

INSTITUTE
OF ECONOMICS



Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM

WORKING PAPER SERIES

Behind the Italian Regional Divide: An Economic Fitness and Complexity Perspective

Angelica Sbardella ^a
Andrea Zaccaria ^b
Luciano Pietronero ^c
Pasquale Scaramozzino ^{d,e}

^a Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy.

^b ISC-CNR—Institute of Complex Systems, UOS Sapienza, Rome, Italy.

^c Centro Studi e Ricerche Enrico Fermi, Rome, Italy.

^d School of Finance and Management, SOAS University of London, London, UK.

^e Department of Economics and Finance, University of Rome Tor Vergata, Rome, Italy

2021/30

September 2021

ISSN(ONLINE) 2284-0400

Behind the Italian Regional Divide: An Economic Fitness and Complexity Perspective

Angelica Sbardella^{1*}, Andrea Zaccaria², Luciano Pietronero³, and Pasquale Scaramozzino^{4,5}

Institute of Economics, Sant'Anna School of Advanced Studies, Pisa, Italy

²ISC-CNR—Institute of Complex Systems, UOS Sapienza, Rome, Italy

³Centro Studi e Ricerche Enrico Fermi, Rome, Italy

⁴School of Finance and Management, SOAS University of London, London, UK

⁵Department of Economics and Finance, University of Rome Tor Vergata, Rome, Italy

Abstract

This paper applies the Economic Fitness and Complexity approach to analyse the underlying factors behind the wide and persistent economic disparities across the Italian regional units. Measures of regional fitness are obtained from their revealed comparative advantage and from their patent performance. Southern regions tend to be characterised by a lower level of complexity than the regions in the Centre-North of the country. We interpret these results as indicating a lower level of capability endowment in the South. The system-wide approach of the paper is able to identify some critical sectors which display a rich pattern of connections with other sectors and which could play a pivotal role to create additional capabilities and foster a more balanced regional development.

JEL classification: L60, O11, O14, O47

Keywords: Italian regional divide, Economic Fitness and Complexity, Productive and technological capabilities.

[§] We are grateful to Maria Enrica Virgillito, Andrea Boltho and Beniamino Pisicoli for their helpful comments.

^{*} The author acknowledges the support from European Union's Horizon 2020 research and innovation programme under grant agreement No. 822781 GROWINPRO – Growth Welfare Innovation Productivity.
©< 2019 >.

Introduction

The analysis of regional economic development is usually conducted in terms of mainstream models of economic growth, both of the exogenous and endogenous growth varieties (Aghion and Howitt, 1998; Barro and Sala-i Martin, 2004; Acemoglu, 2009). The main drivers of regional growth are identified in the accumulation of physical and human capital, investment in infrastructure, expenditure on R&D, openness to trade, and migration flows; agglomeration dynamics could also play a role by affecting the spatial interactions among the regional units (Arbia *et al.*, 2010). Exogenous and endogenous growth theories however yield sharply contrasting predictions about the comparative growth of regions, with the former predicting convergence of the poorer areas of the country to the richer ones whilst the latter predict that relative regional inequalities could persist over time.

Recent years have seen the emergence of the Economic Complexity (EC) perspective, a new approach to economic growth which describes the economy as a complex evolving system characterized by several interacting domains (Kirman, 1992; Arthur, 2013; Dosi *et al.*, 2013). Rather than analysing large amounts of different types of data – and thus dealing with a significant level of noise in the information available – the focus here is on maximizing the signal-to-noise ratio by designing novel network-based empirical algorithms tailored to the economic problem of interest. This approach is able to separate random noise from the desired signal in an optimal way and to summarize the relevant information into synthetic indicators.

The traditional analysis of regional economic development also tends to focus on aggregate, macroeconomic variables. Even micro-founded analytical approaches – such as those that build on Arrow's (1962) seminal learning-by-doing analysis – fail to fully account for the diversities across firms and for the complex nature of their complementarities and interactions. EC goes beyond simple aggregate indicators of economic inputs and outputs and instead considers a more fine-grained and structural vision of the productive possibilities of the economy, which emphasizes the importance of export specialization patterns for long-run growth (Hausmann and Rodrik, 2003; Hausmann and Klinger, 2006). The profile of trade specialization of a country can in fact be regarded as a reflection of its underlying productive capabilities, which are defined as the skills that enable its economy to expand into new production requirements and to adopt new technologies (Cimoli *et al.*, 2010; Teece *et al.*, 1997). This line of research thus seeks to explain the growth potential of an economy in terms of the complexity of its production structures. It does so by articulating a granular approach to the determinants of growth that, by opposing quality to quantity, focuses more on *which* than on *how many* productive outputs a macroeconomic actor is able to be competitive in.

In particular, this paper applies the Economic Fitness and Complexity approach (Tacchella *et al.*, 2012) to the analysis of the production patterns and technological development of Italian

regions. Our system-wide empirical strategy allows us to define two rankings of regions in terms of their productive and technological competitiveness. Moreover, it allows to identify some critical sectors which could play a crucial role in regional economic development and could therefore be a useful target for regional and industrial policy.

A more balanced growth performance among the different areas of the country would have a beneficial effect on the national economy. Because of the interdependencies between the various regions, it is estimated that an increase by €100 of income in the South of Italy leads to an increase in demand in the Centre-North by €40; by contrast, an increase of income by €100 in the Centre North only leads to an increase in demand in the South by €5 (Viesti, 2021). It is therefore apparent that the development of the Southern regions would have a strong multiplier effect on the national economy.

The structure of the paper is as follows. The next section summarises the background to the analysis in the paper. Section 3 explains the methods used in the analysis and describes the data. Section 4 discusses our empirical findings and comments on the results. Section 5 concludes.

Background

From a theoretical point of view, the economic complexity approach can be related to the structuralist work initiated by Hirschman (1958) and Penrose (1959) and to the evolutionary tradition (Nelson and Winter, 1982; Dosi *et al.*, 1990; Dosi and Nelson, 1994), specifically to their emphasis on the role of productive capabilities.

Productive capabilities can be seen as a set of skills which, in addition to allowing the implementation of the existing technologies, also facilitate the adaptation to new production requirement and stimulate the introduction of new technologies (Lall, 1992; Kremer, 1993) by linking different innovation regimes to different organisational routines. Teece *et al.* (1994) explain that intangible, non-transferable capabilities in an organisation are essential for a coherent, competence-based growth of the firm. The capability-based theory of the firm (Dosi and Nelson, 2010; Winter, 1997) describes firms as behavioural entities which continuously evolve and learn (Simon, 1991), whose economic performance is determined by a set of organisational capabilities which are embedded in their procedural know-how and which are persistent over time.

Abramovitz (1986) shifted the focus away from a firm-centred towards a location-based notion of capabilities, and referred to social capabilities as all those attributes that affect a country's ability to operate modern and large-scale businesses. These would also include their political and social characteristics, such as physical and human capital endowments, institutions, skills, technology. Capabilities should therefore not be interpreted in a narrow individualistic sense (see also Fagerberg and Srholec, 2017).

An important recent advance in our understanding of how capabilities can account for differences in output per capita across economies is due to Sutton (2012) and Sutton and Trefler (2016). According to their analysis, capabilities enable firms to maximise the returns from their investment by enhancing their labour productivity and by creating the conditions for further investment. A crucial insight of this analysis is that the capabilities required to produce high-quality products will always be scarce. There will therefore only be a limited number of firms in each industry which hold a dominant position.

The arguments by Sutton *et al.* can fruitfully be applied to regional economic growth. In the context of a national economy, capabilities may tend to concentrate in some specific regions in order to take full advantage of agglomeration externalities. Talents will therefore end up being unequally distributed on the national territory, and this will result in systematic regional disparities. These would take the form of significant differences in both the levels and the rates of growth of regional GDP and income per capita, which could become permanent and even amplify over time.

In a unified country, common national institutions and the mobility of labour and capital – as well as freely tradeable goods and services – could in principle contribute to a convergence of living standards across the regional units. Regional policy can support the poorer areas of a country both through transfers and subsidies and by encouraging private investment. Convergence to the more advanced regions, however, also requires the existence of capabilities for growth in the poorer areas. In the absence of these capabilities, convergence of the poorer regions towards the richer ones may fail to materialise.

Since the beginning of the current century the Italian economy has suffered from a decline in productivity, high unemployment, decreasing real wages, and unfavourable demographic trends. The global financial crisis which started in 2008 affected the country particularly hard because of these prior existing vulnerabilities (Rodano, 2018). Italy is however also characterised by a profound dualism between the more advanced Centre-North regions and the South of the country (the *Mezzogiorno*), and the financial crisis has hit the South of the country particularly hard (Viesti, 2021).

Historically, the level of GDP per capita in the Southern regions had only briefly converged towards Northern Italian levels during the 1960s, despite massive regional policy efforts aimed at stimulating growth and reducing the development and income gap between the two macro-areas of the country. Since the 1970s there has effectively been no convergence of the South, and the income and GDP gaps with the Centre-North have remained stubbornly high (Boltho *et al.*, 1997, 2018). Since the beginning of the new century, Southern regions have experienced an absolute decline in the size of their population, also in part due to significant outmigration of young educated workers to the Centre-North and abroad (De Angelis *et al.*, 2017; Ballatore and Mariani, 2019). One possible reason for this persistent disparity is the slowdown in the process

of accumulation of capabilities. Over time this has resulted in more modest endowments of capabilities in the Southern regions, with a negative impact on their economic performance, on their growth potential (Boltho *et al.*, 2018) and productivity distribution (Basile *et al.*, 2019).

The main difficulty with the empirical analysis of capabilities for economic growth is that they are unobservable. It is therefore necessary to develop a methodology to obtain information about them from observable variables. Hausmann, Hwang and Rodrik (2007) have devised a clever approach to infer the unobservable capabilities of an economy from a country's export specialization profile. The rationale behind this approach is that the production of complex goods requires the existence of a broad set of advanced skills, as well as the ability to combine them effectively. The production of complex goods therefore conveys relevant information about the underlying, unobservable capabilities in the economy. Analysis of Revealed Comparative Advantage (RCA) in export profiles using methods of complex network analysis can identify those countries and regions which specialise in complex products. This information constitutes a reliable indicator of the underlying set of capabilities in the economy and can help explain the potential for growth in the economy.

Two main measures for the complexity of an economy based on a country's revealed comparative advantage profiles have been proposed: the Economic Complexity Index (ECI) introduced by Hidalgo and Hausmann (2009), and the Economic Fitness and Complexity metric (EFC) proposed by Tacchella *et al.* (2012). In this paper we rely on the EFC framework, which is the optimal expression for economic complexity in terms of information theory as established by several contributions in the literature (Caldarelli *et al.*, 2012; Tacchella *et al.*, 2013; Pietronero *et al.*, 2017). The validity of measures of complexity as predictor of growth has been confirmed, among others, by Hausmann *et al.* (2007); Tacchella *et al.* (2012); Ferrarini and Scaramozzino (2016); Cristelli *et al.* (2017); Sbardella *et al.* (2018).

