Economic complexity and regional labor productivity distribution: evidence from Italy

Roberto Basile[*] Gloria Cicerone[†] Lelio Iapadre[‡]

February 26, 2019

Abstract

In this paper we analyze the role of economic complexity as a driver of regional labor productivity dynamics in Italy. Economic complexity is expressed in the composition of an economy's productive structure and reflects the emerging combination of the multiplicity of knowledge embedded in it. Measures of economic complexity (ECI) come from the structure of the network connecting economies to the products they export. We assess the impact of ECI on the distribution dynamics of labor productivity by combining growth regression analysis with conditional density estimates. Counterfactual analysis results suggest that ECI plays a key role in the observed tendency to polarization of regional labor productivity.

Keywords: Productivity; Economic complexity; Distribution dynamics.
Jel codes: R11, R12

[*]Department of Industrial and Information Engineering and Economics, University of L'Aquila, Italy, and Luiss. roberto.basile@univaq.it

[†]GSSI, L'Aquila. gloria.cicerone@gssi.it

[‡]Department of Industrial and Information Engineering and Economics, University of L'Aquila, Italy. lelio.iapadre@univaq.it

1 Introduction

In this paper we analyze the role of economic complexity as a driver of regional labor productivity dynamics in Italy. Economic complexity is expressed in the composition of an economy's productive structure and reflects the emerging combination of the multiplicity of knowledge embedded in it (Hidalgo and Hausmann, 2009). Measures of economic complexity (ECI) come from the structure of the network connecting countries/regions to the products they export[1]. The main argument is that the positioning of an economy's tradeable sectors within the global trade network is crucial for its economic growth trajectory (Hausmann and Klinger, 2006; Hausmann, Hwang, and Rodrik, 2007; Hidalgo, Klinger, Barabási, and Hausmann, 2007; Hidalgo and Hausmann, 2009).

During the last 30 years, models of economic growth have often included the assumption that the variety of inputs used by an economy affects its overall productivity (Aghion and Howitt, 1998; Grossman and Helpman, 1991; Howitt, 2000). However, the concept of variety does not include neither the specificity of productive inputs nor their complementarity, and therefore empirical research has not moved forwards along this line because of the absence of adequate measures of complexity of the productive structure.

Only recently interest in the relationship between economic complexity and growth has been revived by the seminal contributions of Hausmann and Rodrik (2003), Hausmann, Hwang, and Rodrik (2007) and Hidalgo and Hausmann (2009). They develop measures of economic complexity capturing information about the set of capabilities[2] available in an economy, jointly with the set of capabilities that a product requires. Hence, the productivity of an economy depends on the diversity of the set of capabilities it has locally available, and therefore, cross-country

[1]Economic complexity is measured by using international trade data. This choice is justified by the fact that, especially at subnational level, export and import data can be more finely disaggregated at sectoral level than value added or employment data, and is also theoretically supported by the observation that an economy's trade structure reflects quite well its production structure.

[2]The capability theory, which provides the logical underpinning of the economic complexity framework, abstracts capabilities as building blocks of production and economic development as the accumulation of those capabilities.

or cross-regional differences in labor productivity can be explained by differences in economic complexity, as measured by the diversity of capabilities present in a country/region and by the sophistication ("ubiquity") of the products that this economy exports. In a nutshell, international trade openness is expected to bring about both static and dynamic gains through different channels (Frankel and Romer, 1999): *i*) specialization in production according to the economy's comparative advantages (*specialization effect*); *ii*) exploitation of scale economies both in production processes and in R&D activities (*market size effect*); *iii*) international transmission of know-how and dissemination of technological progress (*spillover effect*); *iv*) firms' efficiency increase forced by the exposure to international competition (*competition effect*). The level of trade openness measures, however, does not allow us to discriminate between an economy producing and exporting a large variety of goods and another one highly specialized in a few sectors, or between an economy specialized in traditional sectors and another one exporting sophisticated goods whose production requires the combination of more complex competences.

A recent strand of research, much more closely associated with agglomeration literature, has shed light on the role that the quality of trade patterns plays in economic growth. This literature moves from the debate focused on whether specialization (localization externalities) or diversification (Jacobs's externalities) induce knowledge spillovers and therefore local economic growth (Glaeser, Kallal, Scheinkman, and Shleifer, 1992; Feldman and Audretsch, 1999; Van Oort, 2017). What seems to be crucial is the kind of inter-sectoral linkages emerging from the structure of local economies. The driving idea is that knowledge is more likely to spill from one sector over another when a certain degree of cognitive proximity ensure interactive learning and innovation to take place (Nooteboom, 2000; Boschma, 2005). In other words, the inter-industry transmission of knowledge easily occurs when sectors are related in terms of shared or complementary competences. Therefore an economy characterized by a relatively large presence of related sectors - sectors that are related in terms of shared or complementary competences - performs better than specialized economies, as well as than economies diversified in unrelated sectors. Measures of related variety are based on the concept of entropy (Theil, 1972), and have been applied to the study of trade patterns of Italian provinces (Boschma and Iammarino, 2009) and Spanish regions (Boschma, Frenken, Bathelt, Feldman, Kogler, et al., 2012). These studies have demonstrated that the most promising pathways for a region's growth are found by diversifying its economic structure into sectors which are closely related to the existing dominant technologies (McCann, 2013). From a policy point of view, the related-variety approach allows for the identification of key, growth-promoting sectors of the economy. However, related variety, in itself, is not a reliable predictor of long-run productivity growth. For any given level of related variety, different growth paths can emerge, depending on specific features of the interconnected sectors characterising the structures of different economies. More precisely, what matters is the cognitive sophistication of the products that an economy makes (and exports), which reflects its underlying capabilities. In this sense, the notion of economic complexity - and its operational measure (ECI) - represents an important improvement, since it considers both the degree of variety and the sophistication of the productive structure.

In our empirical analysis we use data on Italian provinces over the period 2000-2015. We aim at assessing the long-run impact of ECI on the dynamics of labor productivity distribution. Our approach combines growth regression analysis with the estimation of conditional density functions. Specifically, we start using a continuous state-space approach for the estimation of the intra-distributional dynamics of labor productivity, and compute the ergodic distribution to identify the long-run properties of the observed distribution dynamics. The most significant stylized fact emerging on the evolution of the cross-province labor productivity distribution is the shift from unimodality in 2000 to bimodality in 2010 (i.e. after the great economic crisis). Evidence from the ergodic distribution suggests that this polarization process continues in the long run. Then, to evaluate the distributional impact of economic complexity, we use a generalized method of counterfactual analysis, based on the comparison between actual and counterfactual ergodic distributions to identify long-run tendencies. The counterfactuals are based on the cross-sectional heterogeneity of growth determinants, i.e. on how the ergodic distribution would have looked like if there were no heterogeneity in economic complexity across provinces. Counterfactual analysis results suggest that ECI plays a key role in the observed tendency to polarization of regional labor productivity in Italy.

The rest of the paper is structured as follows. In Section 2 we discuss the methodology used to measure economic complexity according to the Hidalgo and Hausmann's approach, and then we present some evidence on the provincial distribution of ECI and other relevant variables in Italy. Section 3 focuses on the counterfactual analysis described above to assess the role of ECI in determining the shape of the ergodic distribution of labor productivity. Section 4 concludes.

2 Measuring Economic Complexity

The Economic Complexity Index (ECI) measures the diversity and sophistication of the productive structure of a country or region, and thus it reflects the emerging combination of the multiplicity of knowledge embedded in it (Hidalgo and Hausmann, 2009; Hausmann, Hidalgo, Bustos, Coscia, Simoes, and Yildirim, 2014).

Here, we compute ECI following Hidalgo and Hausmann (2009)'s approach, and using export data for the NUTS3 Italian regions (provinces) for the period 1995-2005.[3] We define $RCA_{i,c}$ as a matrix of the Revealed Comparative Advantages (RCA) showing the specialization pattern of each province (i.e. the products they export more than what we expect based on a province's total exports and a product's global market). We use Balassa (1965)'s definition of RCA as the ratio between the export share of product i in province c and the share of product i in the world market. Formally, RCA_{ic} is defined as:

$$RCA_{ic} = \frac{\frac{X_{ic}}{\sum_i X_{ic}}}{\frac{\sum_c X_{ic}}{\sum_{ic} X_{ic}}} \tag{1}$$

where $X_{i,c}$ denotes export values of province c in product i. This represents the first step to

[3]These data are drawn from ISTAT Coeweb Section, disaggregated according to the ATECO Classification at the three-digit level, providing the regional value exported to the world for 118 product classes for each Italian province. All of the export sectors in our provincial trade dataset are manufacturing sectors, which in 2016 accounted for almost 97% of Italy's total exports - net of refined petroleum products and coke oven products.

build the bipartite network in which provinces are connected to the products they export. A bipartite network (or graph) is a set of nodes and links in which nodes can be separated into two groups, or partitions, such that links only connect nodes in different partitions. Mathematically, we represent this network using the adjacency matrix $M_{i,c}$, where $M_{i,c}$=1 if RCA is larger than a certain RCA* threshold – in other words, we set $M_{i,c} = 1$ if province c is a significant exporter of product i and 0 otherwise. Finally, we make Y_c equal to the number of connections, or degree $k_{c,0}$, that province c has in the network. $k_{c,0}$ is therefore a measure that comes only from the structure of the network. Formally these transformations are:

$$\begin{aligned} M_{ic} &= 1 \ if \ RCA \geq RCA^* \\ Y_c &= k_{c,0} \end{aligned} \tag{2}$$

where $k_{c,0}$ is given by:

$$k_{c,0} = \sum_i M_{i,c} \tag{3}$$

and represents the *diversification* of province c (the number of products in which that province reveals a comparative advantage). Additionally, Hidalgo and Hausmann (2009) define the *ubiquity* of a product in the bipartite network as:

$$k_{i,0} = \sum_c M_{i,c} \tag{4}$$

that is the number of provinces that are specialised in product i. Measures of knowledge complexity for both provinces and products can be found by sequentially combining these measures of diversity and ubiquity in the following two equations over a series of n iterations:

$$k_{i,n} = \frac{1}{k_{i,0}} \sum_c M_{i,c} k_{c,n-1} \tag{5}$$

$$k_{c,n} = \frac{1}{k_{c,0}} \sum_i M_{i,c} k_{i,n-1} \tag{6}$$

This is the so-called Method of Reflections. It allows us to extract relevant information about the availability of capabilities in the Italian provinces.

