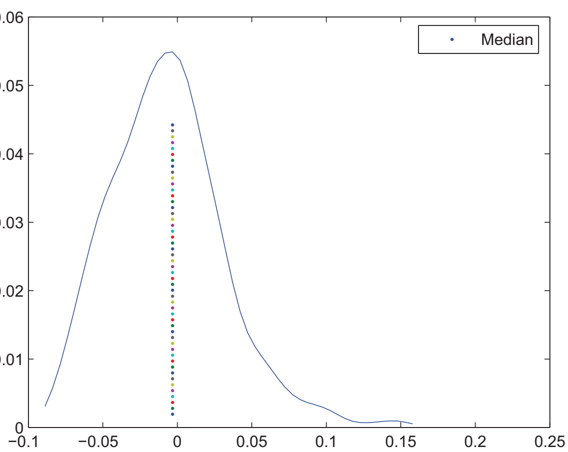
Figure 7: GDP per worker Technological Conditional Convergence



(a) Kernel density

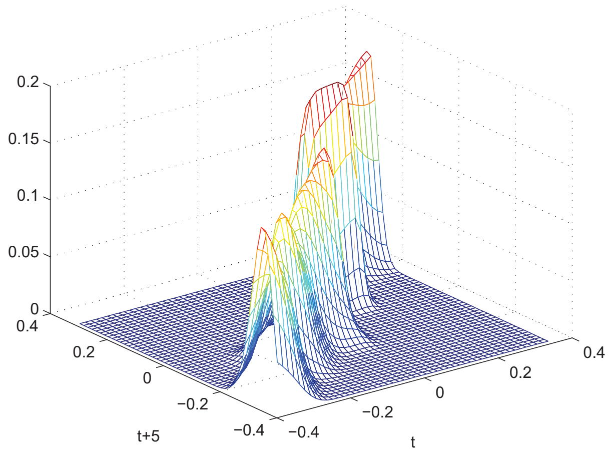


(b) Contour plot

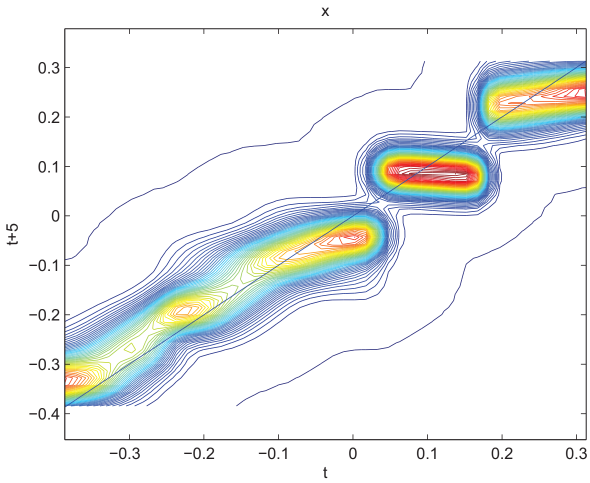


(c) Ergodic distribution

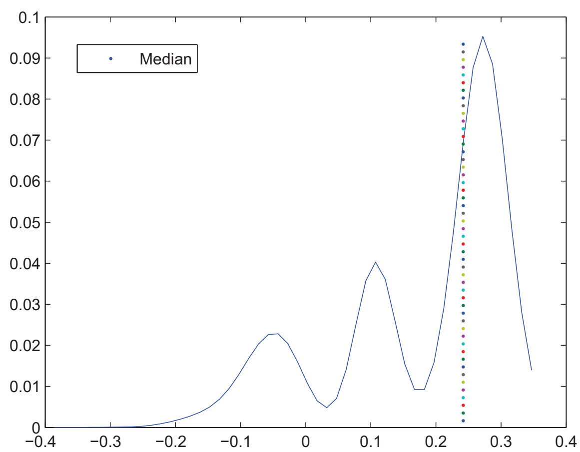
Figure 8: GDP per worker Absolute-Neoclassical Convergence