By building on dynamical system theory, Cristelli *et al.* (2017) developed the Selective Predictability Scheme (SPS), an original methodology that makes use of the Fitness metric and ideas from dynamical systems theory for producing more realistic macroeconomic forecasts. SPS is used to forecast the economic evolution of countries by looking at the trajectories they follow in a bidimensional Fitness–GDP per capita plane. In fact, by using only monetary figures, it is easy to mistakenly assess the development stage of a country: only when the “quantitative” information of GDP is integrated with the “qualitative” one on country capabilities conveyed by Fitness, does it become possible to disentangle heterogeneous development paths otherwise seen as similar. The technique was later revisited by (Tacchella *et al.*, 2018) and their long-run growth outlooks outperformed the accuracy of IMF growth forecasts by over 25%. Pugliese *et al.* (2017) show that Economic Fitness plays a key role at the onset of the industrialization process, and empirically prove that more complex economies face lower GDP per capita barriers when starting the transition towards industrialization. Sbardella *et al.* (2017) show

that complexity can also help explain the relationship between wage inequality and economic development at local levels. In the specific context of comparative regional growth, Boltho *et al.* (2018) argue that a key explanation for the rapid convergence of East Germany to the West following unification, in contrast to the sluggish performance of the Italian Mezzogiorno relative to the Centre-North, may have been the more complex economic structure of the former East German *Länder*.

The field of research, with its interdisciplinary scope, is supported by a very active scientific community that is reshaping its boundaries by extending the possibilities of analysis to more and more areas, and various indices of economic complexity have been adopted by international institutions, such as the World Bank, the European Commission and the OECD, and by national and local governments. In particular, in recent years, one of the most successful fields of application of the EC framework has been the study of local or national or local innovation systems, proving to be particularly effective in quantifying information on technological capabilities at various levels of aggregation. Looking through the lenses of EC at the geographical distribution, quality and relatedness of the innovative activities into which economic actors specialize, as proxied by patent data, allows one to characterize firms (Breschi *et al.*, 2003; Nesta and Saviotti, 2005; Pugliese *et al.*, 2017), as well as regions or cities (Boschma *et al.*, 2013; Balland and Rigby, 2017), recently also in relation to environmental technologies (Sbardella *et al.*, 2018; Napolitano *et al.*, 2019).

By embracing the structural approach brought back to the fore by the economic complexity literature, that identifies the accumulation of capabilities as an essential element in the development process of local areas, the present paper applies the Economic Fitness methodology to the analysis of Italian regions. Bearing in mind the benefits and the shortcomings of using patent data for the study of technology development (see e.g., Arts *et al.*, 2013; Griliches, 1998; Lanjouw and Mody, 1996), we propose two regional complexity measures based respectively on international trade and patent data, Export Fitness and Technology Fitness, to analyse in detail the productive and technological capability structure of Italian regions. Furthermore, we define the Product Progression Network (Albora *et al.*, 2021) of Italian regions, a network methodology that allows us to map the relatedness of different regional productive structures.

Methods and Data

Regional Productive and Technology Economic Fitness

The basic intuition of the Economic Fitness and Complexity (EFC) approach adopted in this paper is that specific products or technologies are important because they constitute different learning opportunities and therefore development possibilities. EFC explicitly builds on the

heterogeneity and interactions between different economic actors, assuming that the level of technological and scientific knowledge of a geographical area – be it a country, a region or a city – cannot be reached at an intensive level: knowledge grows not by accumulating “more”, but by adding new and different elements to existing capacities.

In a nutshell, the EFC is a recursive algorithm based on the observation of an empirical network connecting countries to the products they export with a Revealed Comparative Advantage (RCA) (Balassa, 1965) greater than 1. The output of this algorithm defines two indexes: Fitness (F), an intensive metric for the competitiveness of the productive system of countries, and Complexity (Q), a measure of the sophistication of the products they export.

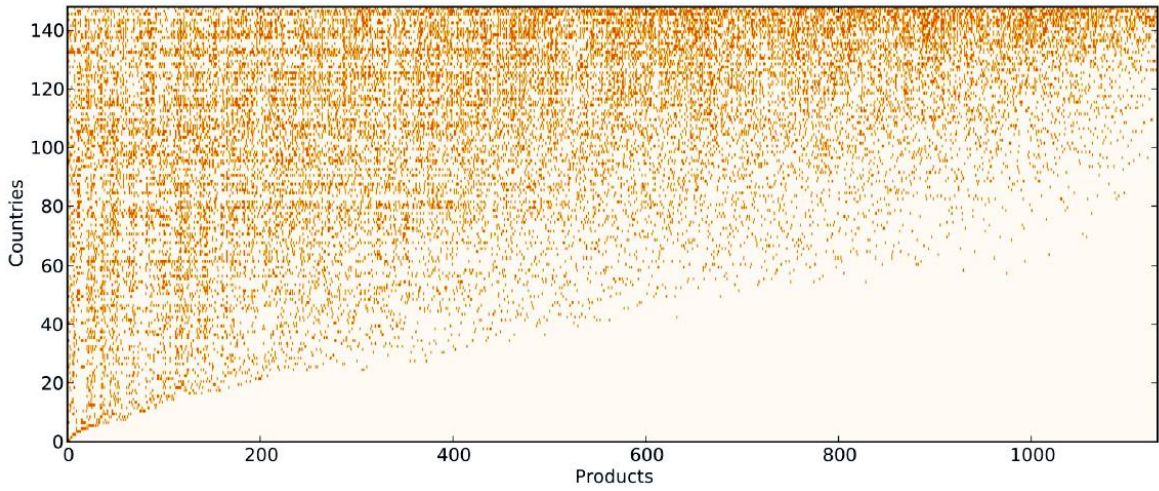


Figure 1: The triangular matrix describing the Country-Product network on which the Economic Fitness and Complexity algorithm is based.

As can be observed in Figure 1, the adjacency matrix of the country-exported product network shows a nested structure, i.e. a triangular shape. This suggests that, contrary to the specialisation predicted by Ricardian or Heckscher-Ohlin models, fitter countries are highly diversified and able to export competitively all products, from the least to the most complex, whereas less competitive countries are able to export only few, ubiquitous, and low-complexity products. To capture this nested structure, the Complexity of products is bounded by the Fitness of its least competitive exporters, while the Fitness of countries is defined as complexity-weighted diversification. In formulae, for country c and product p :

$$\left\{ \begin{array}{l} \tilde{F}_c^{(n)} = \sum_p M_{cp} Q_p^{(n-1)} \\ \tilde{Q}_p^{(n)} = \frac{1}{\sum_c M_{cp} \frac{1}{F_c^{(n)}}} \end{array} \right. \left\{ \begin{array}{l} F_c^{(n)} = \frac{\tilde{F}_c^{(n)}}{\langle \tilde{F}_c^{(n)} \rangle_c} \\ Q_p^{(n)} = \frac{\tilde{Q}_p^{(n)}}{\langle \tilde{Q}_p^{(n)} \rangle_p} \end{array} \right. \quad (1)$$

where $\langle \cdot \rangle_x$ denotes the arithmetic mean with respect to the possible values assumed by the variable dependent on x , and the initial condition is $\sum_p Q_p^{(0)} = 1 \quad \forall p$. The fixed point of Equation 1 defines the non-monetary metric that quantifies F_c , the fitness of country c , and Q_p , the complexity of product p .

Depending on the structure of the input matrix, the EFC algorithm is known to possibly converge to zero Fitness and zero Complexity for a subset of geographical areas and activities respectively (Pugliese *et al.*, 2016). However, this is not an issue because it is always possible to define a consistent ranking along both dimensions.

Nestedness is not only apparent in the international trade network; on the contrary, is a pervasive characteristic of different socio-economic and natural environments (Bustos *et al.*, 2012; Napolitano *et al.*, 2018; Mariani *et al.*, 2019). Hence, the Fitness and Complexity algorithm, which is purposely designed to capture nested patterns, may be successfully applied to other domains of human activity that can be described in terms of location-activity networks. As mentioned above, since its inception, the literature on economic complexity has expanded its scope and object of analysis, and it now comprises a diverse set of studies that make use of complex system methodologies, pay particular attention to disaggregated dynamics and emerging properties, are able to analyze and measure the intrinsic complexity of economic systems at different geographical scales and in different areas of human activity at macro- and micro-levels.

Drawing from different applications of the EFC metric to local areas (Operti *et al.*, 2018; Sbardella *et al.*, 2017) and from various definitions of complexity indexes based on patent data at local (Balland and Rigby, 2017) or national levels (Pugliese *et al.*, 2019; Sbardella *et al.*, 2018), in the present paper we reframe capabilities as local attributes in order to analyse the technological and industrial capability profiles of Italian regions (NUTS2 level), and propose two measures of regional Export Fitness and Technology Fitness. These two generalizations of the EFC algorithm are based on bipartite networks connecting Italian regions to respectively the product they export, or the technology fields in which they patent with $RCA > 1$.

In practice, with the aim of providing a more realistic assessment of productive and technological competitiveness, we employ the Exogenous Fitness methodology, an approach proposed by Operti *et al.* (2018) for calculating the Fitness of sub-national entities. Thus, in the Fitness computations instead of computing the complexity of a product or technology only on the Italian regional subsample we insert the complexity obtained considering international trade and global patent data.

Product Progression Network

To analyse in detail the characteristics of the Italian regional productive structure we rely on another technique from the Economic Complexity toolbox and construct the Product Progression Network profile of regional exports. We draw from recent contributions on the product space (Hausmann and Klinger, 2006; Hidalgo *et al.*, 2007) and in particular from Zaccaria *et al.* (2014) and Pugliese *et al.* (2019), who employ product-level export data to measure the relatedness of products through statistically significant patterns of co-exporting in the international trade network and argue that countries that are able to successfully export a product have developed a set of capabilities that would enable them to diversify into related goods. These analytical tools can be traced back to Jaffe (1986) and to the measure of corporate coherence introduced by Teece *et al.* (1994). A more detailed description of this methodology can be found in Zaccaria *et al.* (2018). The Product Progression Network connects related products in the export basket of the geographical area taken into consideration with a directed link, thus making it possible to trace the most profitable trajectories for entering into a new production line on the basis of the pre-existing endowment of productive capabilities.

This measure of relatedness, or proximity, between two products is based on the observation of their empirical co-occurrences in the export basket of different countries, and is connected to the probability that having a comparative advantage in the first product will also lead to a comparative advantage in the second (see Hidalgo *et al.*, 2007, but see also Tacchella *et al.*, 2021 for a multi-product, non-linear approach based on machine learning). With respect to previous efforts, the Product Progression Network methodology has two main advantages: i) it is dynamical, as it takes explicitly into account time by comparing time-delayed co-occurrences, instead of contemporaneous ones; ii) it filters each link of the resulting network using a suitable null model, so that only the statistically significant links are observed. Indeed, the statistical significance of a co-occurrence is a key issue, because the presence of a link may be simply due to the ubiquity of a product or to the diversification of a country. In order to take into account this effect, we adopt as null-model the Bipartite Configuration Model (Saracco *et al.*, 2015, 2017), a novel maximum-entropy algorithm for the randomisation of bipartite networks. Intuitively, products that share similar inputs will be situated close to each other in the network and, once we have filtered the empirical co-occurrences with the null model, proximity indicates a relatively high probability of jumping from a product to a neighboring one. Therefore, its observation allows us to trace the most profitable trajectories for entering into a new production line on the basis of the pre-existing endowment of productive capabilities.

In this paper we thus propose a version of the Product Progression Network for Italian regions that, by contrast to the product space based on the observation of Italian provinces proposed by Cicerone *et al.* (2020), is able to filter for product ubiquity and diversification of

the geographical areas considered, as well as to avoid the occurrence of a possible simultaneity bias in the observation of the empirical co-occurrences.

Data

International Export

The international export data employed to compute product complexity are taken from the UN-COMTRADE database², which provides the product-level trade flows between countries in millions of US-Dollars. Products are organised according to the Harmonized System (HS), a hierarchical classification that goes from a two-digit (about 100 different products) to sixdigit aggregation level (about 5000 different products). Since the declarations of importers and exporters do not always fully coincide, a data sanitation procedure is required to obtain a harmonised and coherent dataset. To overcome this issue, we employ the denoised dataset produced by Tacchella *et al.* (2018) using a global Bayesian optimization approach that, as mentioned above, has allowed a notable increase in the predictive power of the SPS methodology.