ECI is $k_{c,n}$ with n going to infinity, each additional iteration (or reflections) in $k_{c,n}$ provides increasingly more relevant information about the knowledge complexity of a region. Indeed, although higher-order iterations in this technique become progressively more difficult to define, the method of reflections provides more and more precise measures of the ECI as the noise and size effects are eliminated. The iterations are stopped when the ranking of regions and products is stable from one step to another (i.e., no further information can be extracted from the structure of the region-product network). For practical purposes, Hidalgo suggests to take $n \geq 12$ as being large enough. We stop our iteration process exactly at $n = 12$ as all of the ECI provincial values converge at the same level. Following Hidalgo and Hausmann (2009), we extract information from the tiny deviations of these converging values.

In order to construct the Economic Complexity Index, Hidalgo and Hausmann (2009) assume that each product requires a varied and potentially large set of different complementary non-tradable inputs, which they call capabilities. Economies differ in the capabilities that are present in their territory while products differ in the capabilities they require. More formally, the bi-partite network connecting provinces to the products in which they are specialized can be taken to be the result of the product of two other matrices or networks: a province-capability network that expresses the endowment of capabilities of each province and a capability-product matrix that contains the technological requirement of products (Hausmann and Hidalgo, 2011). As a consequence, provinces with more capabilities will be more diversified, and products that require more capabilities will be accessible to fewer provinces, and hence will be less ubiquitous. Also, provinces with more capabilities will be able to make products that require more capabilities, and are therefore less ubiquitous[4]. The reverse holds true for provinces with few capabilities.

Economic complexity theory predicts traps in the process of economic diversification.

[4]This logic explains the negative relationship between the diversification of provinces and the average ubiquity of the products in which they are specialized.

Economies with few capabilities will be able to make few products, and will have scant benefits from accumulating any individual additional capability. This is because the likelihood that a new capability will be able to synergize with existing capabilities and become useful for the production of a new product is low in the absence of the other requisite capabilities. Therefore, the demand for any randomly selected additional capability is likely to be zero in economies with few capabilities. By contrast, economies with many capabilities would be able to produce many new products by combining any new capability with different subsets of the capabilities they already possess. In other words, the model generates increasing returns in terms of diversification to the accumulation of capabilities. This suggests that the mix of products made by an economy increases gradually through the addition of capabilities, and that this gradual development leaves fingerprints in the structure of the network connecting economies to products.[5]

The incorporation of information on product ubiquity differentiates ECI from the traditional measures of diversification, such as the Herfindahl-Hirschman (HH) index or entropy-based measures. Neither the HH-index nor entropy include any information on products, making their measures of diversification identical for any two baskets of goods that have the same distribution of shares (i.e. they do not distinguish an economy exporting 10% bananas and 90% mangos from another one exporting 90% bananas and 10% mangos, or another exporting 10% motorcycles and 90% aircraft engines). In other words, ECI considers how many economies are specialized in a particular product, going beyond the analysis of economic growth focused just on the study of aggregate output (Hausmann and Hidalgo, 2011).

Insert Figure 1 about here

Figures 1 shows the time trends of Italian average levels of productivity, degree of openness, related variety and degree of economic complexity (ECI). The four variables present positive

[5]The results presented in Hidalgo, Klinger, Barabási, and Hausmann (2007) and in Hausmann and Klinger (2006) show that countries patterns of comparative advantage evolve by moving from existing goods to related goods in the Product Space. However, the ability to add a product to the production set of a country depends not only on how close a given product is to an already existing one, but also on how many other capabilities are present in the country and used in other, potentially more distant, products.

trends, with related variety and ECI trends not directly affected by the economic crisis.

Insert Figure 2 about here

Figure 2 depicts the spatial distribution of relative labor productivity growth, economic complexity index, openness and related variety for the Italian provinces. A North-South pattern emerges across Italy for all variables, with few provinces diverging from this pattern.

3 Economic complexity and labor productivity distribution: evidence for Italian provinces

3.1 Empirical strategy

The empirical strategy used to identify the impact of trade complexity (*ECI* as measured above) and of other competing growth determinants on the intra-distribution dynamics (*IDD*) of manufacturing labor productivity in Italy combines counterfactual analysis and conditional density estimates, in the spirit of recent empirical contributions to the growth analysis (Basile, 2009; Cheshire and Magrini, 2005; Fiaschi, Gianmoena, and Parenti, 2018, 2011). More specifically, this strategy consists of a three-step approach.

In the first step, we analyze the actual *IDD* using the well-known method, firstly proposed by Quah (1997), based on the estimation of conditional densities and ergodic distributions. In a nutshell, this method consists of studying the dynamics of the entire distribution of the level of labor productivity of a set of economies. In particular, if the process describing the evolution of the distribution is time-invariant and first-order Markov, then the relationship between the distribution of labor productivity at time $t+\tau$ and t can be written as

$$\phi_{t+\tau}(y) = \int_0^\infty f_\tau(y|x)\phi_t(x)dx \quad (7)$$

where $f_\tau(y|x)$ is the probability density function of y (labor productivity at time $t+\tau$) conditional upon x (labor productivity at time t). In other words, the conditional density $f_\tau(y|x)$ describes the probability that a given economy moves to a certain state of relative productivity (higher or lower) given that it has a certain relative productivity level in the initial period. In this case convergence must be studied by visualizing and interpreting the shape of the productivity distribution at time $t+\tau$ over the range of productivity levels observed at time t.

The corresponding ergodic distribution is the limit of (7) as t tends to infinity (Johnson, 2005):

$$\phi_\infty(y) = \int_0^\infty f_\tau(y|x)\phi_\infty(x)dx \tag{8}$$

This function describes the long-term behavior of the productivity distribution: it is the density of what the cross-economy productivity distribution tends towards, should the system continue along its historical path.

The second step of the empirical strategy consists of estimating a growth regression equation, including on the r.h.s. (in matrix G of equation 9) both the proximate determinants suggested by the neoclassical and the Schumpeterian growth literature (i.e. initial conditions, investment rates in physical and in human capital, average annual employment growth rates, and R&D investment rate) as well as determinants related to the trade behavior of the provinces (*in primis* the measure of economic complexity, but also a measure of openness, and a measure of trade variety):

$$\gamma = g(G) + \varepsilon \tag{9}$$

where $\gamma \equiv \dfrac{ln(y) - ln(x)}{\tau}$ is the average annual growth rate of labor productivity [i.e. $y \equiv x \times \exp(\gamma\tau)$], and ε is a stochastic error term. To keep the empirical strategy as general as possible, the function $g(G)$ is an additive function of smooth terms, possibly including nonlinear terms and interaction terms, in the spirit of Durlauf, Johnson, and Temple (2005, p. 560). In order to account for spatial spillover effects and for unobserved spatial autocorrelation, matrix G may also include spatial autoregressive terms as proposed, for example, by Ertur and Koch (2007, 2011), and the error term may be a spatial AR1 process. Moreover, in order to control for the endogeneity of the growth determinants, an instrumental variable (IV) method can be used to estimate equation (9).

The last step of the empirical strategy consists of using the results of the growth regression analysis to estimate counterfactual distributions with respect to individual variables of interest, in order to identify their distributional effect, that is their contribution to convergence, divergence and polarization. More precisely, let decompose matrix G in two parts: the first one refers to a specific variable G_k and the second one ($G_{-k} = G_1, ..., G_{k-1}, G_{k+1}, ..., G_K$) includes all other $(K-1)$ variables. Assuming that the values of variable G_k are all the same, for example they are all equal to the sample mean (i.e. $G_k = \overline{G_k}$), the simulated predicted growth rate is calculated as:

$$\widehat{\gamma_{\bar{k}}} \equiv \widehat{\gamma}(\overline{G_k}) + \widehat{\gamma}(G_{-k}) \tag{10}$$

and, thus, we can compute the counterfactual final (i.e. end of period) productivity level for the *k-th* variable as

$$\widehat{y_{\bar{k}}} \equiv x \times \exp(\widehat{\gamma_{\bar{k}}}\tau) \tag{11}$$

This variable simulates the levels of labor productivity that the economies would attain at the end of the period if there were no differences within the sample in the level of the *k-th* growth determinant. Finally, counterfactual univariate, conditional and ergodic densities are estimated using expressions (7) and (8) but replacing the actual variable y with the constructed variable $\widehat{y_{\bar{k}}}$. By comparing actual and counterfactual distributions, we can assess the distributional effect of variable k. We also quantify the distributional impact of variable k by using measures of inequality (Theil index) and polarization (BIPOL index, proposed by Anderson, Linton, and Leo, 2012).

