Economic Complexity

Luciano Pietronero, Sapienza, General Coordinator – **Masud Cader**, IFC-WBG, Country Analytics
Andrea Gabrielli, CNR-ISC, Science & Technology – **Matthieu Cristelli**, IFC-WBG, Algorithms & Forecasting
Roland Kupers, NEM, Managing Director – **Emanuele Pugliese**, CNR-ISC, Technology & Development
Andrea Tacchella, CNR-ISC, Algorithms & Machine Learning – **Andrea Zaccaria**, CNR-ISC, Networks & Companies

What is Economic Complexity?

Economic Complexity (EC) is a new field of research that consists of a radically new methodology. It describes economics as an evolutionary process of **ecosystems** made of industrial and financial technologies as well as infrastructures that are all **globally interconnected**. The approach is multidisciplinary addressing emerging phenomena in economics from different points of view: analysis of complex systems, scientific methods for systems and the recent developments in **big data** (in the spirit of Google Page Rank, deep learning and beyond). This approach offers new opportunities to constructively describe technological ecosystems, analyse their structures, understand their internal dynamics, as well as to introduce **new metrics**. It provides a new paradigm for a **fundamental economic science based on data** without any role of ideologies or interpretations, which is becoming a necessary choice in a highly interconnected and globalized world, especially after the great financial and economic crisis of recent years.

Big Data and Economic Complexity

A crucial element of our methodology is a radically new approach to the problem of **big data**. Big data is often associated with **"big noise"** as well as a subjective ambiguity related to how to structure the data and how to assign them a value that should reflect many arbitrary parameters. In the case of the evaluation of the industrial competitiveness of a country, the required parameters for such an analysis could be more than one hundred. A key feature of EC is to go **from 100 parameters to zero parameters** and obtain results which can be tested scientifically. This is done by focusing on the data in which the **signal to noise** ratio is optimal and developing iterative algorithms in the spirit, but other than Google, and optimized to the economic problem in question. In particular the study of a country or a company is not done at the individual level but through the global network to which it belongs. In this way you get the **Fitness** of the countries and the **Complexity** of the products, together with many other novel information and strategic insight.

Dynamics and Forecasting

The dynamics in the new GDP-Fitness space opens up to a completely new way for monitoring and **forecasting**. The trajectories of countries in this new space show regions of laminar and turbulent behavior which lead to a heterogeneous, non linear approach to forecasting. This is similar to physical dynamical systems and modern weather forecasting. Then, the **network of products** and their evolutionary dynamics is built using machine learning methods. Finally, the same approach is applied to **scientific production, patents and technologies**, so to open up the possibility of analyzing the core elements of the **innovation process**. All this provides a disciplined set of indicators that can be used for the industrial planning of countries and regions. The application range of our general method is not limited to the field of economics, since it can be fruitfully adapted to other problems connected with big data too, for instance medical diagnostics and computational biology. Various exploratory studies are currently underway in these fields.

World Bank Group and Policy Makers

Economic Complexity – in addition to a new vision for a data-based scientific approach for fundamental economics – offers a new set of metrics able to quantify the competitiveness of countries and of technological sectors, measuring future development prospects for nations as well as for large companies. Those metrics have already shown to have a major impact for **policy makers** and for **industry**. The **World Bank (WB)** has shown a great interest in these new methods. Hence, the World Bank developed a close collaboration with the research group led by Pietronero (WBs Senior Advisor) and the WB has recently adopted this new methodology in its strategic analysis. In **Beijing** (China), the WB together with Pietronero's group presented the Economic Complexity methodology to the DRC (Development Research Center of the State Council), a Chinese governmental think-tank that will adopt these methodologies **for planning China's further industrial development**.

The Data

The selection of data is focused on those variables leading to the best signal to noise ratio and which also permit global algorithmic and scientifically testable analysis. We perform our analysis by looking at the activities of countries and firms on three layers: **industrial production and export**, **technological change**, **scientific activity**. A crucial aspect of our research is not only to have the most updated data, but also to clean the raw data to extract and organize information in a way that allows us merge multiple sources of data in one analysis. In the following, our main sources of raw data are listed.

Industrial Competitiveness

UN COMTRADE A database hosted by the **United Nation** with detailed data on countries and products export at different level of aggregation: 2 digits products (total 96 products), 4 digits (∼1100 products), 6 digits (∼6000 products). The group has treated the data with original methodologies so to check for their consistency while also estimating the intrinsic noise. The complete data is released to the public once per year, with one year delay. However we have developed a network reconstruction technique to have the full database in real time on a monthly basis using partial information which allows has a faster release.