Regional Export

For assessing the export profiles of Italian regions, we employ the *ISTAT-Esportazioni per ripartizione territoriale e regione* data-set³ that comprises regional accounts of aggregated exports in millions of Euros. Furthermore, to calculate our regional Export Fitness measure we use a more detailed ISTAT data-set for identifying the products present in the export basket of each region, that are classified according to the Harmonized System.

Regional Patenting Activity

To analyse the patenting activity of Italian regions we employ patent data from PATSTAT 2020a (European Patenting Office, 2020). PATSTAT is published biannually and contains several tables linking over 100 million patents filed at over a hundred national and regional patent offices to the patent families they belong to, their filing date, inventor's address, and technological content as described by Cooperative Patent Classification (CPC) codes attributed to the application. To build our data-set first we geolocalise patents by assigning them to the NUTS3 2013 regions of residence of their inventors exploiting PATSTAT and de Rassenfosse's Geocoding of worldwide patent data (de Rassenfosse *et al.*, 2019). Second, we associate to each

² <https://comtrade.un.org>

³ <https://www.istat.it/it/archivio/esportazioni+delle+regioni>; https://www.istat.it/it/files//2017/03/comreg_new_IV_trim2016.pdf

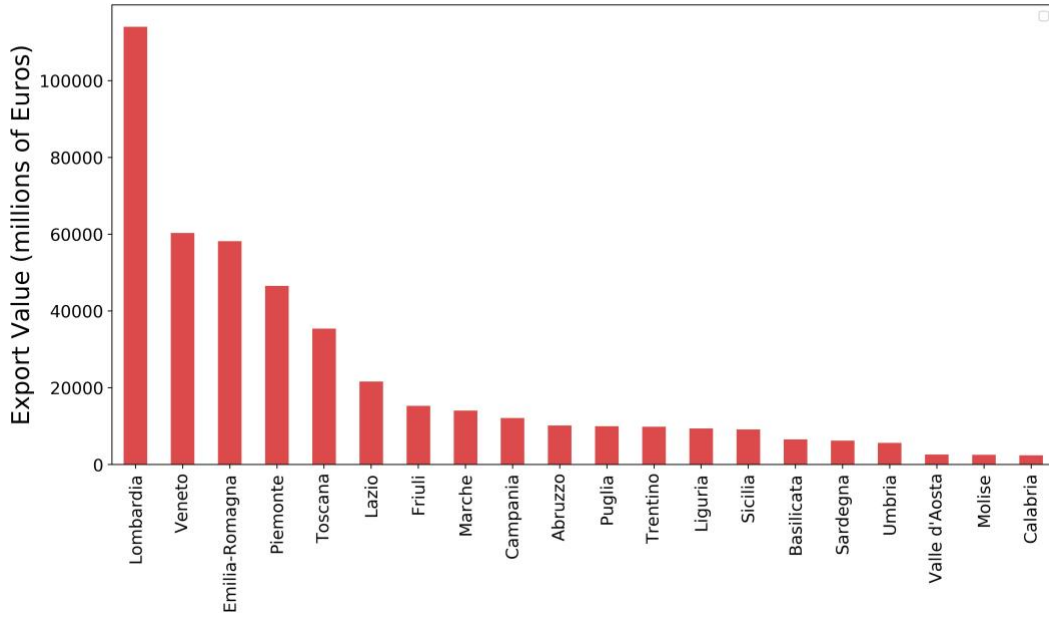
patent a CPC code and group patents in INPADOC patent families, each of which represents a collection of patent documents covering a technology. Third, we define a fractional share of innovative contributions and assign it to each technological code and Italian NUTS2 region present in the family by following a similar strategy to Napolitano *et al.* (2019).

Results

The Export and Technology Fitness of the Italian regions is described in detail for 2016, which is the most recent year in our sample. We however also conduct a dynamic analysis of the evolution of the exporting and patenting performance of the regions since 1998. It is apparent from the data that the regional units display significant heterogeneity according to all the metrics considered. As can be observed in Figure 2 (a), Italian exports are strongly dominated by Lombardy, which is leading with numbers that are orders of magnitude higher than most of the other regions. The second group found immediately after in the bar-plot comprises Veneto, Emilia-Romagna and Piedmont, the backbone of the Italian productive system, which is followed by Tuscany and Lazio, the only two Central regions that exhibit large export volumes.

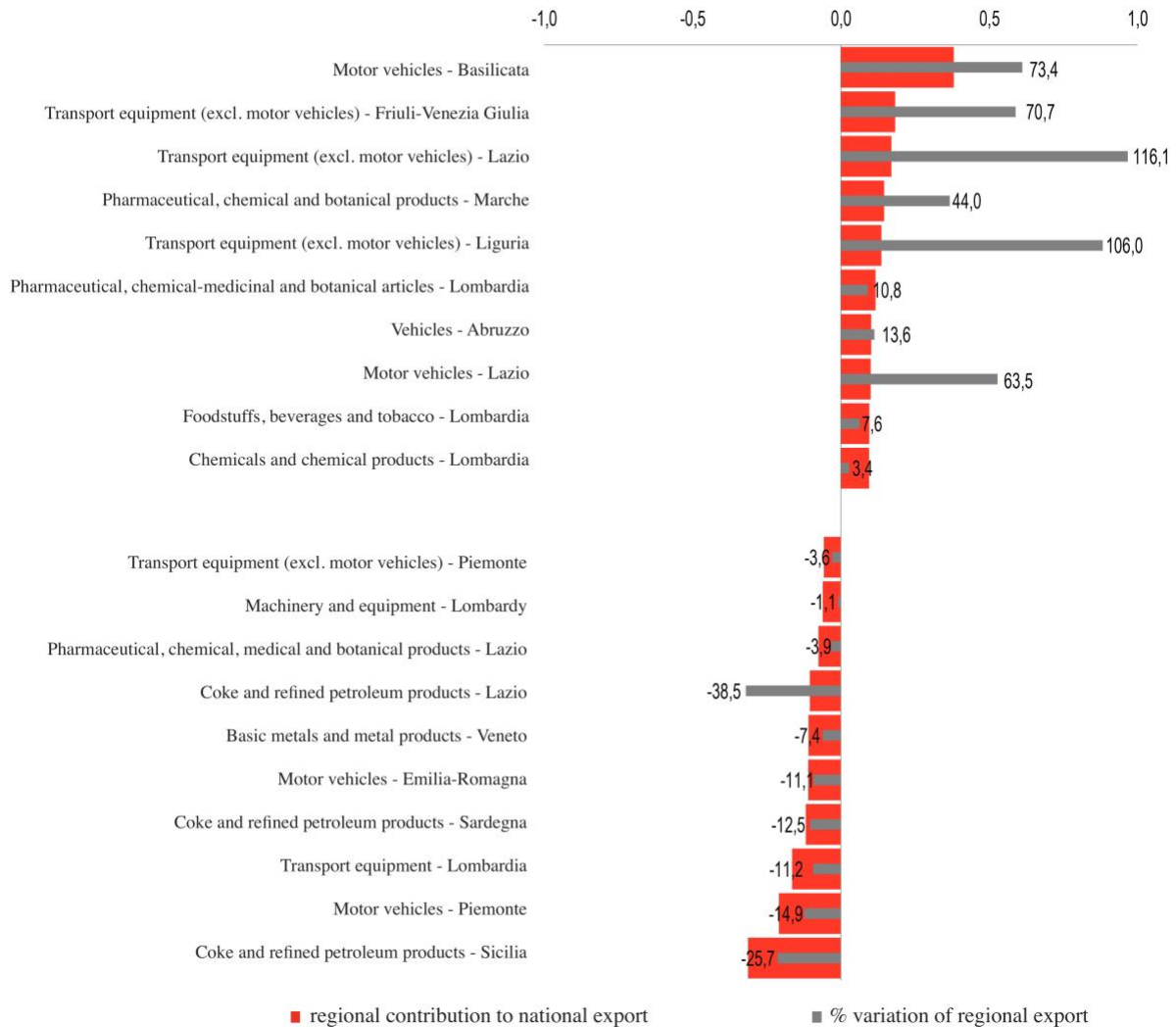
On the one hand, this pattern is consistent throughout the 1998-2016 observation period, during which an overall increase in national exports (+1.2%) was registered. Among the regions that provide the largest positive contributions we find Basilicata, Liguria, Emilia-Romagna, Trentino, Sardinia and Lazio, while the Northern competitive core shows a more moderate growth, and the Southern regions together with Aosta Valley present the lowest export growth. On the other hand, as can be noticed from Figure 2 (b), when looking only at the last year in our time window, the picture is more complex. The increase registered in national exports from January to December 2016 reflects both the increase recorded for Southern (+8.5%), Central (+2.1%), and North-Eastern (+1.8%) regions and the large drop in the Insular area (-15.0%), while the North-Western area is stationary. The joint analysis by sector and region of origin of the goods shows that in 2016 an increase in sales of motor vehicles from Basilicata (+73.4%), means of transport (excluding motor vehicles) from Friuli-Venezia Giulia (+70.7%), Lazio (+116.1%) and Liguria (+106.0%), as well as in pharmaceutical, chemical-medicinal and botanical articles from Marche (+44.0%). In the same period, the sharp decrease in exports of refined petroleum products from Sicily (-25.7%) and motor vehicles from Piedmont (-14.9%) had a negative contribution to domestic sales on foreign markets.

Italian Regional Export (2016)



(a)

Contribution to change in Italian export by region and sector of activity (2016)



(b)

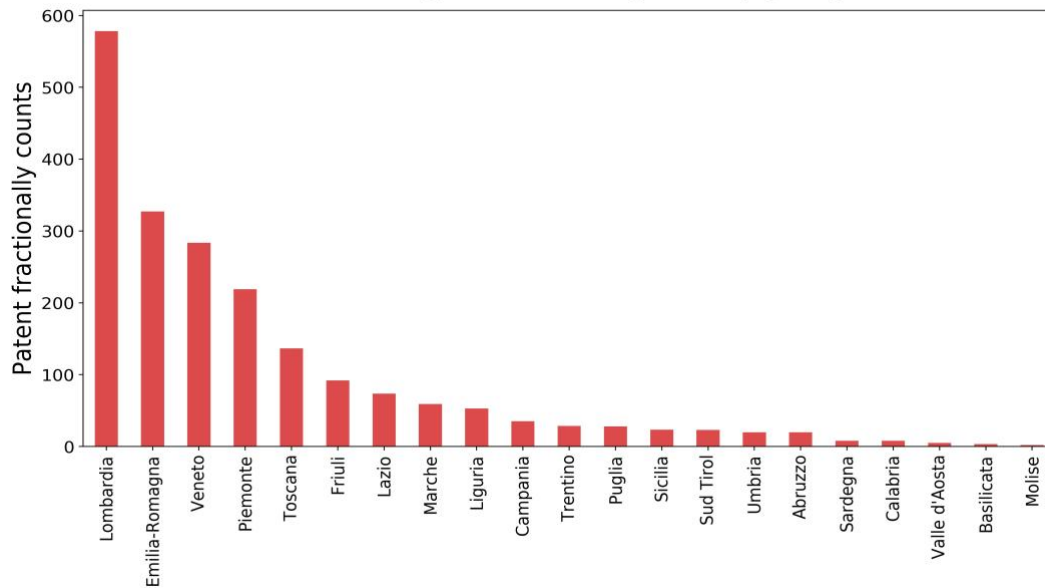
Figure 2: (a) Italian regional export volumes in 2016, reported in millions of Euros. (b) Variation and contribution to change in national export by sector of activity and region in 2016, reported in percentage values. Source of the figure: https://www.istat.it/it/files/2017/03/comreg_new_IV_trim2016.pdf.