3.2 Distribution dynamics of regional labor productivity in Italy

In this section we report the results of the analysis of the intra-distribution dynamics (*IDD*) of regional labor productivity differentials in the manufacturing sector (net of refined petroleum products and coke oven products) in Italy over the period 2000-2015. The spatial unit of the analysis is the province (NUTS3 level), and all data are drawn from ISTAT on line dataset (*dati.istat.it/*). Labor productivity is computed as the ratio between the gross value added (at current prices) and the number of employees. The following analysis is performed using relative measures of labor productivity, that is the ratio between provincial and national (weighted) average labor productivity in order to remove co-movements due to the national-wide business cycle and trends in the average values.

We start from the observation that the distribution of relative labor productivity across the 103 Italian provinces is characterized by a North-South pattern: high (low)-productivity provinces are concentrated in the Northern (Southern) part of the country (Figure 3). Except for a few cases, this spatial pattern appears quite stable over the sample period.

Insert Figure 3 about here

Nevertheless, the estimated univariate densities (Figure 4) and the statistics reported in Table 1 reveal important changes in the distribution of relative productivity levels over time. In particular, the degree of inequality (measured by the Theil index) increased between 2000 and 2010 from 0.020 to 0.033. Also the degree of polarization (measured by the BIPOL index) increased from zero in the first two periods (2000 and 2005) to values higher than 0.4 subsequently. Figure 4 (and more in details Figure 5) shows indeed that the distribution of labor productivity was initially unimodal (in 2000 and in 2005), but then it became bimodal (since 2010): the first mode in 2015 is at about 0.7 while second peak is at about the national average level.

Insert Figures 4 and 5 and Table 1 about here

Over the whole period 2000–2015, the intra-distribution dynamics has indeed generated long-run multiple equilibria with the formation of two clusters (or convergence clubs). The *IDD* is analyzed through the estimation of conditional densities using stochastic kernel estimators with variable bandwidth (Hyndman, Bashtannyk, and Grunwald, 1996; Hyndman, 1996; Basile, 2010).[6] In order to estimate conditional density functions $f_\tau(y|x)$, we consider the first year ($t = 2000$) and the last year ($t + \tau = 2015$) of the available data. Thus, y and x are vectors of 103 observations.

The results of the conditional density estimates are shown using the graphical methods developed by Hyndman (1996), and firstly used by Basile (2010) for the *IDD* analysis. The first graph (Figure 6, a) is the *stacked conditional density* plot, displaying the changes in the shape of the distributions of labor productivity observed in 2015 over the range of labor productivity observed in 2000. Thus, like a row of a transition matrix, each univariate density plot in Figure (6, a) describes transitions over the analyzed period from a given productivity value in 2000.

Insert Figure 6 about here

The second graph is the *highest conditional density region* (*HDR*) plot (Figure 6, b): each vertical band represents the projection on the *xy* plan of the conditional density of y on x. In each band the 25% (the darker-shaded region), 50%, 75% and 90% (the lighter-shaded region) *HDR* are reported.[7] These regions allow a visual summary of the characteristics of a probability distribution function. The *HDR* plot is particularly suitable to analyze *IDD*. If the 45-degree diagonal crosses the 25-50% *HDR*s, it means that most of the elements in the distribution remain where they started (there is '*persistence*'). If the horizontal line traced at the one-value of the vertical axis crosses *all* the 25-50% *HDR*s, we can say that there is '*global convergence*'. If the vertical line traced at the one-value of the horizontal axis crosses *all* the 25-50% *HDR*s, we can say that there is '*global divergence*'. Finally, the presence of nonlinearities in the modal regression functions (shown in the plot as bullets) can be interpreted as evidence in favor of the

[6] Conditional density functions have been estimated using the R software (library *hdrcde*). The two smoothing parameters (a and b) of the conditional kernel density estimator (controlling the smoothness between conditional densities in the x direction and the smoothness of each conditional density in the y direction, respectively) have been selected using a generalized cross validation method. The bandwidth a is a variable (or nearest-neighbor) bandwidth accounting for the heterogeneous distribution of the data over the sample space, while the bandwidth b is fixed.

[7] A high density region is the smallest region of the sample space containing a given probability.

'*club convergence*' hypothesis, according to which regions catch up with one another but only within particular sub-groups.

Figure (6, b) shows the existence of two convergence clubs: provinces sufficiently close to each other over the x-domain converge towards each other. Based on the simple visual inspection of Figure (6, b) we could classify regions in two clusters. However, to make a more accurate identification of the two clubs of regions, we also rely on the methodology proposed by Phillips and Sul (2007, 2009). The results of this analysis brought us to classify Italian provinces in the two convergence clubs as shown in Table 2 and in Figure 7 in green and red colors. The same colors have been used to show the observed values of labor productivity of the provinces in Figure (6, b). Using this classification, we then decompose the Theil index in the two components (within and between). Table 1 clearly shows that the increase in the overall level of productivity inequality from the period 2010 can be ascribed to the increase in the level of between clubs inequality, while the level of within club inequality remains stable over time. In a nutshell, the formation of the two clubs is at the origin of the enlargement of labor productivity differentials across provinces in the manufacturing sector in Italy.

Finally, the shape of the ergodic distribution (computed using the transition matrix extracted from the conditional density estimates) suggests that the dynamics of polarization tends to be persistent in the long run (Figure 8): the values of the BIPOL index slightly increases from 0.417 in 2015 to 0.467 in the ergodic distribution.

Insert Figure 8 about here

3.3 Growth analysis

In the second step of our empirical strategy we perform the growth regression analysis. The dependent variable in the growth regression equation is the annual average growth rate. The set of explanatory variables includes the proximate determinants suggested by the neoclassical (Mankiw, Romer, and Weil, 1992) and the Schumpeterian (Howitt, 2000; Ertur and Koch, 2011) growth literature (i.e. initial conditions, investment rates in physical and in human capital, average annual employment growth rates, and R&D investment rates), as well as determinants related to the international trade behavior of the provinces (namely, degree of economic complexity, openness degree, and degree of related variety).

As mentioned above, the functional form of the growth equation (9) is very general and allows for nonlinearities and spatial autoregressive terms. A linear model may indeed suffer from misspecification bias. However, preliminary results based on the estimation of semiparametric additive models do not reveal any significant evidence of nonlinearities. Moreover, the results from linear OLS estimates do not reveal any residual spatial autocorrelation.[8] All in all, these preliminary results suggest that the manufacturing productivity growth equation for Italian provinces can be specified, in matrix form, as a standard linear model such as:

$$\begin{aligned}\gamma &= \alpha\mathbf{1} + \beta_1 \ln \mathbf{y}_0 + \beta_2 \ln \mathbf{s}_k + \beta_3 \ln \mathbf{s}_h + \beta_4 \ln \mathbf{s}_A + \beta_5 \ln(\mathbf{n} + 0.05) + \\ &\quad \beta_6 \ln \mathbf{eci} + \beta_7 \ln \mathbf{open} + \beta_8 \mathbf{rv} + \varepsilon\end{aligned} \tag{12}$$

where $\mathbf{1}$ is the intercept term, $\ln \mathbf{y}_0$ is the log of the initial (i.e. in 2000) levels of value added per worker, whose effect proxies for technological catching-up and/or decreasing marginal productivity of capital; $\ln \mathbf{s}_k$ is the log of the average annual investment rate (investment to value added ratio); $\ln \mathbf{s}_h$ is the log of the average annual level of human capital, measured in terms of average years of schooling of the provincial population;[9] $\ln \mathbf{s}_A$ is the log of the average annual investment in R&D, measured in terms of R&D employment on total employment ratio; $\ln(\mathbf{n} + 0.05)$ is the log of the average annual growth rate of employment augmented by the depreciation rate and the exogenous rate of technological progress (equal to 0.03 and 0.02 respectively); $\ln \mathbf{eci}$ is the log of the average annual level of trade complexity;[10] $\ln \mathbf{open}$ is the log of average annual rate

[8] These preliminary results are available upon request.

[9] The average years of schooling Ays is measured as $Ays = \sum Sh_l \times D_h$, where Sh_l is the proportion of the population with a level of education l, and D_l is the official duration in years of the $l-th$ level. Specifically, we consider four education levels: 1) up to primary school; 2) lower secondary school; 3) upper secondary school; 4) tertiary education. The weighting scheme is based on the duration of the educational cycle, as suggested by the human capital literature (Barro and Lee, 1993; Psacharopoulos and Arriagada, 1986). In our case the duration of each level is 5, 8, 13, and 18 years, respectively.