Services, Finance, Input-Output and Labor data While UN COMTRADE deals only with physical exports, we recently added data on services to our analysis, from an experimental internal source produced by the **International Monetary Fund** and provided to us by the **World Bank Group**. For selected countries, mostly in the extended OECD sample, we can complement our analysis using data about both internal and foreign industrial production in different sectors, as well as their interaction, by looking at Input-Output data and Labor participation data from the **WIOD** and other sources. This has the advantage that all activities are covered and that we are able to look at the relationship with Value Added and Global Value Chains.

Bureau van Dijk, ORBIS This global database on firms allows us to move our analysis to microdata, from countries to regions and **individual companies**. It also allows us to connect production data to technology data.

Technologies

EPO PATSTAT A database organized by the **European Patent Office**, aggregating data from all the patent offices in the World (over 100 different offices, among witch the US patent office and the European Patent Office). Patents are classified with respect to their technological field at various aggregation, from sections (8 codes) to classes (∼130 codes) to groups (∼7000 codes) and sub-groups (∼70000 codes), with additional intermediate steps. Each patent can be assigned to countries accordingly to the nationality of the applicants or the inventors, and we further organized the data so that it can be pinpointed in space up to the **smallest geographical aggregations**.

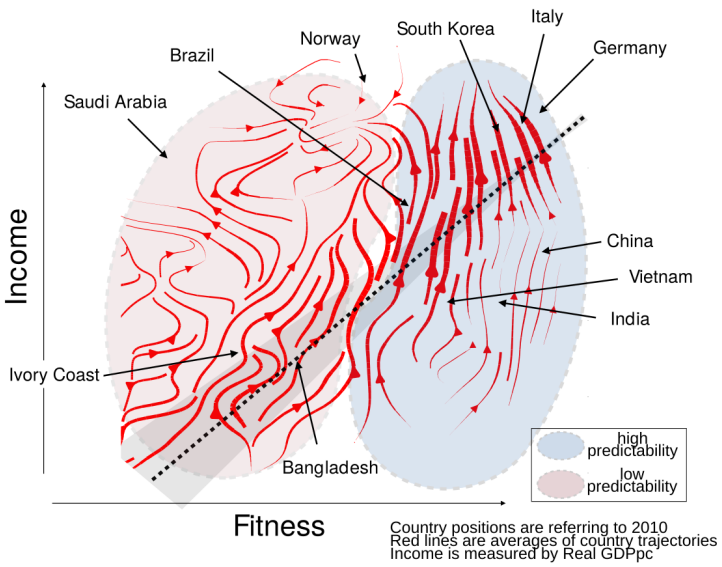
Science

SciVal The SciVal platform aggregates data from **Elsevier-Scopus**, which covers journals, trade publications, book series, conference proceedings, and books. Data cover years from 1996 to 2015, and each scientific publication is assigned to a category at two levels of aggregations: scientific sectors (28 codes) and subsectors (∼300). These data also provide a direct linkage between **scientific publications** and **patents** according to direct citations.

Innovative Methodologies

We developed several **novel algorithms to extract key information** from our data. These algorithms work by looking at each layer of the Complex System individually, i.e. the countries capabilities that can be inferred by their position in the trade network. Other algorithms look at the information relating to multiple aspects, i.e. how the technologies produced by a firm affects its export competitiveness. The following is a list of some quantitative research questions we addressed in the past. **The validation of our answers have been done internally and independently by the World Bank Group**, that adopted our methodology. Indeed, five members of our team are now consultants for the World Bank Group and our methods are used extensively by policy makers all around the World. We are now considering also the possible use of these results for the **financial industry** for long term investments.

Fitness and Complexity Algorithm This was the first problem our group considered [1]. By exploiting the matrix connecting **the countries and the products** that a country exports, we compute at the same time the **Fitness** of the country, measured as the weighted sum of the **Complexity** of the exported products, and the Complexity of the products, by looking at the Fitness of the countries exporting it with a suitable non linear relation. We solve the apparently circular reasoning by looking at fixed point of the coupled equations of the non-linear algorithm. This example shows that Google Page Rank is not the only algorithm to introduce a ranking in a complex network and, for each big data problem, a suitable new algorithm should be introduced.



The country Fitness is a measure of the country intangible capabilities and it has been proved to be a crucial dimension to forecast growth potential. With non-linear techniques of heterogeneous dynamics,

similar to **weather forecasting**, we assess the expectation for the future growth of countries [2]. Advancements based on similar techniques have been also tested by the World Bank Group and they are now used extensively by our group to make forecasting of growth in the medium-long period, 5 to 10 years. **Forecasting favorable compares with the standard IMF forecasting**. The concept of Fitness-GDP dynamics has already proven very useful in the analysis of **BRIC countries**. By looking at their Fitness, one could have observed already ten years ago that these four countries are very different for their position and even more for for their trends. These concepts are now acknowledged by the standard analysis, but with a delay of ten years. Using the same technology, we argued that **China** has not yet exploited its full growth industrial potential. These concepts provide also a firm ground for the interpretation of the **secular stagnation** of advanced countries.