The scenario for patenting activity is not dissimilar: in terms of volumes (see Figure 3 (a)), Lombardy still dominates and is followed, albeit at a considerable distance, by Emilia-Romagna, Veneto and Piedmont, whose activity ranges between 56% and 38% of that of Lombardy; the percentages go down to between 25% and 10% for Tuscany, Friuli-Venezia Giulia, Lazio, Marche, Liguria and between 5% and 0.3% for Apulia, Sicily, Umbria, Aosta Valley, Basilicata and Molise. Whilst the geographical gap between Northern and Southern Italy has remained unchanged over our period of observation, also in this case, we witness a general increase in the number of patents filed in Italy and the variability of these numbers is greater than that of exports. While little variation in the figures of the top competitive areas is observed, a growth in the patenting activity of various Southern regions, in particular Apulia, Calabria, Molise and Campania is recorded, whereas Sicily and Lazio are the only two regions that show negative growth and Abruzzo displays almost zero growth. As Figure 3 (b) shows, if we disregard the ubiquitous Section A which contains a miscellany of different patent types, when looking at the composition of the patenting activity across regions in 2016, we can observe a considerable diversification across CPC sections in almost all regions – albeit in different ranges. Lombardy, Emilia-Romagna, Veneto and Piedmont are mainly focused on Chemistry (C), Mechanical Engineering (F), Performing Operations and Transport (B) and, to a lesser extent, on Physics (G), Electricity (H) and, in even lower proportions, in New Cross-sectional Developments (Y), while they show almost no contribution in Textiles and Paper (D). The latter is significantly present especially in Friuli-Venezia Giulia and Marche. In central Italy we find an important share of patents in chemistry and mechanical engineering, especially in Tuscany and Umbria, while in Lazio a significant role is played also by physics and electricity. The combination of physics, electricity and new cross-sectional developments – which also include new environment-related patents mainly associated to renewable energies and climate change mitigation technologies – characterises the Aosta Valley, Molise, Basilicata and the insular regions. The latter two, together with Apulia, also display a significant presence of chemistry, although in Apulia the prevalent technological area is mechanical engineering.

In Figure 4 and Figure 5 we explore the dynamics and geographical distribution of our two measures of Export and Technology Fitness across Italian regions between 1998 and 2016. To visualise the geographical patterns of productive and technological capabilities, panel (a) of both figures display an Italian regional map coloured according to the respective fitness

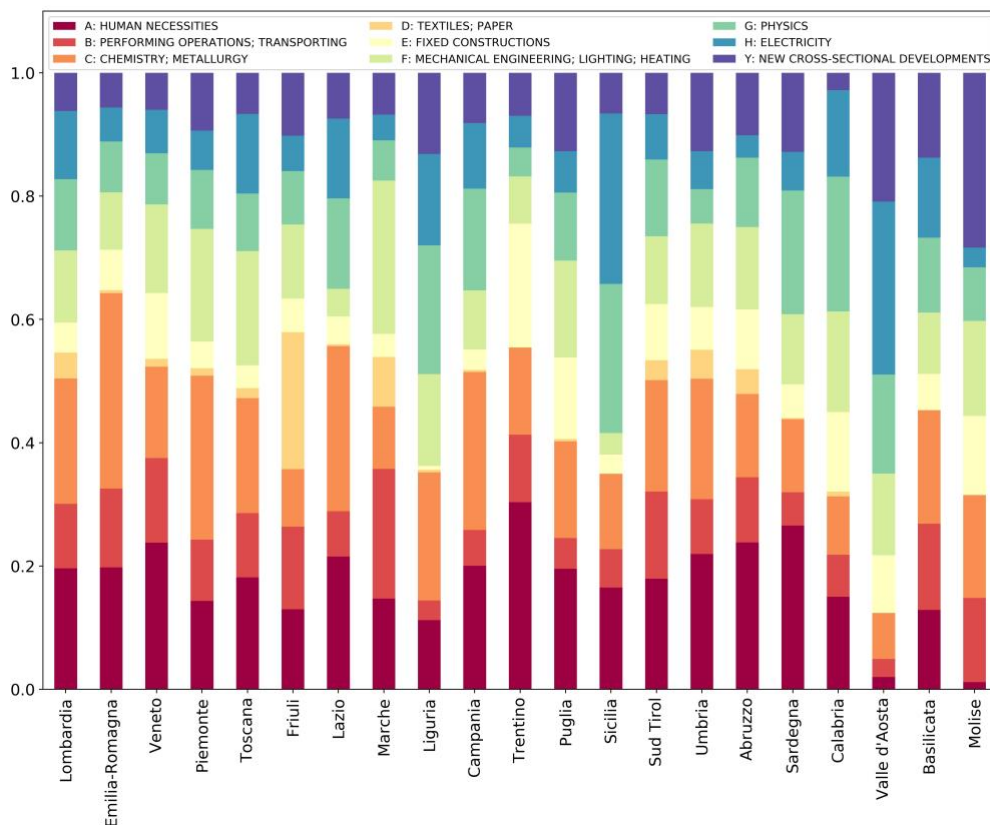
ranking at the start and the end of the period of observation, the darker the red the higher the Fitness.

Italian Regional Patenting Activity (2016)



(a)

Italian Regional Patenting Activity Composition (2016)

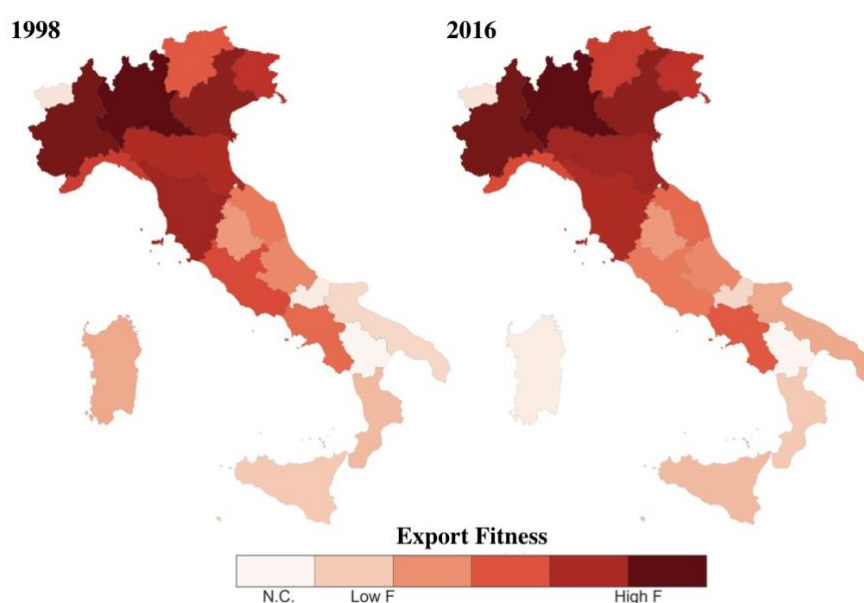


(b)

Figure 3: (a) Patenting activity in Italy by region in 2016. (b) Patenting activity composition across CPC sections in Italy by region in 2016.

In order to better appreciate the dynamics, in panel (b) of both figures we show the time evolution of the regional rankings over the three time periods 1998-2004, 2005-2010 and 2011-2016. The colour of the label and curve associated to each region reflect the initial ranking in 1998: regions that start in the top third feature a dark red label; regions in the middle third a light red label; regions in the bottom third are labeled in salmon. It is important to notice that, being Fitness a relative metric, when a region reaches a higher position it always does so at the “expense” of another region.

From the colour distribution in Figure 4 (a), a stark North-South dichotomy and lack of convergence in terms of Export Fitness emerge clearly. In panel (b), where the time evolution of the ranking is shown, we can immediately notice that the gap is aggravated by its persistence over time. This should not be surprising: in fact, we observe very little variation in colour between the left and right vertical axis and the top five positions are consistently occupied by Lombardy, Piedmont, Veneto, Emilia-Romagna and Tuscany. It is worth emphasising the monotonic loss of competitiveness of Lazio and the even more dramatic fall of Sardinia, which starts from an intermediate position but at the end of the period is second-to-last in the ranking. Together with the central regions Umbria and Marche, we find Abruzzo and Campania, the latter being the Southern region with the highest endowment of productive capabilities showing a fairly stable intermediate position in the ranking. In the last positions of the ranking, we find consistently almost all Southern regions, Calabria, Sicily, Apulia, Molise, and Basilicata, with the exception of the Aosta Valley which is the only Northern region to show a persistent low level of Export Fitness.



(a)



(b)

Figure 4: (a) Italian regional map coloured according to Export Fitness levels in 1998 and 2016. (b) Export Fitness ranking time evolution from 1998 to 2016 by Italian region.

By contrast, the Italian technological landscape has been far from static in the past twenty years. Indeed, in Figure 5 (a) the North-South divide is still evident, but the colour distribution is more animated, as also suggested by the color mixing of the labels on the right vertical axis in Figure 5 (b) which shows that the race to the top has seen some new entries, some of which coming from far behind. The top innovative core still appears unaltered: Emilia-Romagna maintains the highest position followed by Lombardy and Piedmont, which remain solidly in the top five. By contrast, Veneto improves its relative position and, starting from the ninth position, enters the top five in 2016.

In this case, in addition to the leading role of Emilia-Romagna, the regions of Central Italy, especially at the beginning of the period, are better positioned with respect to the Export Fitness ranking. In fact, in 1998, Lazio, Tuscany and Marche show high Technology Fitness; however, while Tuscany and Emilia-Romagna maintained their role as top players, Lazio and Marche lost their relative importance and ended up occupying intermediate positions in the ranking, leaving room, in particular, for Campania, Apulia and Sicily, which gained ground and in 2016 show high levels of fitness. Remarkable are the upward shifts of Apulia which, through a fairly eventful path, in 2016 reaches the sixth position, and of Campania which, starting as second last in 1998, is seventh. This suggests that the specialisation profiles of these regions have gradually moved to more complex technologies and that to increase their Technology Fitness they could be aiming at making jumps into related technologies on the basis of their existing knowledge base.

To quantify the differences in the movement of the ranking of Export and Technology Fitness, we can measure the overall degree of mobility among Italian regions in terms of the

Markov transition probabilities across the quintiles of the distribution. Table 1 (a) shows the transition matrix for the quintiles for Export Fitness from 1995 to 2016, and Table 1 (b) the corresponding transition matrix for Technology Fitness. The largest values of the transition probabilities for the Export Fitness are on the main diagonal of the matrix, indicating a high degree of persistence in the quintile ranking. The values on the main diagonal are much lower for Technology Fitness, suggesting a greater degree of mobility among the regions over the period considered.