[10] In order to compute the logarithm of **ECI** (measured as described in section 2), we first project it onto the [0,1]

of trade openness, measured as *(Export+Import)/Value added*; **rv** is the average annual level of export related variety measured as in Boschma and Iammarino (2009); ε is the error term, assumed to be independently and identically distributed. Finally, α and $\beta_1 - \beta_8$ are the parameters to be estimated. Averages are computed over the whole 2000-2015 period. All variables are computed in deviation to the national mean.

Growth models like model 12 are likely affected by endogeneity bias due to potential simultaneity between the dependent variable (the growth rate) and the right-hand side regressors, reverse causality and measurement errors (Durlauf, Johnson, and Temple, 2005). In our context, the problem looks quite relevant especially for the inclusion of trade variables. Following the recent literature on trade and growth treating trade openness as an endogenous variable (e.g. Frankel and Romer, 1999), we claim that model (12) can be consistently estimated by means of a linear two-stage least squares (linear-TSLS) estimator, i.e. instrumenting each potentially endogenous trade variable ($\ln \mathbf{eci}$, $\ln \mathbf{open}$, **rv**) with its time lag (computed as average values over the five years 1995-1999 preceding the sample period 2000-2015), plus their squared values.

Estimation results from the linear model (12) are reported in Table 3 along with some additional tests about the quality of the estimates. The Sargan-J test gives statistical support to the chosen set of instruments at the conventional nominal level of significance; both the tests for weak- and under-identification reject the null hypothesis.

Insert Table 3 about here

The first specification includes only initial conditions and the three trade variables ($\ln \mathbf{eci}$, $\ln \mathbf{open}$, and **rv**). We document a positive relationship between the response variable (the growth rate over the period 2000-2015) and the three indicators of trade performance: the estimated parameters of $\ln \mathbf{eci}$ and $\ln \mathbf{open}$ are statistically significant at the 5 per cent nominal level of

interval according to the following monotonic transformation: $\mathbf{eci} = ((\tanh(\mathbf{ECI}) + 1)/2) + 1)/2$, where $\tanh(.)$ stands for the hyperbolic tangent function, which plots the transformed values in the $[0,1]$ interval. The variable included in the growth equation is then the log transformation of **eci**. The monotonic transformation we have chosen looks preferable to alternatives like those based on the standardized normal distribution or the normalization $(\mathbf{ECI} - \mathbf{ECI}_{min})/(\mathbf{ECI}_{max} - \mathbf{ECI}_{min})$, where $\mathbf{ECI}_{min}$ and $\mathbf{ECI}_{max}$ denote the sample minimum and maximum value, respectively, because it turns out to be less dependent on possible outliers in the sample.

significance according to the heteroskedasticity consistent standard error á la White; the parameter of **rv** is only weakly significant (at 10 per cent nominal level of significance). The positive and significant marginal effect of ln **open** is consistent with a wide empirical literature mainly based on cross-country data (e.g. Frankel and Romer, 1999; Daumal et al., 2010; Busse and Königer, 2012). The positive and significant marginal effect of **rv** is in line with Saviotti and Frenken (2008) on 20 OECD countries, Rebelo and Silva (2017) on Portugal, and Boschma and Iammarino (2009) on Italian regions who have shown the existence of a nexus between diversification and economic growth. These studies find that including in the export basket products 'related' to the pre-existing ones positively influences labour productivity growth. Finally, the positive and significant marginal effect of economic complexity (ln **eci**) corroborates Hidalgo, Klinger, Barabási, and Hausmann (2007)'s prediction according to which what countries or regions produce today matters in explaining their subsequent economic growth pattern: the production of goods that are at the 'core' of the so-called 'product space' - typically more sophisticated manufacturing products - leads to higher growth. In other words, our results confirm the existence, at a local level, of a causality nexus between the sophisticatedness of local production and subsequent economic performances in the medium-long run. It is worth pointing out that this positive marginal effect is net of the marginal effect of the overall trade openness and of the degree of related variety, implying that, *ceteris paribus*, the composition of the export basket matters.

The second specification includes also the traditional proximity growth determinants. Confirming the predictions of the neoclassical growth model, initial conditions exerted a negative effect on the growth rate of productivity, while all the other proximate growth determinants do not have any significant impact on growth. Net of the effect of initial conditions, ln **eci** and ln **open** still enter positively and significantly the cross-section growth equation, while **rv** is not significant any more. Thus, our final estimation is a more parsimonious model, including only initial conditions, the degree of economic complexity and the degree of openness. These three variables now enter highly significantly in model 3, and the adjusted R^2 of this model turns out to be higher than the one computed with models 1 and 2. Thus, on the basis of this last parsimoniuos specification, we compute simulations (counterfactual analysis) to assess the effect of $\ln \mathbf{eci}$ and $\ln \mathbf{open}$ on the distribution dynamics of labor productivity in the short and in the long run.

3.4 Counterfactual analysis

In this section we report the results of the counterfactual analysis (the third step of our empirical strategy) for each significant growth determinant, namely the initial level of per worker value added ($\ln \mathbf{y}_0$), the degree of economic complexity ($\ln \mathbf{eci}$), and the degree of trade openness ($\ln \mathbf{open}$).

Starting from the counterfactual analysis of initial conditions, Table 1 shows that the counterfactual distribution in the final year (2015) is characterized by a value of the Theil index (0.038) slightly higher than the one computed for the actual distribution (0.033). Such increase is evenly distributed in the two components (within and between) of the Theil index. Thus, initial conditions contribute to the reduction of labor productivity inequalities among Italian provinces. In the long run, initial conditions seem also to weakly reduce the degree of polarization: BIPOL for the counterfactual ergodic distribution (0.497) is slightly higher than the one computed for the actual ergodic distribution (0.467). Figure 9(a) displays indeed that the counterfactual ergodic distribution is markedly bimodal. Thus, if all provinces had had the same values of initial conditions, the long run degree of polarization would had been higher. In particular, the simulated distribution shows a lower density around the second peak, i.e. a lower fraction of provinces belonging to the medium-high productivity club with respect to the actual ergodic distribution. Thus, we may conclude that technological catching up, proxied by the initial level of labor productivity, has contributed to reduce inequality and polarization.

Trade related variables ($\ln \mathbf{eci}$ and $\ln \mathbf{open}$) are instead strongly responsible for inequality in the short run and for polarization in the long run. The counterfactual distribution in the final year (2015) is characterized by a value of the Theil index (0.02) lower than the one computed for the actual distribution (0.033), implying that these two variables tend to increase the overall degree of inequality. This effect is, however, observable only in the between component. Coherently, the polarization index is zero in the counterfactual ergodic distribution for the two variables; thus, if all provinces had had the same values of trade openness and economic complexity, polarization in the long run would had disappeared. Figures 9(b-c) display indeed unimodal distributions.

Insert Figure 9 about here

4 Conclusions

In this paper we have analyzed the effect of economic (trade) complexity on regional productivity growth in Italy. The results of our counterfactual analysis suggest that ECI contributed to increase spatial productivity inequalities both in the short and in the long run, and was responsible for the polarization of regional productivity levels in the long run. On the contrary, the initial level of labor productivity slightly contrasted the long run tendency towards polarization and slightly reduced inequalities.

Our findings contribute to the debate on the effect of economic complexity on productivity growth at the sub-national scale in Italy. A previous study by Coniglio, Lagravinese, and Vurchio (2016) had already identified a positive relationship between economic complexity and regional growth using Italian provincial data. Nevertheless, the complexity measure they used, the so-called $EXPY$, is strongly criticized today. The use of GDP per capita income information in the creation of $EXPY$ is indeed problematic: sectors with high values of $PRODY$ are, by construction, those where high income countries play a major role in production, relative to the other participants in world exports in that sector. The observation that ‘rich countries export rich country goods’ is close to being a circular argument (Hidalgo and Hausmann, 2009), and this is also the case for regions. ECI represents a better measure of economic complexity overcoming this issue.

Under a different perspective, Boschma and Iammarino (2009) have shown the existence of a positive relationship between diversification and regional economic growth in Italy. Their study has demonstrated that including in the export basket products ‘related’ to the pre-existing ones positively influences economic outcomes, and in particular labor productivity, employment and value added growth. However, the economic complexity theory suggests that ECI largely captures the effect of the variety by adding an essential, further information: the ubiquity of products. Indeed, in order to have a significant impact on growth, variety - related or not - needs to embed the measure of the number of capabilities required by the products that countries make (and export), that is the products’ ubiquity. This argument is consistent with our econometric results, according to which, once ECI is controlled for, the parameter of related variety looses significance.

Finally, in contrast to the works of Coniglio, Lagravinese, and Vurchio (2016) and Boschma and Iammarino (2009), our study does not only focus on the transitional growth (regression) analysis, but it also try to assess the contribution of ECI to the long run ergodic distribution of productivity by province. Our findings confirm that, in the long run, the productivity of regions is strongly determined by the variety and sophistication of the products they make. In other words our results show that regions tend to converge to the level of productivity dictated by the complexity level of their productive structures, suggesting that development efforts should focus on generating the conditions that would allow complexity to emerge in the long run.