(a) Kernel density

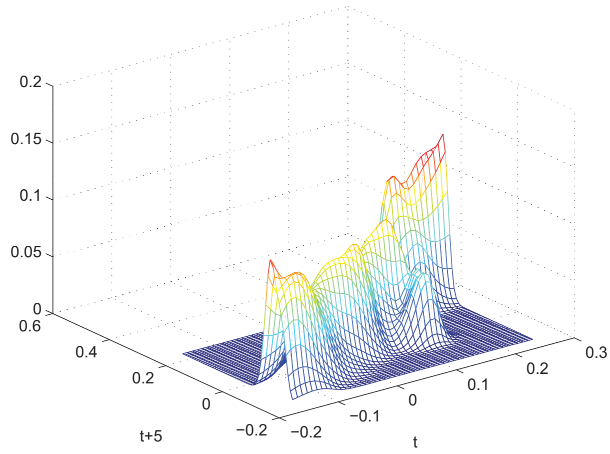


(b) Contour plot

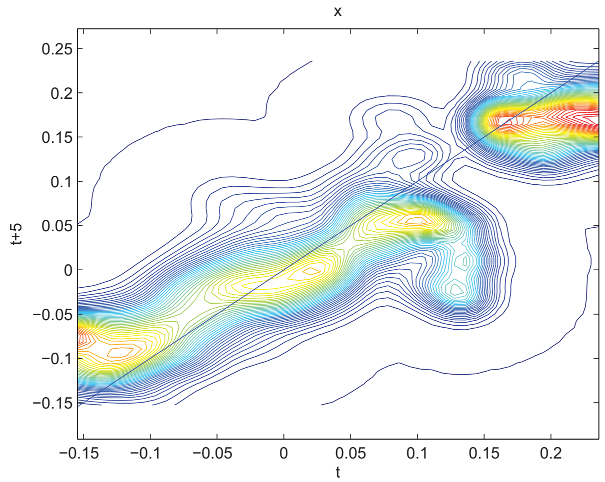


(c) Ergodic distribution

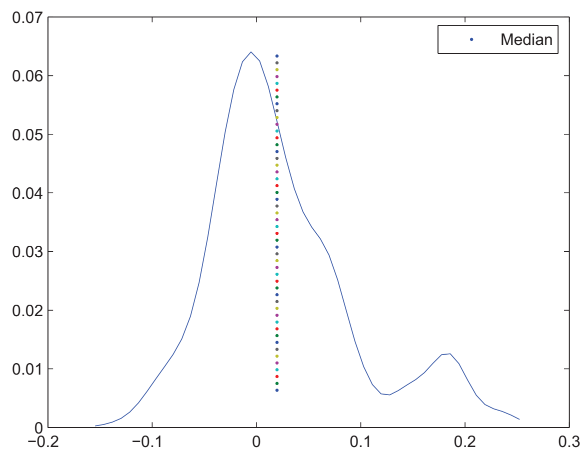
Figure 9: GDP per worker Absolute-Technological Convergence



(a) Kernel density



(b) Contour plot



(c) Ergodic distribution

Appendix

Following Diewert (1976), Caves et al. (1982b) derives an index number that allows TFP comparisons among countries. This index is *superlative*, meaning that is exact for the flexible aggregator function chosen (i.e. translog production function); and *transitive*, so that the choice of base country and year is inconsequential.[40]

Formally, I assume that value added of a generic region i, at time t, is a function of capital stock and employment; that is translog with identical second-order term; that constant returns to scale apply and that inputs are measured perfectly and in the same units for each observation. In symbols:

$$\ln y_{it} = \alpha_0 + \alpha_1 lnl_{it} + \alpha_2 lnk_{it} + \alpha_3 (lnl_{it})^2 + \alpha_4 (lnk_{it})^2 + \alpha_5 (lnl_{it} * lnk_{it})$$

Where constant returns to scale hypothesis requires $\alpha_1 + \alpha_2 = 1$ and $2\alpha_3 + \alpha_5 = 2\alpha_4 + \alpha_5 = 0$.

I review Caves et al.(1982) contribution, beginning with TFP index number for bilateral comparisons.

There are two economies, b and c; b is the basis of comparison and the distance function $D_c(y_b, l_b, k_b)$ represents the minimum proportional decrease in y_b such that the resulting output is producible with the inputs and productivity levels of c. Or, $D_c(y_b, l_b, k_b)$ is the smallest input bundle capable of producing y_b using the technology of c. In symbols:

$$D_c(y_b, x_b) = \min \{\delta \in \Re_+ : f_c(\delta x_b) \geq y_b\}$$

where $x_b = (k_b, l_b)$.[41] Assuming that producers are cost-minimisers and price takers in input markets, it can be shown that the Malmquist index (i.e. the geometric mean) of two distance functions for any two countries c and b gives the following TFP index:

$$TFP_{cb} = \frac{y_c}{y_b} \left(\frac{\bar{l}}{l_c}\right)^{\sigma_c} \left(\frac{\bar{k}}{k_c}\right)^{1-\sigma_c} \left(\frac{l_b}{\bar{l}}\right)^{\sigma_b} \left(\frac{k_b}{\bar{k}}\right)^{1-\sigma_b}$$

where a bar denotes an average over countries and $\sigma_i = (\alpha_i + \overline{\alpha})/2$, where (α_i) stands for labor's share in total costs for region i.