	Q1	Q2	Q3	Q4	Q5		Q1	Q2	Q3	Q4	Q5
Q1	0.50	0.25	0.25	0.00	0.00	Q1	0.75	0.25	0.00	0.00	0.00
Q2	0.25	0.50	0.25	0.00	0.00	Q2	0.00	0.00	0.50	0.50	0.00
Q3	0.25	0.25	0.50	0.00	0.00	Q3	0.25	0.25	0.25	0.00	0.25
Q4	0.00	0.00	0.00	0.75	0.25	Q4	0.00	0.25	0.00	0.00	0.75
Q5	0.00	0.00	0.00	0.25	0.75	Q5	0.00	0.20	0.20	0.40	0.20

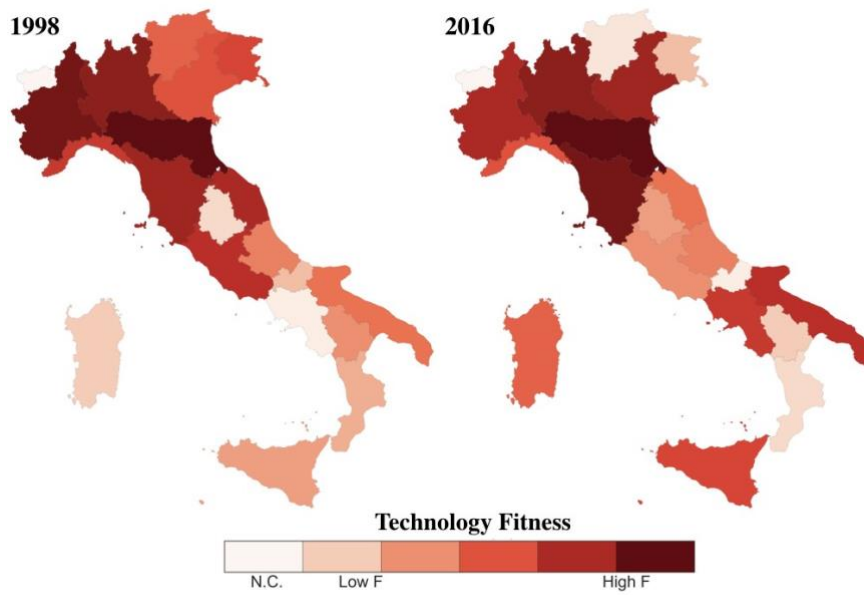
(a) (b)

Table 1: (a) Transition matrix for the quintiles for Export Fitness from 1995 to 2016. (b) Transition matrix for the quintiles for Technology Fitness from 1995 to 2016.

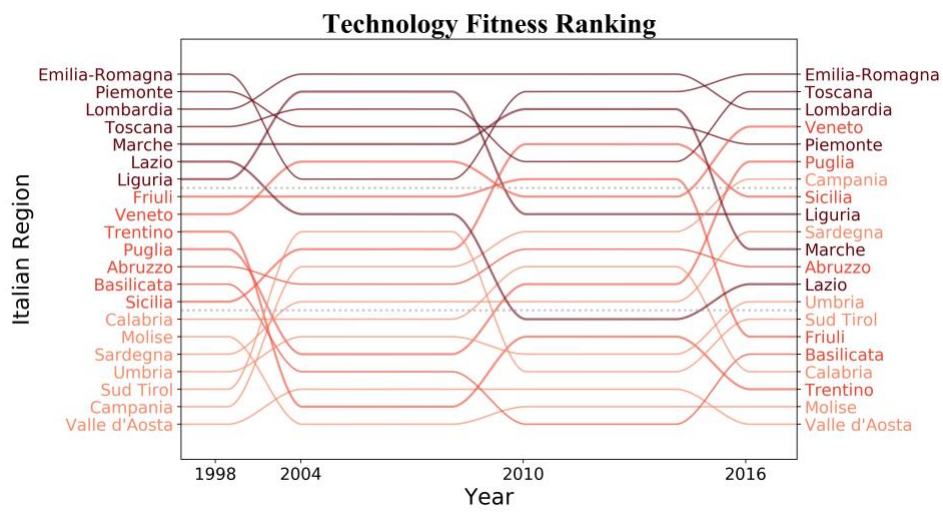
Formal measures of mobility can be obtained from the indicators M^1 and M^2 . These measures were originally proposed by Shorrocks (1978) and are defined as follows:

$$\begin{cases} M^1 = \frac{K - \sum_i \lambda_i}{K - 1} \\ M^2 = 1 - \prod \lambda_i \end{cases} \quad (2)$$

where K is the order of the transition matrix and λ_i are its eigenvalues. The indicator M^1 is based on the relative magnitude of the diagonal and off-diagonal elements of the matrix, whereas M^2 is based on the determinant of the transition matrix. High values of the indices reveal a large degree of mobility across the fitness ranking. The value of M^1 is equal to 0.5 for Export Fitness and to 0.950 for Technology Fitness. There is therefore much greater inertia in the profile of regional comparative advantage in export than there is for technological fitness. The corresponding values for M^2 are 0.969 for Economic Fitness and 0.986 for Technological Fitness: this indicator therefore provides a less sensitive measure in this context. Jointly considered, though, both indices reveal that the relative degree of Technological Fitness has varied considerably among the regions of Italy over the period considered.



(a)



(b)

Figure 5: (a) Italian regional map coloured according to Technology Fitness levels in 1998 and 2016. (b) Technology Fitness ranking time evolution from 1998 to 2016 by Italian region.

Export Fitness-GDPpc Country trajectories (1999-2016)

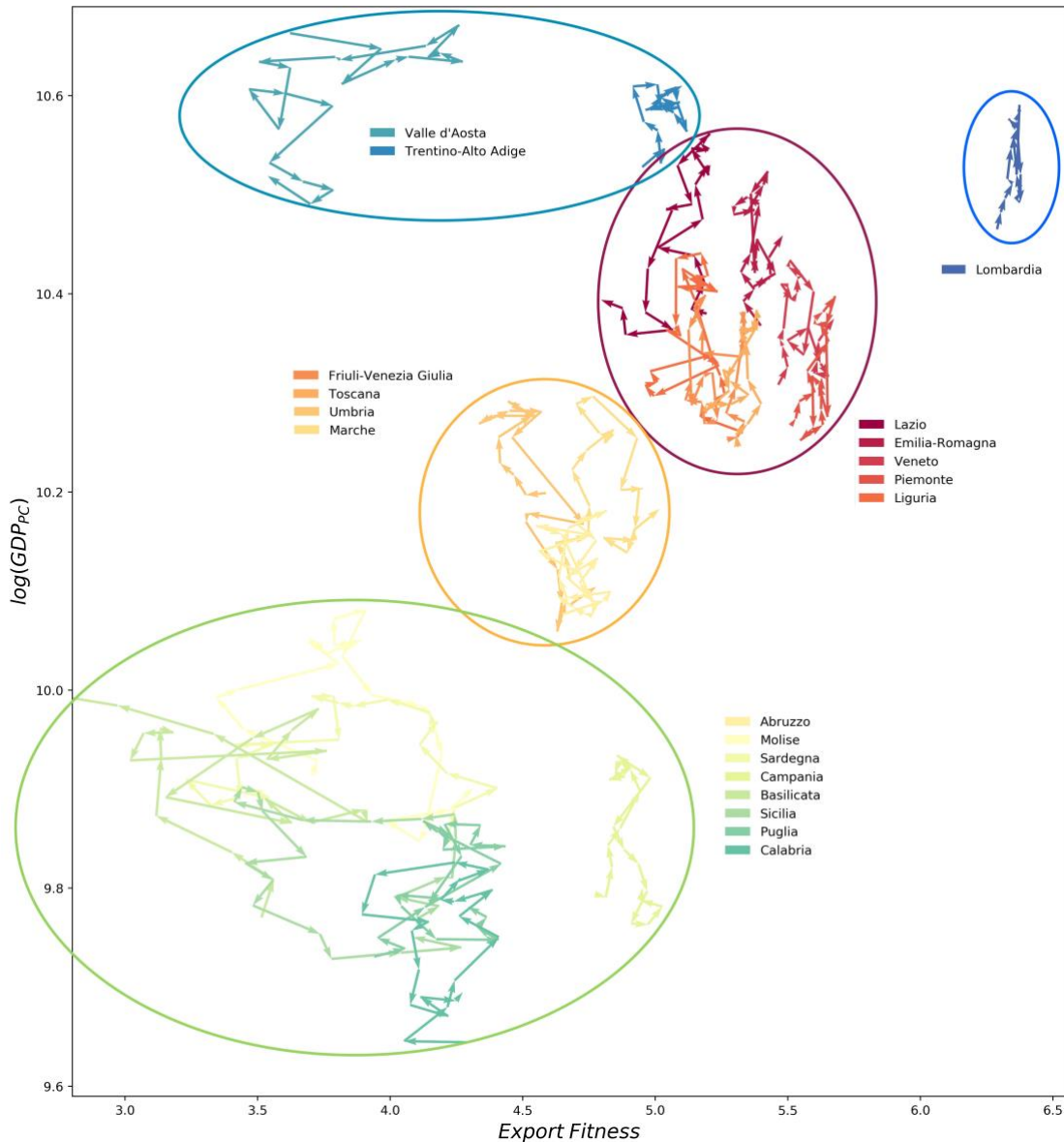


Figure 6: Trajectories of the Italian regions in the Export Fitness–(log) GDP per capita plane over the period 1999-2016.

Figure 5 shows the trajectories of the Italian regions in the Export Fitness-GDP per capita plane (Cristelli *et al.*, 2015) over the period 1999-2016. As analysed above, for most regions, Export Fitness did not change much over the period considered. GDP per capita tended to increase in the first half of the period, but then decreased after the financial crisis of 2008; for several regions, GDP per capita was actually lower, or not much higher in 2016 than it was in 1999.

Whilst the role of Fitness as a predictor of GDP growth cannot be clearly established from the graph because of the recessionary consequences of the financial crisis and of its aftermath, Figure 6 reveals a number of homogeneous regional clusters that group regions with similar

endowments of productive capabilities, and that strongly confirm the North-South divide. Lazio, Emilia-Romagna, Veneto, Piedmont, and Liguria in the Centre-North are characterised by high values of fitness and GDP per capita, whereas Abruzzo, Molise, Sardinia, Campania, Basilicata, Sicily, Apulia and Calabria in the South show low levels of fitness and of GDP per capita (it should though be noted that Campania exhibits higher levels of Fitness than the other regions in this group). Friuli-Venezia Giulia, Tuscany, Umbria and Marche present intermediate levels of both variables. Lombardy stands out with his very high levels of both indicators. Trentino-Alto Adige and the Aosta Valley are somewhat exceptions, with higher levels of GDP per capita than would be warranted by their Export Fitness. All in all, the figure supports a positive association between Export Fitness and per capita income, although we also observe some interesting non-linear deviations.

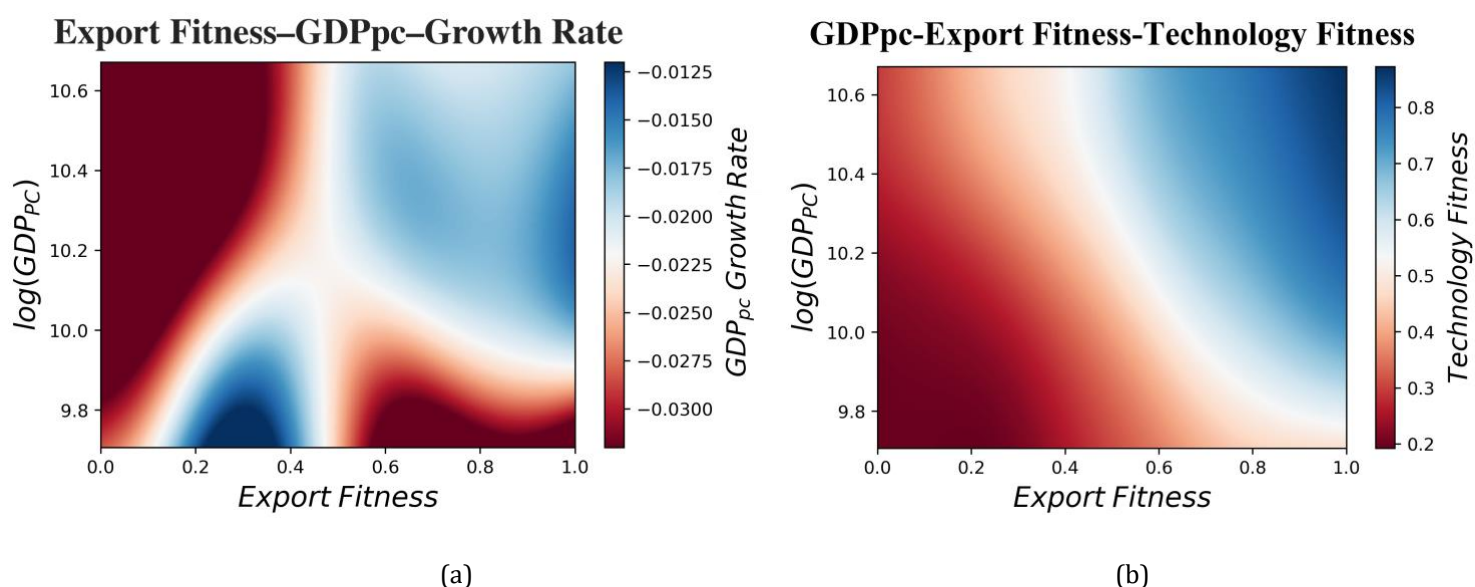


Figure 7: The three-dimensional relationship between Export Fitness, GDP per capita, and (a) the growth rate or (b) Technology Fitness of Italian regions. The colour maps are obtained with a non-parametric Nadaraya-Watson kernel estimation by pooling all regions over the time interval 1998–2016.