The same technique to compute Fitness and Complexity of countries and products have been used for science and technology data, to estimate **scientific and technological capabilities** of countries. The analysis of the multilayer space led to forecast the **long run growth** of countries, up to 20-30 years and includes the forecasting of radical innovation. These long times are intrinsic of the process: **Science-Technology-Product**.

Taxonomy of Products By looking at product co-occurrence in the same countries, we obtained a measure of **relatedness between different products** [3]. It was then possible to build a hierarchically **directed** network, in which the taxonomy of products emerges in a natural way. The structure of the network influences the path of countries' development. The same technique has been used on science and technology data, to estimate the relatedness between scientific and technological fields.

Sector-level Forecasting Considering the subset of products in an industrial sector, we estimated the country **capabilities in a specific sector**. By doing so we can focus our analysis at a greater detail, to identify which sectors have more growth potential.

Country Complexity Spectroscopy At an even greater detail, it is possible to look at the individual products in a country and their complexity, to produce a detailed **spectroscopy** of the complexity pattern of the country.

Value Added and Global Value Chain We looked at how the position of a country in the Global Value Chain, the balance between the **import** and the **export** of complex products, can help us identify **long term growth patterns** and which developing countries will **avoid the middle income trap**.

Country Development and Societal Change We observed how **the development transforms a country**, increases or decreases **inequality** and potentially destabilizes internal equilibria, changes its **demographic structure** and affect its **governance**, in different development paths.

Country Development, Resilience and Sustainability Related to the previous point, we looked at how the kind of development path affects the **resilience of a country** to shocks and the long-term sustainability of growth, both **ecological sustainability and social sustainability**. This approach can be extended to develop specific sustainability metrics and forecasts using water, energy and material intensity data for each product.

Coherence of the firm knowledge base By looking at the patent portfolio of different firms, we built a **measure of firm coherent diversification** in its patent portfolio that considers not only the number of different patent fields in which the firm is active, but also their **coherence**. This measure is shown to be a very good predictor for the **efficiency of a firm**, measuring the degree of synergies between firms different activities.

Radical Innovations In a patent, multiple technology fields are used at the same time. By using **machine learning techniques** inspired to neural networks, we developed a metric in the technology space to see which technology fields, even if they have been never used together, will be likely used together in the next future. By doing so, we are able to forecast **which technological fields are ripe for new disruptive innovations**.

From innovation to production We looked at the network connecting technological field to production, and successful competition of a specific product in the international markets. We did so by computing the probability that a **patent innovation affects a specific product**. We are then able to estimate the effect that innovation capabilities in specific technologies have on the future potential in a given production line, which represents the **innovation fingerprints** of future success.

Size effects While our measures are in principle **intensive measures**, in practice social systems are not scale invariant and a country is not simply an aggregation of regions: several effects have a **specific scale dimensions**. By looking at the dynamics of technological capabilities at different aggregation scales, we studied how to tackle such effects so to smoothly move from countries to regions and from industrial clusters to firms.

What can Economic Complexity do for you?

The EC methodology has been used successfully and extensively both for academic purposes and for policy and strategic management by policy makers and firms. We are in the position to provide **invaluable quantitative and novel information previously assumed necessarily qualitative and subjective in nature**. While some of the following tasks might seem qualitative in nature, our answer to the following questions is quantitative: our algorithms offer **precise answers to precise questions**, with a probability distributions of scenarios. We report here a list of the main concrete results that can be obtained with the EC approach. We split the list in **macro** and **micro** economy, as the analysis mostly relates to countries or regions and firms respectively. Obviously many macro-economy variables have an important impact on firms and many studies of firm competitiveness could be relevant for a policy maker.

Macro-economy

Capability assessment Study of the **Fitness** of a country or a region in time, in terms of its capabilities to export complex products or innovate in complex fields. Comparison with the main competitors, cross-sectional and in time.

Co-evolution Dynamics of Fitness and Macro-economic Variables

Through the assessment of the **Fitness** of a country or a region, forecasting of **future macro-economic variables like GDP**, and comparison with other common statistics and scenario analysis.

Product Progression Network

We identified which products are the best **diversification targets** for a country, i.e. products that are feasible given the country capabilities while giving the highest payoff.

Sector Competitiveness

Study of the **Fitness** of a country **in one or more specific industrial sectors**, in terms of the idiosyncratic capabilities specific to export with success in that specific sector.