Similar reasoning can be applied to derive the multilateral version of TFP index, that

[40] *Exact* literally means that the resulting index is not an approximation. For details see Diewert (1976) and its result on the use of Tornqvist-Theil approximation to the Divisia index. *Flexible* is an aggregator function that can provide a second order approximation to an arbitrary twice differentiable linearly homogeneous function.

[41] This notation implies that only one homogeneous output is produced using only one homogeneous input. For further details on productivity measurement in this simple and more complex environments (i.e. multiple output-multiple input technologies), see Diewert (1992).

allows for TFP comparisons among more than two regions. Then, TFP level in economy i at time t is:

$$TFP_{it} = \frac{Y_{it}}{\overline{Y_t}}\left(\frac{\overline{L_t}}{L_{it}}\right)^{\tilde{\sigma}_{it}}\left(\frac{\overline{K_t}}{K_{it}}\right)^{1\tilde{-}\sigma_{it}}$$

where a bar denotes the geometric average over all regions in a given year t and $\tilde{\sigma}_{it} = (\alpha_{it} + \overline{\alpha})/2$, where α_{it} is labor share in region i and $\overline{\alpha}$ is the cross-region average.
Then, taking natural logarithms, the previous expression becomes:

$$TFP_{it} = ln\left(\frac{Y_{it}}{\overline{Y_t}}\right) - \tilde{\sigma}_{it} ln\left(\frac{L_{it}}{\overline{L_{it}}}\right) - (1-\tilde{\sigma}_{it}) ln\left(\frac{K_{it}}{\overline{K_t}}\right)$$

As originally noticed by Harrigan (1997), the variability in *actual* labor shares over value added makes difficult the empirical implementation of Equation (3). To solve this problem *smoothed* and not *actual* labor shares are usually employed.
Smoothed labor shares are simply obtained running a regression of actual labor shares on a constant and the capital to labor ratio:[42]

$$\alpha_{it} = \xi_i + \chi_i ln\left(K_{it}/L_{it}\right)$$

where ξ_i is a time invariant but region specific effect and χ_i is the region specific slope. Previous studies, such as Harrigan (1997,1999) and Griffith *et al.* (2004), considering only developed countries, allow only for slopes' heterogeneity (i.e.χ_i). As I work with regions that have reached different stages of economic development, I improved this sort of specification, considering regional heterogeneity in both intercepts and slopes, ξ_i and χ_i. In particular, to avoid a major loss in data variability, due to many dummies, I grouped Italian regions into three Macro-aggregates (i.e. North, Center and South). The diagnostics employed strongly reject the null hypothesis of non-heterogeneity in both intercepts and slopes among different regions. More precisely, using panel data F-tests, I have detected, separately, intercept heterogeneity. Through Chow type F-statistics, I have tested for both slope and intercepts' heterogeneity.

[42]This reduced form directly comes from the translog production function with constant returns to scale hypothesis.

References

Abramovitz, M. (1986): "Catching-up, forging ahead, and falling behind," *Journal of Economic History*, XLVI, 385–406.

Aiello, F., and V. Scoppa (2007): "Convergence and Regional Productivity Divide in Italy: Evidence from Panel Data," *Mimeo, University of Calabria*.

Alesina, A., and F. Giavazzi (2006): *Goodbye Europa*. Rizzoli.

Arellano, M., and S. Bond (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, 58, 277–297.

Arellano, M., and O. Bover (1995): "Another look at the instrumental-variable estimation of error-components models," *Journal of Econometrics*, 68, 29–52.

Ascari, G., and V. Di Cosmo (2005): "Determinants of Total Factor Productivity in the Italian Regions," *Italian Journal of Regional Science*, 4, 27–49.

Barro, R. J., and X. Sala-i Martin (1995): *Economic growth*. The MIT Press, Cambridge, Massachusetts.

Baumol, W. (1986): "Productivity growth, convergence and welfare: what the long run data show," *American Economic Review*, 76(5), 1072–85.