References

AGHION, P., AND P. HOWITT (1998): *Endogenous growth theory*. MIT press.

ANDERSON, G., O. LINTON, AND T. W. LEO (2012): “A polarization-cohesion perspective on cross-country convergence,” *Journal of economic growth*, 17(1), 49–69.

BALASSA, B. (1965): “Trade liberalisation and “revealed” comparative advantage 1,” *The Manchester School*, 33(2), 99–123.

BARRO, R. J., AND J.-W. LEE (1993): “International comparisons of educational attainment,” *Journal of monetary economics*, 32(3), 363–394.

BASILE, R. (2009): “Productivity polarization across regions in Europe: The role of nonlinearities and spatial dependence,” *International Regional Science Review*, 32(1), 92–115.

——— (2010): “Intra-distribution dynamics of regional per-capita income in Europe: evidence from alternative conditional density estimators,” *Statistica*, 70(70), 3–22.

BOSCHMA, R. (2005): “Proximity and innovation: a critical assessment,” *Regional studies*, 39(1), 61–74.

BOSCHMA, R., K. FRENKEN, H. BATHELT, M. FELDMAN, D. KOGLER, ET AL. (2012): “Technological relatedness and regional branching,” *Beyond territory. Dynamic geographies of knowledge creation, diffusion and innovation*, pp. 64–68.

BOSCHMA, R., AND S. IAMMARINO (2009): “Related variety, trade linkages, and regional growth in Italy,” *Economic geography*, 85(3), 289–311.

BUSSE, M., AND J. KÖNIGER (2012): “Trade and economic growth: A re-examination of the empirical evidence,” .

CHESHIRE, P., AND S. MAGRINI (2005): “Analysing Growth and Distribution Dynamics-Isolating Divergence Factors,” .

CONIGLIO, N. D., R. LAGRAVINESE, AND D. VURCHIO (2016): "Production sophisticatedness and growth: evidence from Italian provinces before and during the crisis, 1997–2013," *Cambridge Journal of Regions, Economy and Society*, 9(2), 423–442.

DAUMAL, M., ET AL. (2010): "The impact of international trade flows on economic growth in Brazilian states," *Review of Economics and Institutions*, 2(1).

DURLAUF, S. N., P. A. JOHNSON, AND J. R. TEMPLE (2005): "Growth econometrics," *Handbook of economic growth*, 1, 555–677.

ERTUR, C., AND W. KOCH (2007): "Growth, technological interdependence and spatial externalities: theory and evidence," *Journal of applied econometrics*, 22(6), 1033–1062.

——— (2011): "A contribution to the theory and empirics of Schumpeterian growth with worldwide interactions," *Journal of Economic Growth*, 16(3), 215.

FELDMAN, M. P., AND D. B. AUDRETSCH (1999): "Innovation in cities:: Science-based diversity, specialization and localized competition," *European economic review*, 43(2), 409–429.

FIASCHI, D., L. GIANMOENA, AND A. PARENTI (2011): "The dynamics of labour productivity across Italian provinces: convergence and polarization," *Rivista italiana degli economisti*, 16(2), 209–240.

——— (2018): "Deep and Proximate Determinants of the World Income Distribution," *Unpublished paper*.

FRANKEL, J. A., AND D. H. ROMER (1999): "Does trade cause growth?," *American economic review*, 89(3), 379–399.

GLAESER, E. L., H. D. KALLAL, J. A. SCHEINKMAN, AND A. SHLEIFER (1992): "Growth in cities," *Journal of political economy*, 100(6), 1126–1152.

GROSSMAN, G. M., AND E. HELPMAN (1991): "Trade, knowledge spillovers, and growth," *European economic review*, 35(2-3), 517–526.

HAUSMANN, R., AND C. A. HIDALGO (2011): "The network structure of economic output," *Journal of Economic Growth*, 16(4), 309–342.

HAUSMANN, R., C. A. HIDALGO, S. BUSTOS, M. COSCIA, A. SIMOES, AND M. A. YILDIRIM (2014): *The atlas of economic complexity: Mapping paths to prosperity*. Mit Press.

HAUSMANN, R., J. HWANG, AND D. RODRIK (2007): "What you export matters," *Journal of economic growth*, 12(1), 1–25.

HAUSMANN, R., AND B. KLINGER (2006): "Structural transformation and patterns of comparative advantage in the product space," .

HAUSMANN, R., AND D. RODRIK (2003): "Economic development as self-discovery," *Journal of development Economics*, 72(2), 603–633.

HIDALGO, C. A., AND R. HAUSMANN (2009): "The building blocks of economic complexity," *Proceedings of the national academy of sciences*, 106(26), 10570–10575.

HIDALGO, C. A., B. KLINGER, A.-L. BARABÁSI, AND R. HAUSMANN (2007): "The product space conditions the development of nations," *Science*, 317(5837), 482–487.

HOWITT, P. (2000): "Endogenous growth and cross-country income differences," *American Economic Review*, 90(4), 829–846.

HYNDMAN, R. J. (1996): "Computing and graphing highest density regions," *The American Statistician*, 50(2), 120–126.

HYNDMAN, R. J., D. M. BASHTANNYK, AND G. K. GRUNWALD (1996): "Estimating and visualizing conditional densities," *Journal of Computational and Graphical Statistics*, 5(4), 315–336.

JOHNSON, P. A. (2005): "A continuous state space approach to "convergence by parts"," *Economics Letters*, 86(3), 317–321.

MANKIW, N. G., D. ROMER, AND D. N. WEIL (1992): “A contribution to the empirics of economic growth,” *The quarterly journal of economics*, 107(2), 407–437.

MCCANN, P. (2013): *Modern urban and regional economics*. Oxford University Press.

NOOTEBOOM, B. (2000): “Learning by interaction: absorptive capacity, cognitive distance and governance,” *Journal of management and governance*, 4(1-2), 69–92.

PHILLIPS, P. C., AND D. SUL (2007): “Transition modeling and econometric convergence tests,” *Econometrica*, 75(6), 1771–1855.

——— (2009): “Economic transition and growth,” *Journal of Applied Econometrics*, 24(7), 1153–1185.

PSACHAROPOULOS, G., AND A. M. ARRIAGADA (1986): “The educational composition of the labour force: an international comparison,” *Int’l Lab. Rev.*, 125, 561.

QUAH, D. T. (1997): “Empirics for growth and distribution: stratification, polarization, and convergence clubs,” *Journal of economic growth*, 2(1), 27–59.

REBELO, F., AND E. G. SILVA (2017): “Export variety, technological content and economic performance: the case of Portugal,” *Industrial and Corporate Change*, 26(3), 443–465.

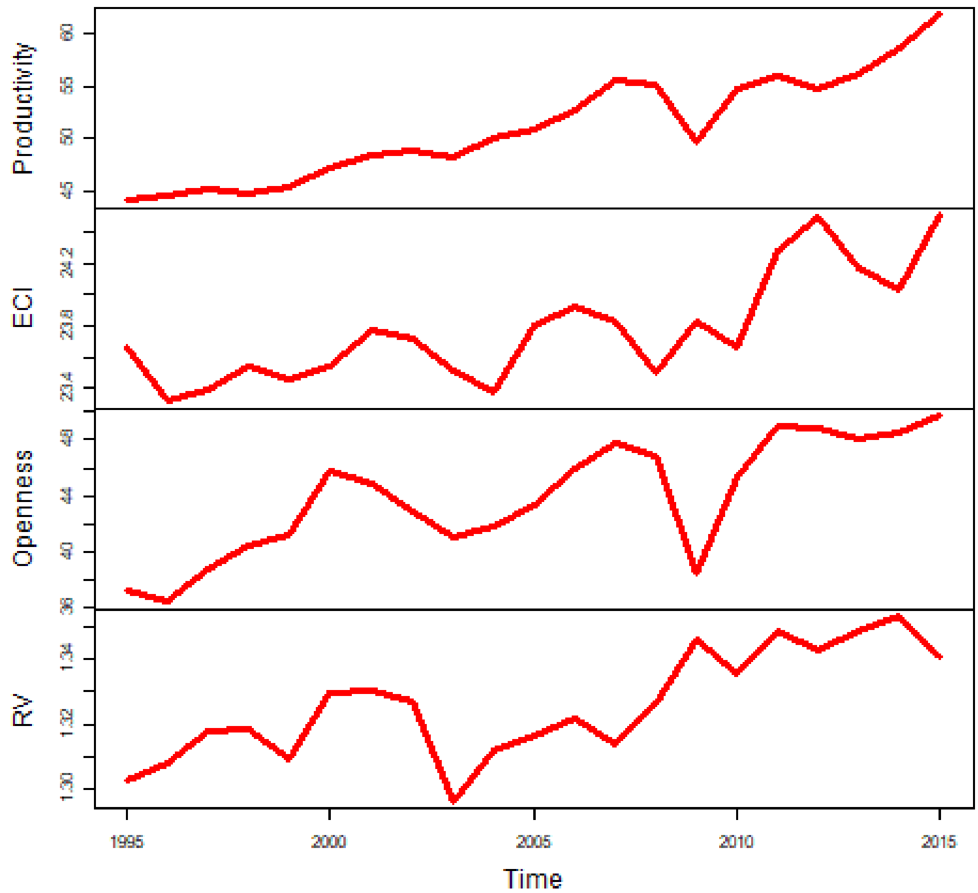
SAVIOTTI, P. P., AND K. FRENKEN (2008): “Export variety and the economic performance of countries,” *Journal of Evolutionary Economics*, 18(2), 201–218.