Therefore, to explore the dynamic and possibly non-linear relationship between Export Fitness, income per capita and economic growth as well as Technology Fitness, we propose two non-parametric graphical analyses which do not assume any *a priori* functional forms for the relationships. The plots in Figure 7 represent a colour-map of the relationship between Export Fitness on the x-axis, the logarithm of GDP per capita on the y-axis, and a non-parametric estimate of GDP per capita growth rate on the z-axis – panel (a) – or of the Technology Fitness ranking – panel (b) – represented through a red-to-blue colour-map. The colour-maps are obtained via a multivariate Nadaraya-Watson regression (Nadaraya, 1964), a continuous non-

parametric method, with a Gaussian kernel. We pool all the countries in the time interval 1998-2016 and, to avoid the risk of a potential simultaneity bias, the explanatory variables GDP and Export Fitness are lagged five years relative to Technological Fitness or the regional growth rate, the dependent variables on the z-axis.

The indicator of Export Fitness helps predict the future growth rate of the regional economy. Figure 7 (a) illustrates how log GDP per capita and Export Fitness are related to the future GDP growth rate. The initial level of income does not seem to play a clear role: GDP growth appears to be unrelated to initial income. If anything, there may be evidence of a weak negative association between income per capita and future growth rate which may be consistent with some forms of catching up of the poorer regions to the richer ones. By contrast, initial Export Fitness tends on the whole to be associated with faster rates of growth of the local economy: greater production complexity can act as a spur to economic activity, irrespective of the initial level of income in the region.

Figure 7 (b) shows how Technology Fitness is related to the log of GDP per capita and to Export Fitness. An interesting feature of the plot is the diagonal movement of colour, which hints at a complementarity between the roles of initial Export Fitness and income per capita levels in unleashing innovation capacity. Higher levels of both GDP per capita and Export Fitness tend to be associated with higher future values of Technology Fitness: the research performance of regions in the production of patents is therefore related not simply to the overall level of economic activity in the region, but also to the degree of complexity of the underlying production structure as measured by the Export Fitness. The latter indicator thus contains important information which can help predict the future research performance of the region.

With the aim of shedding light on possible opportunities for diversification and possible bottlenecks, we show the patterns of inter-sectoral connections of the regional productive structures in Figure 8, where we display the Product Progression Network analysis of four significant regions: Lombardy (a), Veneto (b), Campania (c), and Sicily (d). As specified in the Methods and Data section, the network is built to show the connections between industrial sectors and, in particular, their time-delayed, statistically validated co-occurrences. Two industries p and p^0 are connected if countries show a revealed comparative advantage in p and then, after a fixed amount of time, in p^0 , and if such co-occurrences cannot be explained by industries' ubiquity and countries' diversification. While the network structure is defined by the co-occurrences at the national level, the node coloring depends on the revealed comparative advantages of each selected region. This serves the purpose of graphically depicting both the regional export basket as a whole and the possible future specialization steps. In particular, for each region we use three shades of red to depict low ($0 < RCA < 0.5$), medium-low ($0.5 < RCA < 1$), or high RCA ($RCA > 1$).

Firstly, we notice the presence of three main significant clusters: in the bottom left corner of the network we can find textile-related sectors; in the top left corner are concentrated products related to agrifood industries; while, the top right corner is characterised by a prevalence of high-tech products. Secondly, at a glance, it is possible to observe the extremely higher diversification in the export baskets of the two Northern regions with respect to the Southern ones. More specifically, Lombardy has high *RCA* in many products in the hightech cluster in the upper-right portion of the graph, e.g., in Machinery and Chemicals. In this cluster, Lombardy shows a medium-low *RCA* in Pharmaceutical and Instruments and therefore we can infer that these are *feasible* sectors for Lombardy, i.e., the region already possess many of the capabilities required to successfully export products in these sectors. A similar line of reasoning can be also applied to the case of Footwear and Vehicles.

By contrast, Sicily is in the opposite situation: its poor diversification does not provide many choices for further steps in the Product Progression Network. For instance, being active in Fish, Sicily could move to the Preparations of meat or fish sector; however, this is a peripheral node with not many further openings. Campania and Veneto are shown to be in an intermediate situation, with moderate to good diversification which opens up various possibilities to enter into different new markets.

Conclusions

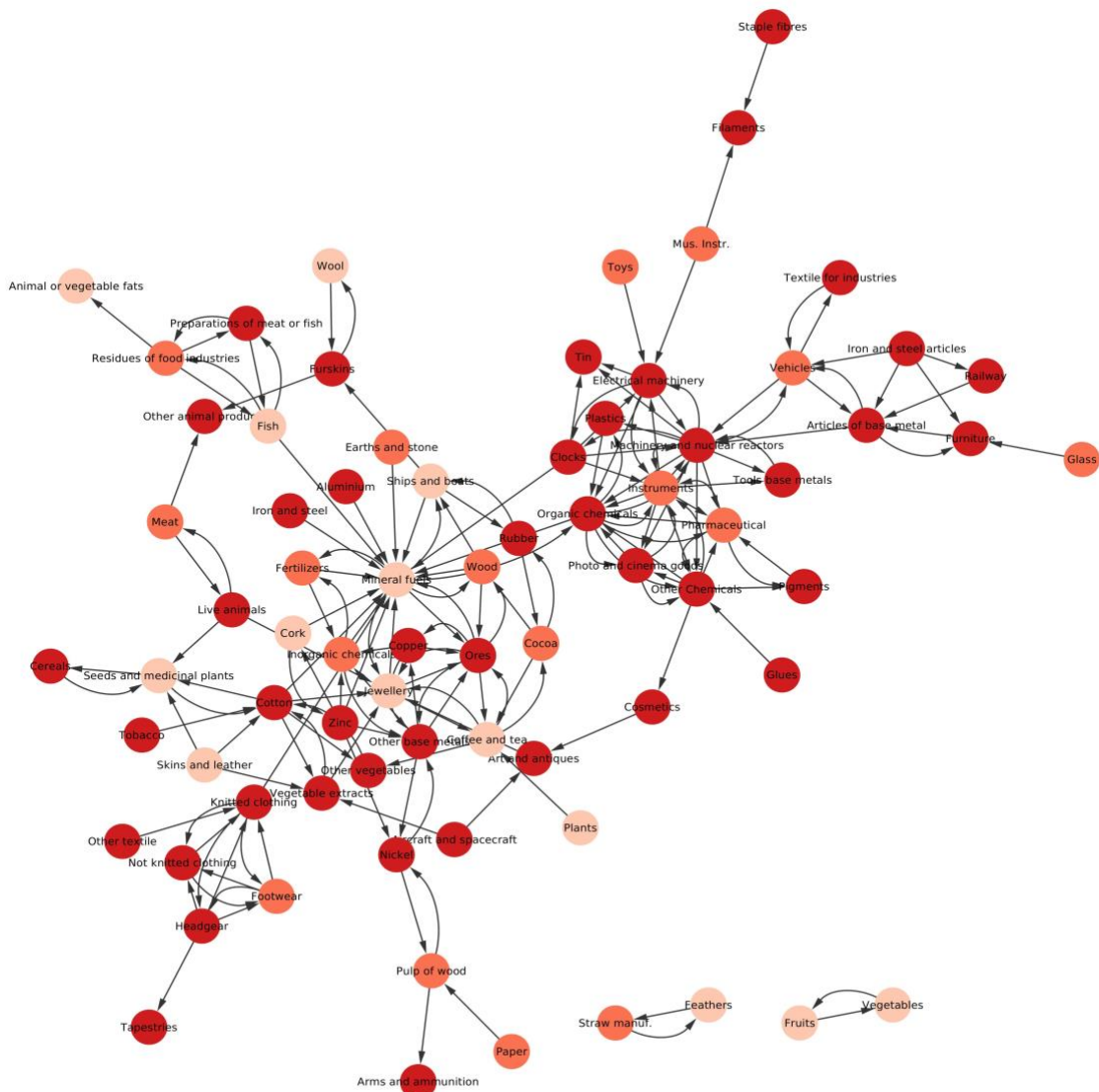
Italy had experienced fundamental regional differences and a North-South divide since well before its unification in the mid-XIX century (Vecchi, 2011; Felice, 2016). These economic disparities have proved stubbornly resistant to policies aimed at reducing the gap between the richer and the less-well-off areas of the country. The Southern regions are still lagging behind the regions in the Centre-North according to most economic and welfare metrics.

The present paper develops a novel analysis of regional inequalities in Italy by applying innovative methods from the recent Economic Fitness and Complexity (EFC) approach. The performance of Italian regions is shown to be positively related to the degree of complexity of their production structure, measured by indicators of both Export Fitness and Technology Fitness. Consistent to the EFC approach, the complexity of the regional economy is seen as reflecting the underlying production and technological capabilities which make it possible for the economy to adapt to new production requirements and to stimulate the introduction of new technologies.