Bernard, A., and S. N. Durlauf (1995): "Convergence in International Output," *Journal of Applied Econometrics*, 10, 97–108.

Bianchi, C., and M. Menegatti (1997): "Differenziali regionali di produttivit e crescita economica: un riesame della convergenza in Italia nel periodo 1970-1994," *Studi Economici*, 63, 15–42.

——— (2005): "Neoclassical versus Technological Convergence: An Empirical Analysis Applied to the Italian Regions," *Quaderni di Dipartimento, Universita' degli Studi di Pavia*, 172 (01-05).

Blundell, R., and S. Bond (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87, 115–143.

Bonaglia, F., and L. Picci (2000): "Il capitale nelle Regioni Italiane," *DSE Working Paper 374, Universit di Bologna*.

Bourguignon, F., and C. Morrisson (2002): "Inequality among World Citizens: 1820-1992," *American Economic Review*, 92 (4), 727–744.

Carmeci, G., and L. Mauro (2002): "The convergence of the Italian regions and unemployment: theory and evidence," *Journal of Regional Science*, 42(3), 509–532.

Caruso, R. (2008): "Spesa Pubblica e Criminalit organizzata in Italia, evidenza empirica su dati Panel nel periodo 1997- 2003.," *Economia e Lavoro*, Forthcoming.

Caselli, F., G. Esquivel, and L. Fernando (1996): "Reopening the convergence debate: a new look at cross-country growth empirics," *Journal of Economic Growth*, 1, 363–389.

Caves, D. (1982a): "Multilateral comparisons of output, input and productivity using superlative index numbers," *Economic Journal*, 99, 73–86.

Caves, D., L. Christensen, and E. Diewert (1982b): "The economic theory of index numbers and the measurement of input, output and productivity," *Econometrica*, 50(6), 1393–1414.

CENSIS (2007): *XLI Rapporto Annuale sulla Situazione del Paese*. Franco Angeli.

Chamberlain, G. (1984): "Panel data," *Hanbook of Econometrics*, 2, 1247–1318.

Dal Bianco, S. (2007): "Alternative hypotheses of cross-country convergence. A nonparametric analysis of manufacturing sectors.," *Quaderni di Dipartimento, Universita' degli Studi di Pavia*, 196.

Del Monte, A., and E. Papagni (1998): "The determinants of corruption in Italy: regional panel data analysis," *European Journal of Political Economy*, 23, 379–396.

Desmet, K., and M. Fafchamps (2006): "Employment concentration across U.S. counties," *Regional Science and Urban Economics*, Forthcoming.

Di Liberto, A. (1994): "Convergence across Italian Regions," *Fondazione ENI Enrico Mattei, W.P. n.6894*.

Di Liberto, A., F. Pigliaru, and R. Mura (2008): "How to measure the unobservable: a panel data technique for the analysis of TFP convergence," *Oxford Economic Papers*, 60(2), 343–368.

Diewert, E. (1976): "Exact and superlative index numbers," *Journal of Econometrics*, 4(2), 115–145.

——— (1992): "Fisher ideal output, input, and productivity Indexes revisited," *Journal of Productivity Analysis*, 3, 211–48.

Dowrick, S., and M. Rogers (2002): "Classical and technological convergence: beyond the Solow model," *Oxford Economic Papers*, 54(3), 369–385.

DURLAUF, S. N., P. A. JOHNSON, AND J. R. TEMPLE (2005): "Growth Econometrics," *Handbook of Economic Growth. P. Aghion and S. N. Durlauf (eds.), North-Holland: Amsterdam*, 1A, 555–677.

EPSTEIN, P., P. HOWLETT, AND M.-S. SCHULZE (2003): "Trade, convergence and globalisation: the dynamics of the. international income distribution, 1950-1998," London School of Economics, STIRCED.

EUROSTAT (2005): "Regions:Statistical Yearbook 2005. Data 1999-2003," Discussion paper, European Commission.

EVANS, P., AND G. KARRAS (1996): "Convergence Revisited," *Journal of Monetary Economics*, 37, 249–265.

FAGERBERG, J. (1988): "Why growth rates differ," in *Technical change and economic theory.* G.Dosi et al.(eds), London, Pinter Publishers.

FRIEDMAN, M. (1992): "Do Old Fallacies Ever Die?," *Journal of Economic Literature*, 30, 2129–2132.

GAMBETTA, D., AND P. REUTER (1995): *The Economics of Organised Crime*chap. Conspiracy among the many: the mafia in legitimate industries. Cambridge University Press.