THEIL, H. (1972): “Statistical decomposition analysis; with applications in the social and administrative sciences,” Discussion paper.

VAN OORT, F. G. (2017): *Urban growth and innovation: Spatially bounded externalities in the Netherlands*. Routledge.

FIGURE 1

Time trends of Italian average levels of labor productivity in manufacturing, degree of openness, related variety and degree of economic complexity (ECI)

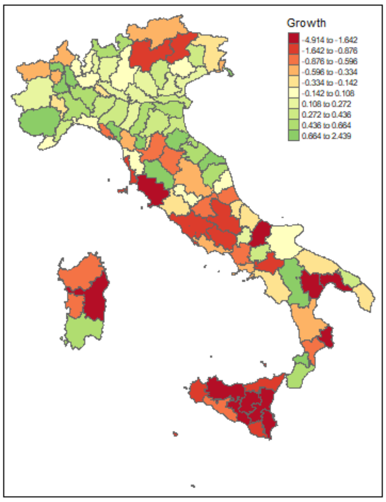


Productivity: Gross Value Added (GVA) per worker - thousands of euros; ECI: Levels; Openness: (Import+Export)/GVA*100; RV: levels

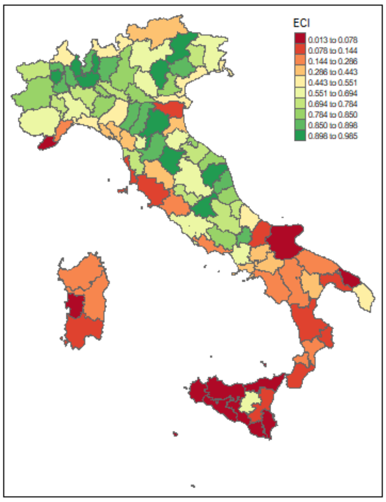
Source: Our elaboration on ISTAT data.

FIGURE 2

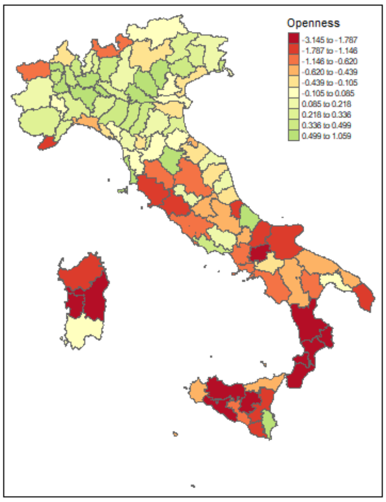
Maps of labor productivity growth, economic complexity index, openness and related variety. Average values over the period 2000-2015. All variables measured in deviation from the national average.



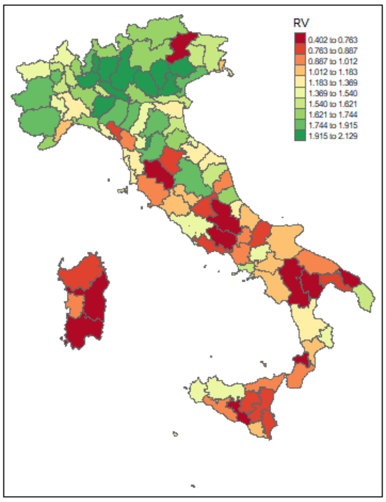
(a) Growth



(b) ECI



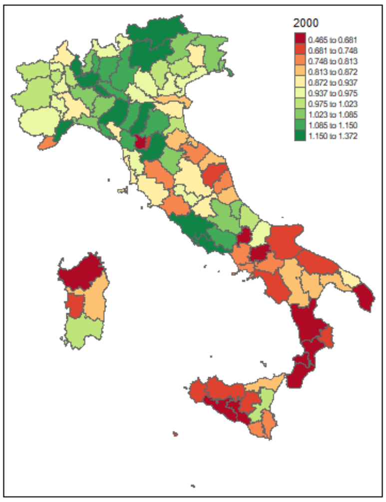
(c) Openness



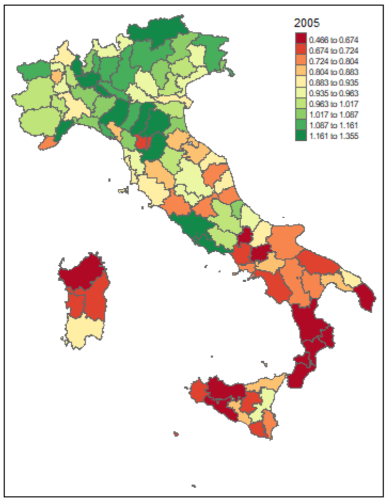
(d) RV

Source: Our elaboration on ISTAT data.

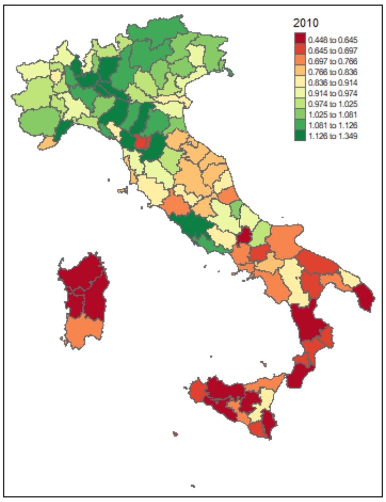
FIGURE 3
Maps of relative labor productivity in manufacturing of Italian provinces



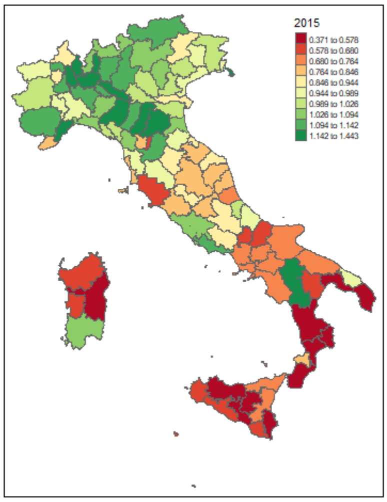
(a) 2000



(b) 2005



(c) 2010

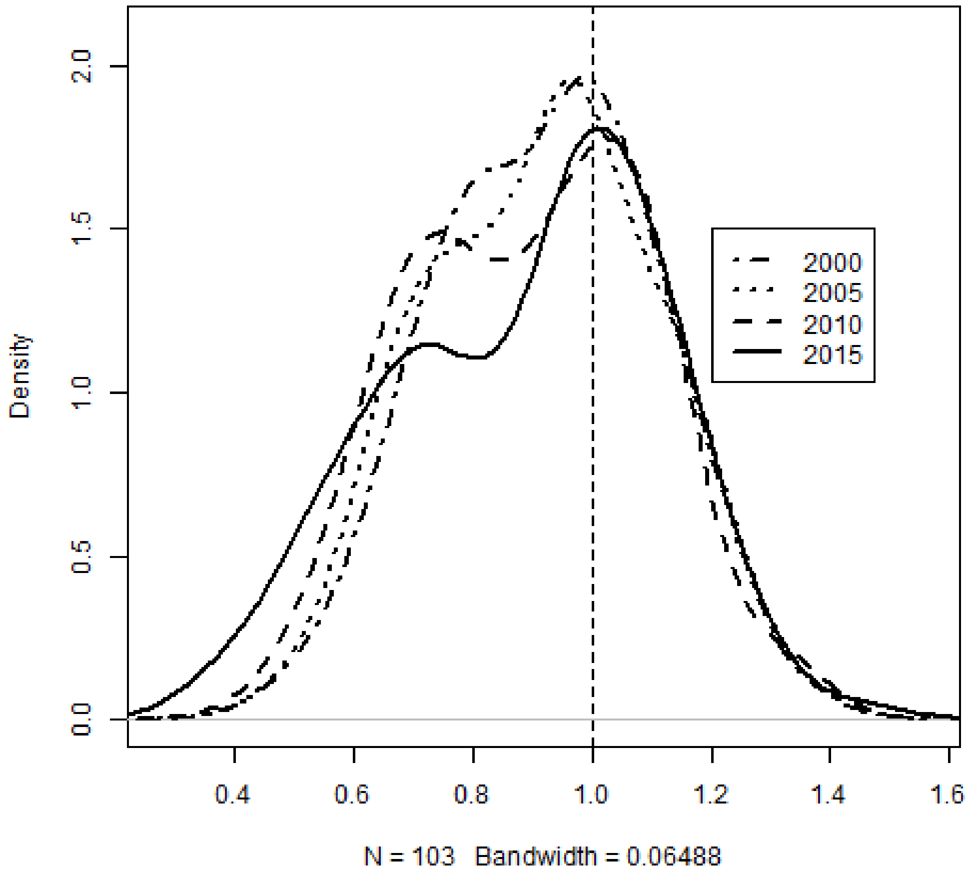


(d) 2015

Source: Our elaboration on ISTAT data.