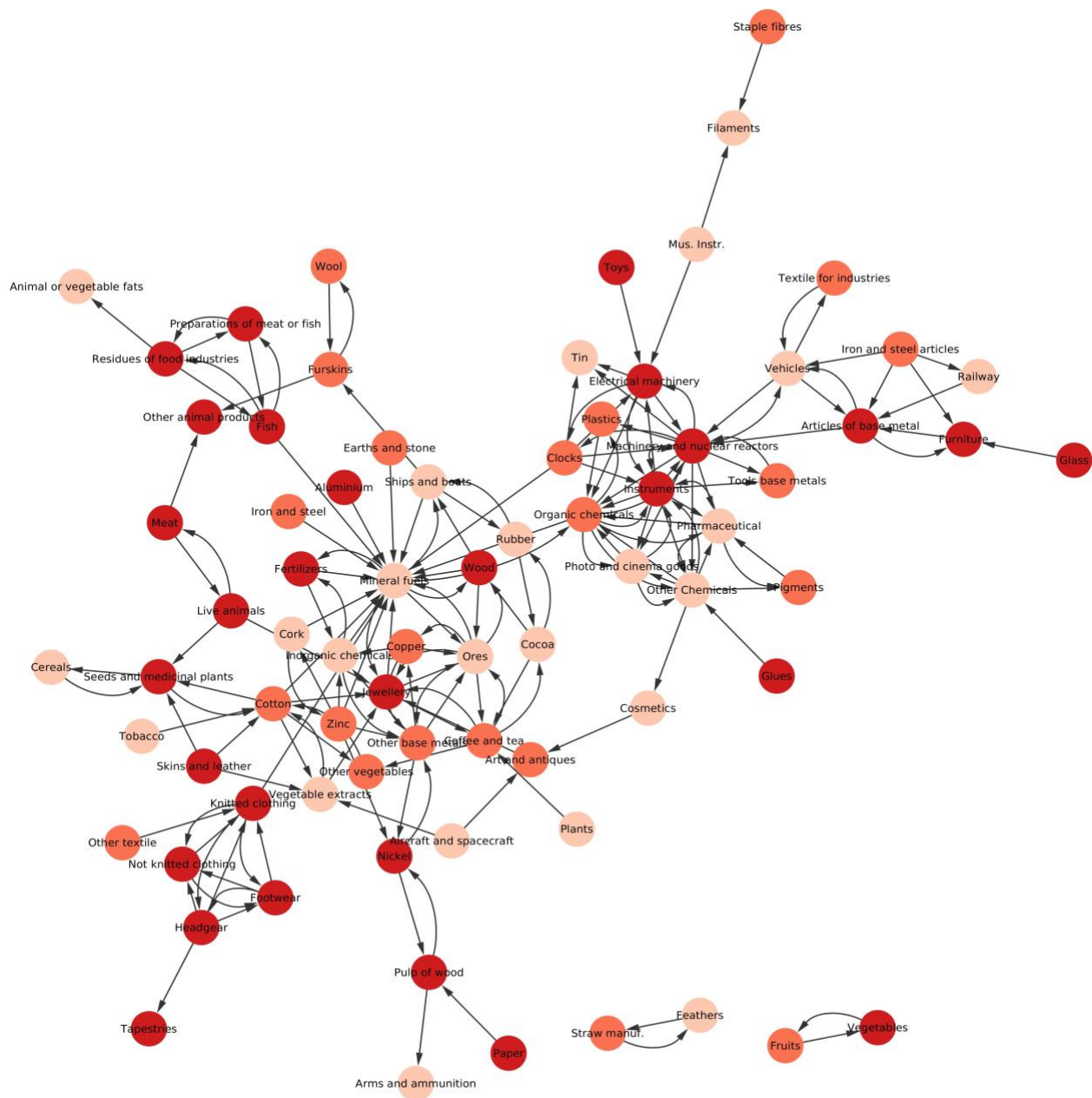
The findings of this paper confirm the North-South divide also in terms of the complexity of their respective production structure, which captures the depth and breadth of the underlying capabilities in the regional economy. Southern regions tend to be characterised by a lower level of complexity than the regions in the Centre-North, according to both the Export Fitness and

the Technology Fitness metrics. The index of Export Fitness is shown to be related to the economic performance of the regional units. The ranking of regions by Export Fitness is however remarkably stable over time, consistent with the stubborn persistence over time of their economic disparities.

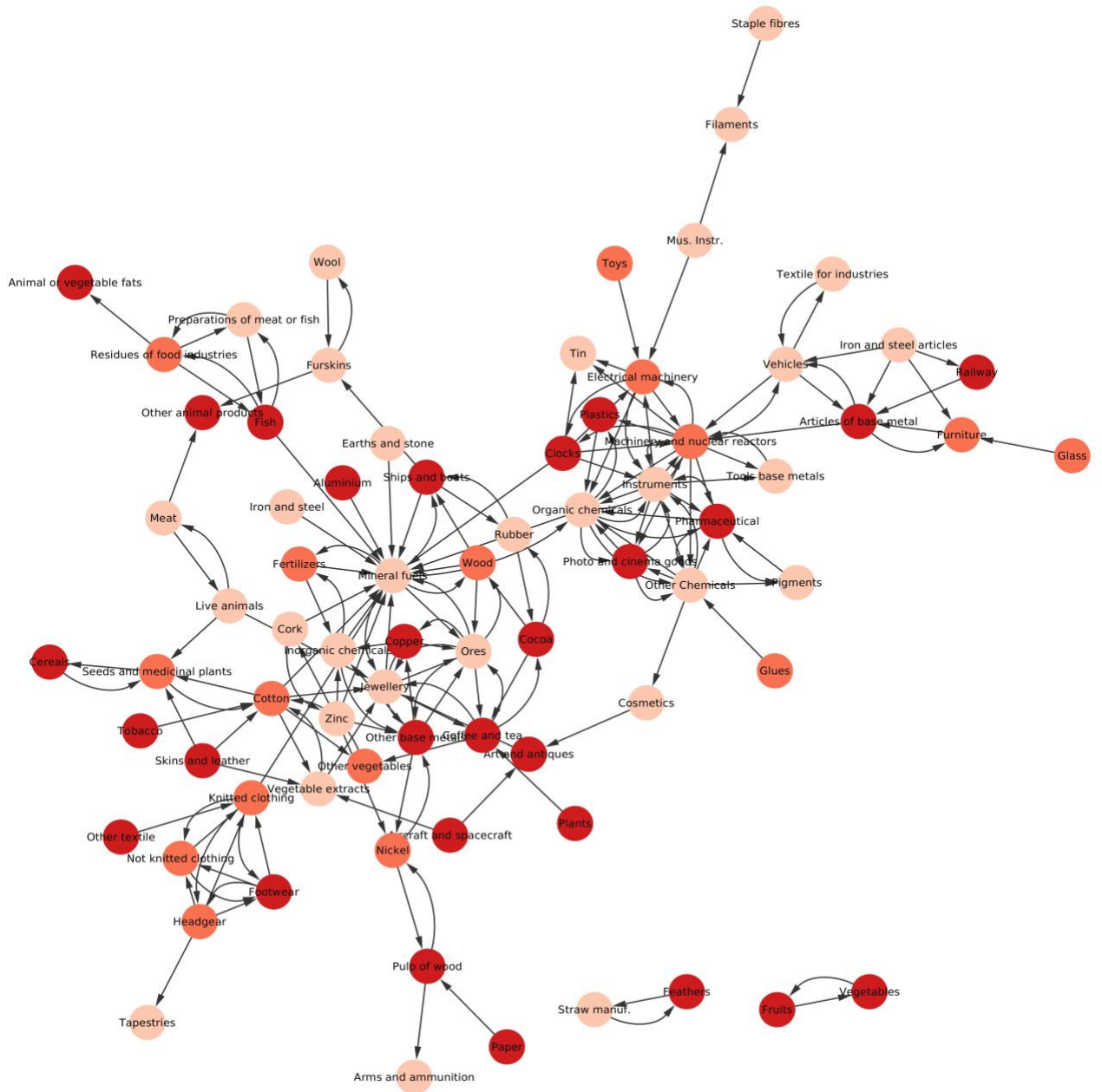
An effective policy seeking to promote convergence of all areas towards the more advanced regions of the country must enhance the set of capabilities in the poorer areas. In the absence of these capabilities, faster growth and convergence of the poorer regions towards the richer ones may never happen. The system-wide approach set out in this paper can identify critical sectors which could play a crucial part in regional economic development, thanks to their rich pattern of connections with other sectors in the regional economy. This analysis can therefore contribute to more effective regional and industrial policies to achieve a balanced economic growth among the different areas of the country.



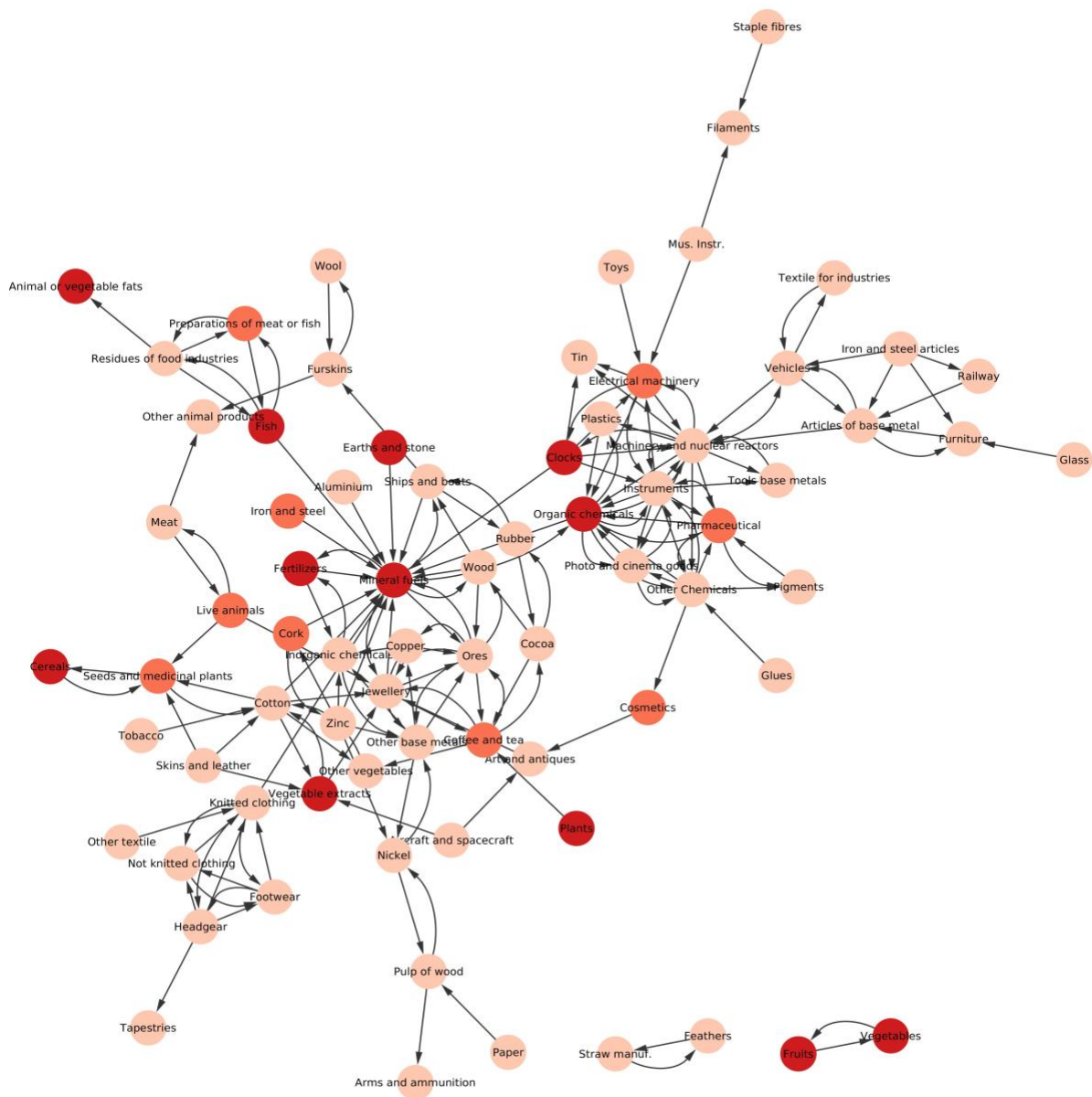
(a)



(b)



(c)



(d)

Figure 8: The Product Progression Network provides a concise representation of the export baskets and the diversification opportunities of the Italian regions. Here, in particular, we show it for Lombardy (a), Veneto (b), Campania (c) and Sicily (d).