GERSCHNKRON, A. (1954): *Economic backwardeness in historical perspective*. Belknap Press.

GRIFFITH, R., S. REDDING, AND J. VAN REENEN (2004): "Mapping the two faces of R&D: productivity growth in a panel of OECD industries," *The Review of Economics and Statistics*, 87, 475–494.

HARRIGAN, J. (1997): "Technology, factor supplies and international specialisation," *American Economic Review*, 87, 475–494.

ISLAM, N. (1995): "Growth empirics: a panel data approach," *Quarterly Journal of Economics*, 110, 1127–1170.

ISLAM, N. (2003): "What have we learned from the convergence debate?," *Journal of Economic Surveys*, 17, 309–362.

JORGENSON, D. W., AND Z. GRILICHES (1967): "The explanation of productivity change," *Review of Economic Studies*, 34, 349–383.

LABMIM (2006): "Innovare Milano: Agenzia Nazionale dell'Innovazione e Palazzo dell'Innovazione," Discussion paper, Camera di Commercio di Milano.

Lall, S. (2000): "The technological structure and performance of developing country manufactured exports, 1985-1998," *Oxford Development Studies*, 28 (3), 337–369.

——— (2001): *Competitiveness, technology and skills*. Edward Elgar, Cheltenham.

Lamo, A. (2000): "On convergence empirics: some evidence for Spanish regions," *Investigaciones Economicas, Fundacion SEPI*, 24(3), 681–707.

Lavezzi, A. M. (2008): "Economic Structure and Vulnerability to Organised Crime: Evidence from Sicily," *Global Crime*, 9(3), 198–220.

Leonida, L., C. Petraglia, and L. Murillo-Zamorano (2004): "Total factor productivity and the convergence hypothesis in the Italian regions," *Applied Economics*, 36, 2187–2193.

Luenberger, D. G. (1979): *Introduction to dynamic systems*. Wiley, New York.

Maffezzoli, M. (2006): "Convergence Across Italian Regions and the Role of Technological Catch-Up," *The B.E. Journal of Macroeconomics*, 6, Iss.1 (Topics).

Magrini, S. (2007): "Analysing convergence through the Distribution Dynamics Approach: Why and How?," *DSE Working Paper 13*.

Mankiw, G. N., D. Romer, and D. N. Weil (1992): "A contribution to the empirics of economic growth," *Quarterly Journal of Economics*, 429, 407–437.

Mauro, L., and E. Podrecca (1994): "The case of Italian Regions: Convergence or Dualism?," *Economic Notes*, 23, 447–468.

Nelson, R., and E. Phelps (1966): "Investment in humans, technological diffusion and economic growth," *American Economic Review*, 56(2), 69–75.

Paci, R., and A. Saba (1998): "The empirics of regional economic growth in Italy, 1951-1993," *Rivista Internazionale di Scienze Economiche e Commerciali*, 45, 515–542.

Quah, D. (1993): "Galton's fallacy and tests of the convergence hypothesis," *The Scandinavian Journal of Economics*, 95(4), 427–443.

——— (1996): "Convergence empirics across economies with (some) capital mobility," *Journal of Economic Growth*, 1, 95–124.

——— (1996a): "Empirics for Economic Growth and Convergence," *European Economic Review*, 40(6), 1353–1375.

——— (1997): "Empirics for growth and distribution: Stratification, polarisation and convergence clubs. Journal of Economic Growth," *Journal of Economic Growth*, 2(1), 27–59.

Rogers, M. (2003): *Knowledge, technological catch-up and economic growth.* Edgward Elgar.

Sala-i Martin, X. (1996): "The Classical Approach to Convergence Analysis," *Economic Journal*, 106, 1019–1036.

Solow, R. (1956): "A contribution to the theory of economic growth.," *Quarterly Journal of Economics*, 70(1), 65–94.

——— (1957): "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, 39(3), 312–320.

Stockey, N. L., R. E. Lucas, and P. E. C (1989): *Recursive methods in economic dynamics.* Harvard University Press, Cambridge MA.

Tullio, G., and S. Quarella (1999): "Convergenza economica tra le regioni italiane: il ruolo della criminalit e della spesa pubblica 1960-1993.," *Rivista di Politica Economica*, 89, 77–128.

Wong, W.-K. (2007): "Economic Growth: A Channel Decomposition Exercise," *The B.E. Journal of Macroeconomics*, 7(1).