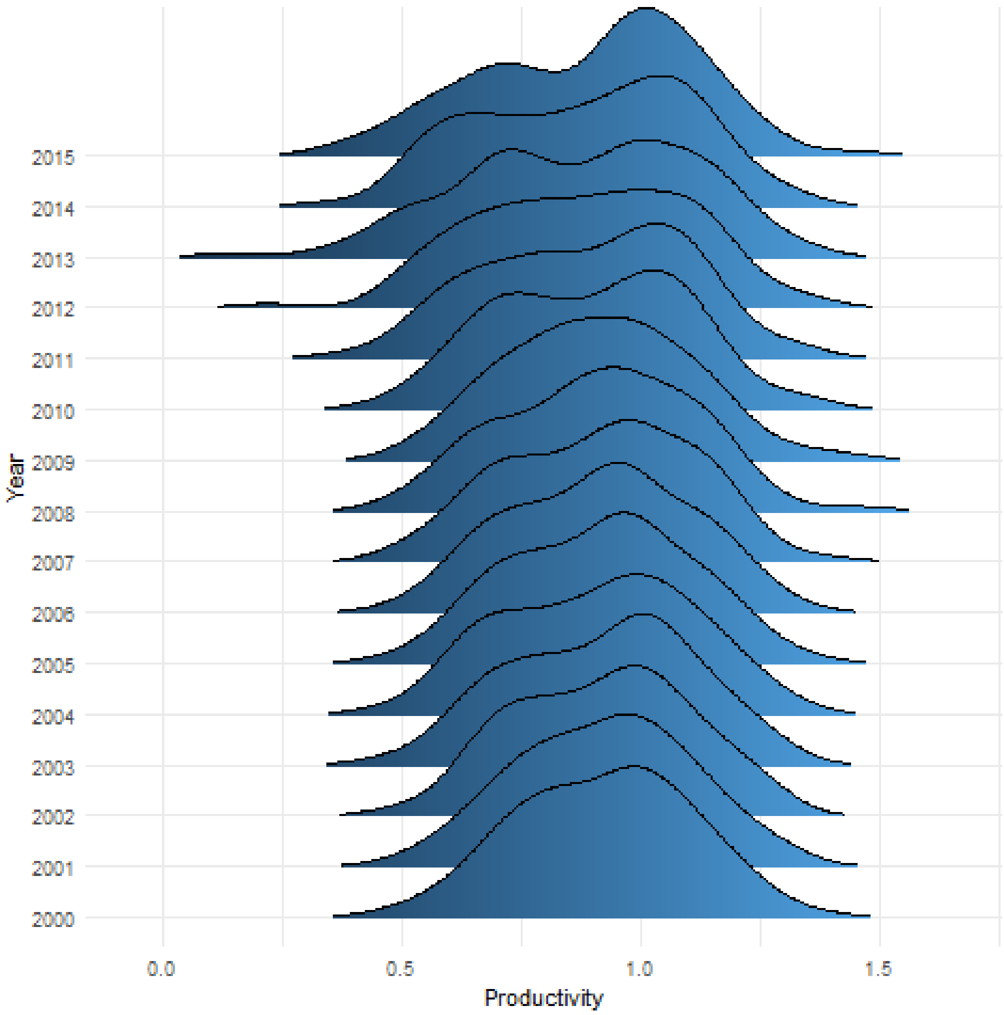
FIGURE 4

Univariate density of relative labor productivity in manufacturing of Italian provinces



Source: Our elaboration on ISTAT data.

FIGURE 5
Univariate density of relative labor productivity in manufacturing of Italian provinces



Source: Our elaboration on ISTAT data.

TABLE 1

Relative value added per worker in manufacturing of Italian provinces. Measures of inequality and polarization. Years: 2000, 2005, 2010, and 2015. Bootstrap standard errors in parenthesis.

Actual density		Theil		BIPOL
	Overall	Within	Between	
Actual distribution 2000	0.020	0.008	0.012	0.000
	(0.003)	(0.001)	(0.002)	(NA)
Actual distribution 2005	0.021	0.008	0.013	0.000
	(0.002)	(0.001)	(0.002)	(NA)
Actual distribution 2010	0.025	0.009	0.016	0.460
	(0.003)	(0.001)	(0.002)	(0.018)
Actual distribution 2015	0.033	0.009	0.024	0.417
	(0.004)	(0.001)	(0.004)	(0.020)
Counterfactual distribution 2015 (Initial conditions)	0.038	0.013	0.025	
	(0.005)	(0.002)	(0.004)	
Counterfactual distribution 2015 (ECI)	0.021	0.008	0.013	
	(0.003)	(0.001)	(0.002)	
Counterfactual distribution 2015 (Open)	0.020	0.009	0.011	
	(0.004)	(0.002)	(0.002)	
Actual Ergodic distribution				0.467
				(0.017)
Counterfactual Ergodic distribution (Initial conditions)				0.497
				(0.024)
Counterfactual Ergodic distribution (ECI)				0.000
				(NA)
Counterfactual Ergodic distribution (Open)				0.000
				(NA)

Source: Our elaboration on ISTAT data.

TABLE 2

Convergence clubs

Club 1 (59 provinces)
SAVONA, PARMA, MODENA, MILANO, POTENZA, TRIESTE, BOLOGNA, VARESE, CREMONA, LECCO, NOVARA, LATINA, BOLZANO, LODI, PAVIA, REGGIO NELL'EMILIA, BERGAMO, VERCELLI, CUNEO, FIRENZE, BRESCIA, PIACENZA, GENOVA, ROMA, VICENZA, RAVENNA, TRENTO, LUCCA, CAGLIARI, COMO, SONDRIO, MASSA CARRARA, TORINO, PESCARA, UDINE, PADOVA, FERRARA, VERONA, TREVISO, LA SPEZIA, MANTOVA, GORIZIA, VENEZIA, ALESSANDRIA, PORDENONE, BRINDISI, SIENA, CHIETI, AOSTA, ASTI, FORLI' - CESENA, PISA, BIELLA, FROSINONE, TERNI, ANCONA, BELLUNO, AREZZO, L'AQUILA
Club 2 (44 provinces)
RIMINI, ROVIGO, VERBANO-CUSIO-OSSOLA, PERUGIA, PESARO E URBINO, IMPERIA, ASCOLI PICENO, PISTOIA, RIETI, LIVORNO, VITERBO, MACERATA, VIBO VALENTIA, FOGGIA, TERAMO, CASERTA, BARI, NAPOLI, SALERNO, BENEVENTO, CATANIA, MESSINA, PRATO AVELLINO, ISERNIA, GROSSETO, ORISTANO, MATERA, RAGUSA, TRAPANI, SASSARI, AGRIGENTO, CAMPOBASSO, NUORO, ENNA, PALERMO, LECCE, CATANZARO, COSENZA, REGGIO DI CALABRIA, CROTONE, SIRACUSA, TARANTO, CALTANISSETTA

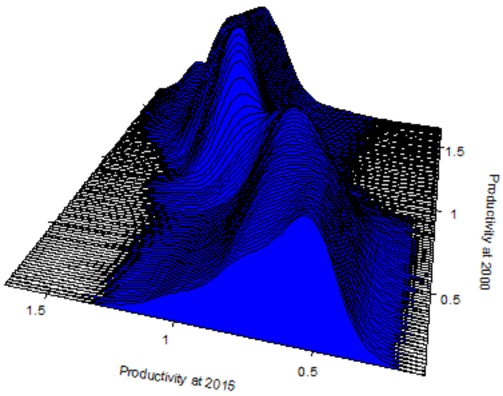
Source: Our elaboration on ISTAT data.

TABLE 3
2SLS estimates

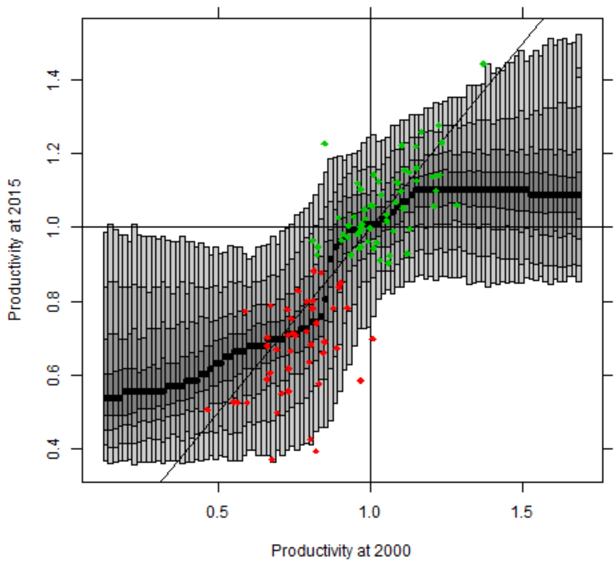
Variable	Model 1	Model 2	Model 3
	Coefficients and heteroskedasticity consistent s.e. (in parenthesis)		
Intercept	0.037	0.007	0.076
	(0.100)	(0.113)	(0.090)
$\ln y_0$	-1.755***	-2.120***	-1.866***
	(0.531)	(0.653)	(0.537)
$\ln(eci)$	0.259**	0.221*	0.300***
	(0.123)	(0.127)	(0.116)
$\ln(open)$	0.321**	0.349**	0.401**
	(0.153)	(0.157)	(0.165)
$\ln(rv)$	0.525*	0.467	
	(0.309)	(0.288)	
$\ln s_k$		-0.114	
		(0.425)	
$\ln(n)$		0.084	
		(0.120)	
$\ln s_h$		0.021	
		(0.028)	
$\ln s_A$		0.106	
		(0.121)	
	Diagnostics		
Weak instr.-$\ln(eci)$	166.377***	128.993***	253.300***
	(0.000)	(0.000)	(0.000)
Weak instr.-$\ln(open)$	85.596***	76.720***	130.441***
	(0.000)	(0.000)	(0.000)
Weak instr.-$\ln(rv)$	52.356***	43.173***	
	(0.000)	(0.000)	
Wu-Hausman	6.684***	6.329***	6.795***
	(0.000)	(0.000)	(0.001)
Sargan	0.927	0.753	0.104
	(0.818)	(0.860)	(0.949)
Multiple R-Squared	0.523	0.520	0.538
Moran I	-0.132	-0.025	0.459
	(0.552)	(0.510)	(0.323)

Source: Our elaboration on ISTAT data.