References

- Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *The Journal of Economic History* 46(2), 385–406.
- Acemoglu, D. (2009). *Introduction to modern economic growth*. Princeton and Oxford: Princeton University Press.
- Aghion, P. and P. Howitt (1998). Market structure and the growth process. *Review of Economic Dynamics* 1(1), 276–305.
- Albora, G., L. Pietronero, A. Tacchella, & A. Zaccaria (2021). Product Progression: a machine learning approach to forecasting industrial upgrading. arXiv preprint arXiv:2105.15018.
- Arbia, G., M. Battisti, and G. Di Vaio (2010). Institutions and geography: Empirical test of spatial growth models for european regions. *Economic Modelling* 27(1), 12–21.
- Arrow, K. J. (1962). The economic implications of learning by doing. *The review of Economic Studies* 29(3), 155–173.
- Arthur, W. B. (2013). *Complexity Economics*. Oxford: Oxford University Press.
- Arts, S., F. P. Appio, and B. Van Looy (2013). Inventions shaping technological trajectories: do existing patent indicators provide a comprehensive picture? *Scientometrics* 97(2), 397–419.
- Balassa, B. (1965). Trade liberalisation and revealed comparative advantage. *The Manchester School* 33(2), 99–123.
- Balland, P.-A. and D. Rigby (2017). The geography of complex knowledge. *Economic Geography* 93(1), 1–23.
- Ballatore, R. M. and V. Mariani (2019). Human capital differentials across urban and rural areas in Italy. the role of migrations. *Italian Economic Journal* 5(2), 307–324.
- Barro, R. J. and X. Sala-i Martin (2004). *Economic Growth. 2nd Ed.* Cambridge (MA): MIT Press.
- Basile, R., G. Cicerone, and L. Iapadre (2019). Economic complexity and regional labor productivity distribution: evidence from Italy.
- Boltho, A., W. Carlin, and P. Scaramozzino (1997). Will East Germany Become a New Mezzogiorno? *Journal of Comparative Economics* 24(3), 241–264.

- Boltho, A., W. Carlin, and P. Scaramozzino (2018). Why East Germany did not become a new Mezzogiorno. *Journal of Comparative Economics* 46(1), 308–325.
- Boschma, R., A. Minondo, and M. Navarro (2013). The Emergence of New Industries at the Regional Level in Spain: A Proximity Approach Based on Product Relatedness. *Economic Geography* 89(1), 29–51.
- Breschi, S., F. Lissoni, and F. Malerba (2003). Knowledge-relatedness in firm technological diversification. *Research Policy* 32(1), 69–87.
- Bustos, S., C. Gomez, R. Hausmann, and C. A. Hidalgo (2012). The dynamics of nestedness predicts the evolution of industrial ecosystems. *PloS one* 7(11), e49393.
- Caldarelli, G., M. Cristelli, A. Gabrielli, L. Pietronero, A. Scala, and A. Tacchella (2012). A network analysis of countries' export flows: firm grounds for the building blocks of the economy. *PloS one* 7(10), e47278.
- Cicerone, G., P. McCann, and V. A. Venhorst (2020). Promoting regional growth and innovation: relatedness, revealed comparative advantage and the product space. *Journal of Economic Geography* 20(1), 293–316.
- Cimoli, M., G. Dosi, and J. E. Stiglitz (2010). The political economy of capabilities accumulation: The past and future of policies for industrial development. In *Industrial Policy and Development: The Political Economy of Capabilities Accumulation*. Oxford: Oxford University Press.
- Cristelli, M., A. Tacchella, M. Cader, K. Roster, and L. Pietronero (2017). *On the predictability of growth*. Washington, DC: World Bank Group.
- Cristelli, M., A. Tacchella, and L. Pietronero (2015). The heterogeneous dynamics of economic complexity. *PloS one* 10(2), e0117174.
- De Angelis, I., V. Mariani, and R. Torrini (2017). New evidence on interregional mobility of students in tertiary education: the case of Italy. *Politica Economica* 33(1), 73–96.
- de Rassenfosse, G., J. Kozak, and F. Seliger (2019). Geocoding of worldwide patent data. *Scientific Data* 6(1), 1–15.
- Dosi, G., G. Fagiolo, M. Napoletano, and A. Roventini (2013). Income distribution, credit and fiscal policies in an agent-based Keynesian model. *Journal of Economic Dynamics and Control* 37(8), 1598–1625.

- Dosi, G. and R. R. Nelson (1994). An introduction to evolutionary theories in economics. *Journal of Evolutionary Economics* 4(3), 153–172.
- Dosi, G. and R. R. Nelson (2010). *Technical change and industrial dynamics as evolutionary processes*, Volume 1. Amsterdam: North-Holland.
- Dosi, G., K. Pavitt, and L. Soete (1990). *The economics of technical change and international trade*. Pisa: Laboratory of Economics and Management (LEM) Book Series, Sant’Anna School of Advanced Studies.
- European Patenting Office (2020). EPO Data Catalog PATSTAT Global 2020, Spring Edition.
- Fagerberg, J. and M. Srholec (2017). Capabilities, economic development, sustainability. *Cambridge Journal of Economics* 41(3), 905–926.
- Felice, E. (2016). *Perché il Sud è rimasto indietro*. Bologna: Il Mulino.
- Ferrarini, B. and P. Scaramozzino (2016). Production complexity, adaptability and economic growth. *Structural Change and Economic Dynamics* 37, 52–61.
- Griliches, Z. (1998). Patent statistics as economic indicators: a survey. In *R&D and productivity: the econometric evidence*, pp. 287–343. University of Chicago 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. *CID Working Paper* (128).
- 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 of the United States of America* 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.
- Hirschman, A. O. A. O. (1958). *The strategy of economic development*. New Haven: Yale University Press.
- Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: evidence from firms’ patents, profits and market value. *NBER Working Paper Series* (1815).

- Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of Economic Perspectives* 6(2), 117–136.
- Kremer, M. (1993). The O-ring theory of economic development. *The Quarterly Journal of Economics* 108(3), 551–575.
- Lall, S. (1992). Technological capabilities and industrialization. *World Development* 20(2), 165–186.
- Lanjouw, J. O. and A. Mody (1996). Innovation and the international diffusion of environmentally responsive technology. *Research Policy* 25(4), 549–571.
- Mariani, M. S., Z.-M. Ren, J. Bascompte, and C. J. Tessone (2019). Nestedness in complex networks: observation, emergence, and implications. *Physics Reports* 813, 1–90.
- Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications* 9(1), 141–142.
- Napolitano, L., E. Evangelou, E. Pugliese, P. Zeppini, and G. Room (2018). Technology networks: the autocatalytic origins of innovation. *Royal Society Open Science* 5(6), 172445.
- Napolitano, L., A. Sbardella, D. Consoli, N. Barbieri, and F. Perruchas (2019). Green innovation and income inequality: A complex system analysis. *SWPS-SPRU Working Paper Series*.
- Nelson, R. R. and S. G. Winter (1982). *An evolutionary theory of economic change*. Cambridge (MA): Harvard University Press.
- Nesta, L. and P. P. Saviotti (2005). Coherence of the knowledge base and the firm’s innovative performance: evidence from the us pharmaceutical industry. *The Journal of Industrial Economics* 53(1), 123–142.
- Operti, F. G., E. Pugliese, J. S. Andrade Jr, L. Pietronero, and A. Gabrielli (2018). Dynamics in the fitness-income plane: Brazilian states vs world countries. *PloS one* 13(6), e0197616.
- Penrose, E. (1959). *The Theory of the Growth of the Firm*. Oxford: Oxford University Press.
- Pietronero, L., M. Cristelli, A. Gabrielli, D. Mazzilli, E. Pugliese, A. Tacchella, and A. Zaccaria (2017). Economic complexity: “buttarla in caciarà” vs a constructive approach. *arXiv preprint arXiv:1709.05272*.
- Pugliese, E., G. L. Chiarotti, A. Zaccaria, and L. Pietronero (2017). Complex economies have a lateral escape from the poverty trap. *PloS one* 12(1), e0168540.

- Pugliese, E., G. Cimini, A. Patelli, A. Zaccaria, L. Pietronero, and A. Gabrielli (2019). Unfolding the innovation system for the development of countries: coevolution of science, technology and production. *Scientific Reports* 9(1), 1–12.
- Pugliese, E., L. Napolitano, A. Zaccaria, and L. Pietronero (2017). Coherent diversification in corporate technological portfolios. *arXiv preprint arXiv:1707.02188*.
- Pugliese, E., L. Napolitano, A. Zaccaria, and L. Pietronero (2019). Coherent diversification in corporate technological portfolios. *PloS one* 14(10).
- Pugliese, E., A. Zaccaria, and L. Pietronero (2016). On the convergence of the fitness-complexity algorithm. *The European Physical Journal Special Topics* 225(10), 1893–1911.
- Rodano, G. (2018). *Elementi di una teoria per la storia economica. Una rilettura dell'Italia dal 1950 a oggi*. Bologna: Il Mulino.
- Saracco, F., R. Di Clemente, A. Gabrielli, and T. Squartini (2015). Randomizing bipartite networks: the case of the World Trade Web. *Scientific Reports* 5, 10595.
- Saracco, F., M. J. Straka, R. Di Clemente, A. Gabrielli, G. Caldarelli, and T. Squartini (2017). Inferring monopartite projections of bipartite networks: an entropy-based approach. *New Journal of Physics* 19(5), 053022.
- Sbardella, A., F. Perruchas, L. Napolitano, N. Barbieri, and D. Consoli (2018). Green technology fitness. *Entropy* 20(10), 776.
- Sbardella, A., E. Pugliese, and L. Pietronero (2017). Economic development and wage inequality: A complex system analysis. *PloS one* 12(9), e0182774.
- Sbardella, A., E. Pugliese, A. Zaccaria, and P. Scaramozzino (2018). The role of complex analysis in modelling economic growth. *Entropy* 20(11), 883.
- Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica* 46(5), 1013–1024.
- Simon, H. A. (1991). The architecture of complexity. In *Facets of systems science*, pp. 457–476. Springer.
- Sutton, J. (2012). *Competing in capabilities: the globalization process*. Oxford: Oxford University Press.
- Sutton, J. and D. Trefler (2016). Capabilities, wealth, and trade. *Journal of Political Economy* 124(3), 826–878.

- Tacchella, A., M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero (2012). A new metrics for countries' fitness and products' complexity. *Scientific Reports* 2(723).
- Tacchella, A., M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero (2013). Economic complexity: conceptual grounding of a new metrics for global competitiveness. *Journal of Economic Dynamics and Control* 37(8), 1683–1691.
- Tacchella, A., D. Mazzilli, and L. Pietronero (2018). A dynamical systems approach to gross domestic product forecasting. *Nature Physics* 14(8), 861.
- Tacchella, A., A. Zaccaria, M. Micheli, and L. Pietronero (2021). Relatedness in the era of machine learning. *arXiv preprint arXiv:2103.06017*.
- Teece, D. J., G. Dosi, and S. G. Winter (1994). Understanding corporate coherence. Theory and evidence. *Journal of Economic Behavior and Organization* 23(1), 1–30.
- Teece, D. J., G. Pisano, and A. Shuen (1997). Dynamic capabilities and strategic management. *Strategic Management Journal* 18(7), 509–533.
- Vecchi, G. (2011). *In ricchezza e in povertà: il benessere degli italiani dall'Unità a oggi*. Bologna: Il Mulino.
- Viesti, G. (2021). *Centri e periferie: Europa, Italia, Mezzogiorno dal XX al XXI secolo*. Bari: Laterza.
- Winter, S. G. (1997). Knowledge and competence as strategic assets. In A. Klein, *The strategic management of intellectual capital*, Volume 187, pp. 165–187. New York, Elsevier Ed.
- Zaccaria, A., M. Cristelli, A. Tacchella, and L. Pietronero (2014). How the taxonomy of products drives the economic development of countries. *PloS one* 9(12), e113770.
- Zaccaria, A., S. Mishra, M. Z. Cader, and L. Pietronero (2018). Integrating services in the Economic Fitness approach. Technical Report WPS8485, The World Bank.