Innovation Competitiveness in one Market

Study of the **Fitness** of a country or a region **in a specific production line**, in terms of the capabilities to make innovations that will be relevant to **future** export competitiveness of a present production line.

Forecasting of Import/Export Markets

Forecast of the future trade network: which countries will represent the **new rising markets** for a specific product? Which countries will enter in a market as a competitor?

Micro-economy

Diversification analysis for Firms

Study of a firm potential, in terms of knowledge base and production capabilities, to decide **which new products** could be a viable addition to a firm production lines.

Vertical and horizontal integration

Study of industrial synergies: which firms could be a **viable partner** for a partnership, a network, or for a direct **merging or acquisition**, with the intent to acquire capabilities to penetrate a new production line, or to reinforce the innovation capabilities in present production lines.

Optimal location

Study of the optimal **location of a new business unit** by looking at the industrial knowledge of the sector and the possibility of technological spill-overs, eventually given various industrial and administrative constraints required by our partner.

Assessment of competitiveness in present Markets

Study of the innovative capabilities of a firm in high technology sector, to **forecast future competitiveness** in its present production lines.

Assessment of competitiveness in *FUTURE* Markets

Study of the potential of a firm, in terms of the capabilities to make innovations that will be relevant to **future competitiveness in future markets**, innovations that are in the **adjacent possible**, likely to happen but still not present, that are possible to penetrate by the firm.

CNR-Institute for Complex Systems

The **Institute for Complex Systems** (ISC) is an institute of the Italian **National Research Council** (CNR). It has been founded by **Luciano Pietronero** and it is currently directed by **Claudio Conti**. The institute is home of most of our research, and it has managed the two main projects on Economic Complexity by this group, **CrisisLab** and **GrowthCom**.

New Economic Metrics

New Economic Metrics (NEM) is an incorporated company that provides customized analysis, case studies and quantitative forecasts to private stakeholders. NEM methodologies complement the Economic Complexity framework with proprietary tools and datasets, that are thought and implemented for providing quantitative answers for specific **business scenarios**.

Partners

Scientific and Business advisors

Many people helped and advised us, to mention a few of our partners: **Roberto Benzi**, Scientist, Tor Vergata – **Guido Chiarotti** Scientist, Entrepreneur, Manager, MaX - European Centre of excellence – **Giulio Cimini**, Scientist, IMT Lucca – **Marco Corradino**, founder and CEO of Lastminute.com Group – **Laura Deitinger** and **Riccardo Masia**, president and vicepresident of Assoknowledge (Confindustria SIT) – **Adriano De Maio**, Scientist, Lombardy Region – **Giovanni Dosi**, Economist, Scuola Superiore Sant'Anna – **Cristiano Esclapon**, Banker and Founder and President of Club Italia Investimenti 2 – **Rodolfo Guzzi**, Scientific board of Space Italy – **William Jeneway**, Venture capitalist and Economist – **Anatole Kaletsky**, Economic Journalist – **Franco Miglietta**, Scientist, IBIMET – **Martin Reeves**, Strategic Institute, BCG – **Francesco Sylos Labini** Scientist, Istituto Fermi.

Policy Makers and Stakeholders

Some of the other institutions that have been applying our methods are: the **Italian Ministry of Foreign Affairs** (i.e. studies of business opportunities between Italy, China and ASEAN countries); the **Ministry of Economics of Singapore**; **OECD** (Paris); **Boston Consulting Group** (New York); the **Institute for New Economic Thinking** (New York); **Royal Dutch Shell** (Den Haag, NL); **Azimut Financial Group** (Milan); **Alibaba Business School** (Hangzhou, China); **Bravofly-Rumbo Group** (Chiasso, CH); **Assoknowledge (Confindustria SIT)**. Recently the **Italian Space Agency (ASI)** has commissioned to our group a study of Space Economy with the EC methodology.

For more information

- Web-site `http://www.lucianopietronero.it/`
- Nature editorial on EC: `http://www.nature.com/news/physicists-make-weather-forecasts-for-economies-1.16963`
- Economic Complexity was presented in a conference dedicated to OECD in Paris. Pietronero and Masud Cader (World Bank) presentations can be downloaded here: `http://www.lucianopietronero.it/presentations/`

References

[1] Tacchella A et al. (2012) Scientific Reports 2, Article number: 723 doi:10.1038/srep00723

[2] Cristelli M et al. (2015) PLoS ONE 10(2): e0117174. doi:10.1371/journal.pone.0117174

[3] Zaccaria A et al. (2014) PLoS ONE 9(12): e113770. doi:10.1371/journal.pone.0113770