FIGURE 6
Conditional density. Kernel density estimates with variable bandwidth using cross-sectional data for 2000 and 2015



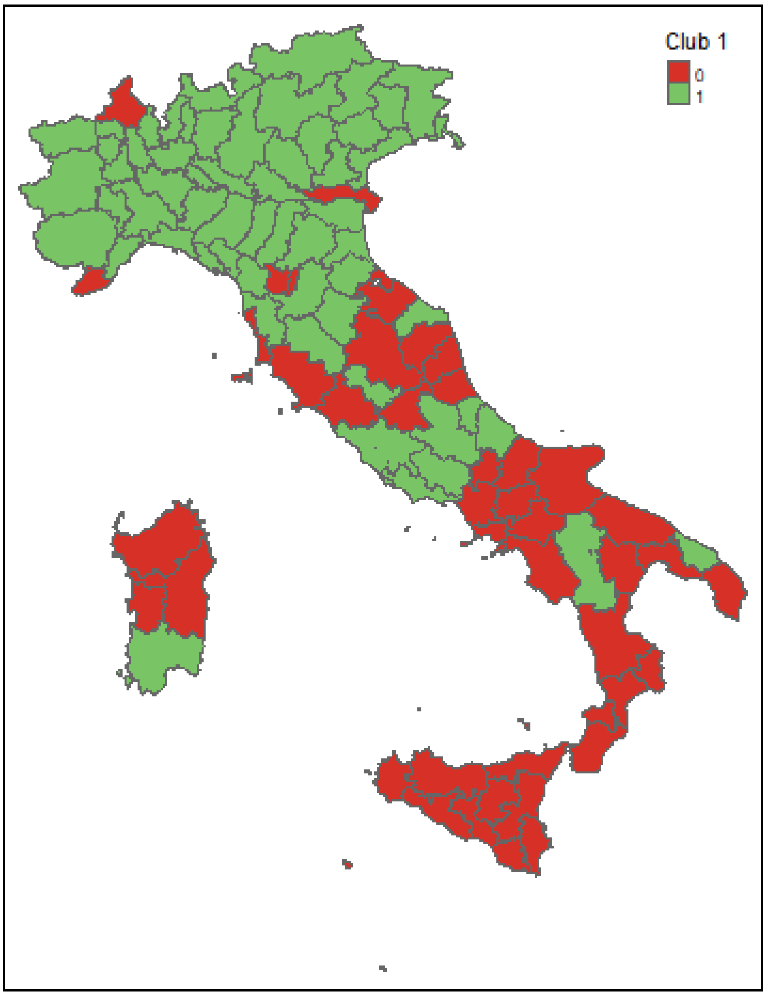
(a) Stacked density plot



(b) HDR plot

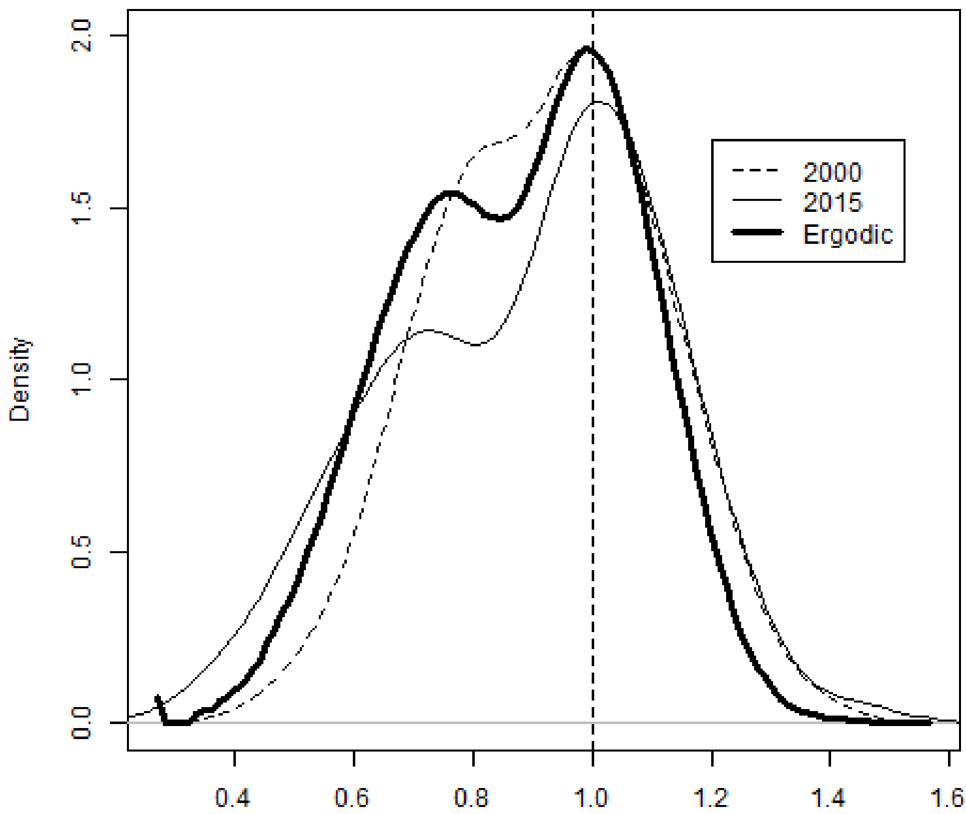
Source: Our elaboration on ISTAT data.

FIGURE 7
Convergence clubs



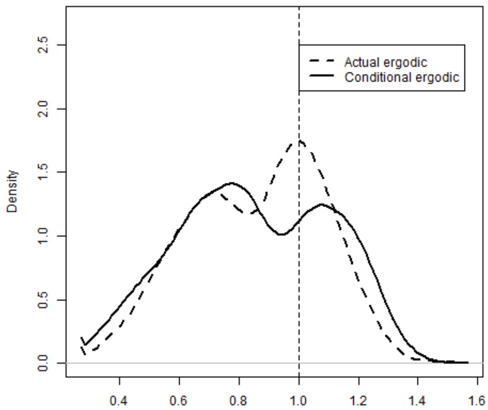
The two cluster have been identified using the methodology proposed by Phillips and Sul (2007, 2009).
Source: Our elaboration on ISTAT data.

FIGURE 8
Initial, final and ergodic distributions using cross-sectional data for 2000 and 2015

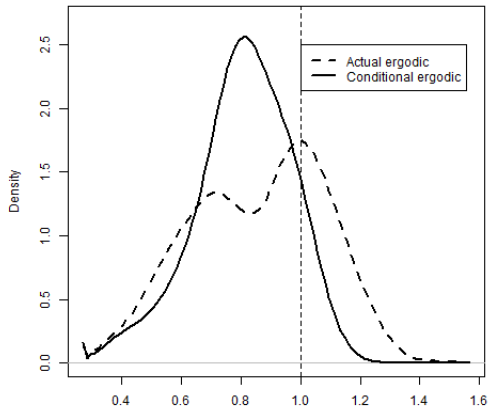


Source: Our elaboration on ISTAT data.

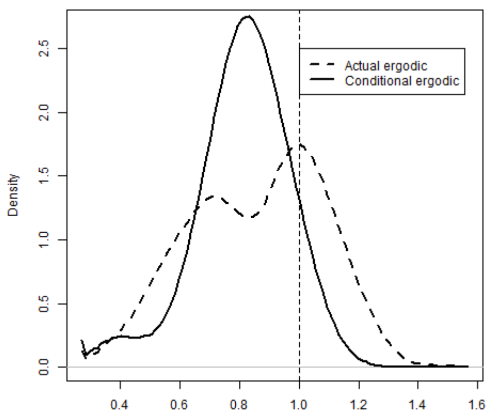
FIGURE 9
Actual and counterfactual ergodic distributions



(a) Counterfactual for lny0



(b) Counterfactual for ECI



(c) Counterfactual for Openness

Source: Our elaboration on ISTAT data.