Growth scenarios for sub-Saharan countries in the framework of economic complexity

Cristelli, Matthieu and Tacchella, Andrea and Zaccaria, Andrea and Pietronero, Luciano

ISC-CNR, V. dei Taurini 19, 00185, Roma, Italy, Dip. di Fisica, Università "La Sapienza", P.le A. Moro 2, 00185, Roma, Italy, LIMS, 35a South Street, W1K 2XF London, United Kingdom.

November 2014

Growth scenarios for sub-Saharan countries in the framework of economic complexity

Matthieu Cristelli[1], Andrea Tacchella[1,2], Andrea Zaccaria[1], Luciano Pietronero[2,1,3]

[1] ISC-CNR, V. dei Taurini 19, 00185, Roma, Italy

[2] Dip. di Fisica, Università "La Sapienza", P.le A. Moro 2, 00185, Roma, Italy

[3] LIMS, 35a South Street, W1K 2XF London, United Kingdom.

17/11/14

Abstract

We present a comparative analysis of the medium-long term perspectives of development for sub-Saharan countries in the framework of *economic complexity*. This analysis is made in comparison with the development of Asian tigers. Economic complexity is a data-driven framework which aims at providing a more scientific basis for the economic theory and it has a specific focus on understanding the determinants of growth by means of two new economic dimensions: the country fitness and the product complexity. We argue that the fitness of countries is a quantitative assessment of those intangible assets, which drive the growth. The comparison of this measure for intangibles with monetary figures provides effective insights on the growth potential of countries and defines the *fitness-income* plane.

The analysis of the dynamics in this plane reveals that most sub-Saharans get stuck in a pre-industrial regime which can be thought as a generalized *poverty trap* where both income and fitness dimensions are considered. Only Senegal, Kenya, Tanzania, Madagascar and Uganda show a behavior compatible with the early steps of a long term stable and sustained growth, which resembles the one of Vietnam and Malaysia at the beginning of the nineties. As expected, South Africa is the most mature economy of the southern part of Africa. However, its trajectory highlights the concrete risk of an incomplete development of its productive system in terms of diversification, which might concretely jeopardize South Africa's chance to reach the level of wealth of fully developed countries and put the country at risk of getting stuck in the so-called *middle-income trap*.

1 Introduction

Understanding the development and growth paths of countries represents one of the key challenges of economics in view of its huge societal and political impact. Although economics has proposed a number of theories for the determinants of the growth in the last two centuries, still nowadays this question appears highly debated and controversial. Even though a systematic review of the literature is beyond the scope of this work, the quest to explain the process generating stable long term growth and industrialization can be traced back to the very founders of the economic theory [Smith (1776); Ricardo (1891)]. In the twentieth century several drivers

for the growth have been invoked. On one hand, inspired by Schumpeterian arguments on *creative destruction* [Schumpeter (1934, 1939)], endogenous determinants such as technology have been proposed in [Aghion & Howitt (1992); Romer (1990)]. On the other hand, using the Ricardian concept of *comparative advantage* in [Heckscher & Ohlin (1991)] country endowments, such as capital and labor are proposed as direct drivers of growth and development. However, in the fifties, the validity of this approach has been strongly challenged by the attempt to empirically verify the Heckscher-Ohlin theory by Leontief [Leontief (1953, 1956)].

The lack of consensus or conclusive answers also characterizes the mechanisms of the very first steps of development of an economic system in which the country exits from a subsistence level (also known as *poverty trap* or malthusian regime [Beinhocker (2006)]) and moves towards industrialization. Standard economic theories invoke a multiple-equilibria process characterized by a step-like development path due to the breakage of successive barriers whereof the nature has been proposed in terms of capital [Solow (1956)], demand [Murphy et al. (1988)] or more comprehensive factors [Galor & Weil (2000)]. However, among many alternative factors and mechanisms thought to be relevant in these very first steps towards industrialization, we can cite culture-based arguments [Weber (2002)], geographic [Diamond & Ordunio (2005)], political [Acemoglu et al. (2005)] and biological ones [Ashraf & Galor (2013)]. Literature is also very debated on the possible outcomes of this exit such as incomplete or partial development: the so-called middle-income trap [Eichengreen et al. (2013); Kharas & Kohli (2011); Ohno (2009)].

The variety of determinants affecting growth is even larger in the empirical literature where hundreds of drivers can be listed by analyzing review literature [Levine & Renelt (1992); Sachs et al. (1995); Bleaney & Nishiyama (2002)] but no real consensus emerges. The factors which have been tested in those works range from capital to other endogenous factors [Auerbach et al. (1993); Kormendi & Meguire (1985)] through the fiscal burden, the level of GDP *per capita*, trade openness, human capital, political and institutional system, foreign capital, inflation, demographic aspect, natural resources, etc.

In this paper, we want to address the issue of economic growth for sub-Saharan countries in a completely different fashion with respect to previous attempts by means of the framework of a novel and strongly data-driven approach, called *economic complexity*. In analogy with finance where price *should* discount all the available information [Fama (1970)], we argue that the productive capacity of countries – as measured by the export basket – is able to, roughly speaking, *discount* the network of capabilities [Lall (2000); Hidalgo & Hausmann (2009); Tacchella et al. (2012)] in which the intangible information of the potential of growth and competitiveness are encoded. Using a recently proposed measure for this intangible assets [Tacchella et al. (2012); Cristelli et al. (2013); Tacchella et al. (2013); Caldarelli et al. (2012)] – the fitness of country – and a forecast scheme introduced in [Cristelli et al. (2014)] – the *selective predictability scheme* – we discuss a comparative analysis of the long term scenario of growth for sub-Saharan countries.

The economic complexity approach permits to cast into a scientific framework key questions as whether and which sub-Saharan countries can achieve a sustained and stable development in the long term [Ansu (2014)], which countries have started the industrialization process, which countries are exiting the poverty trap, and citing [Anyanwu (2014)], whether there are any lesson from China (and from Asian tigers), thought to be prototype of emerging and emerged countries.

The overall picture from the evolution of countries in the plane fitness-income reveals that most of sub-Saharan countries are stuck in a pre-industrial state and the majority is still far from the region where the early steps of s sustained growth are observed. This pre-industrial regime can be thought as a generalized *poverty trap* which is defined both by the income level and the technological level of the country as measured by the fitness [Zaccaria et al. (2014); Cristelli et al. (2014); Tacchella et al. (2014)]. Therefore countries which are outside the *income* poverty trap

might be inside the regime we hereafter simply call *poverty trap* since this regime also depends on the fitness dimension. We are arguing the income dimension gives a partial insight on the exit process from the *poverty trap* and we believe a more complete comprenhension can be achieved by also considering the fitness.

Only five countries, namely Kenya, Madagascar, Senegal, Uganda and Tanzania, are exiting from the poverty trap. The very early steps of their industrialization process are similar to the path followed by most Asian tigers in the past two-three decades. In particular the level of sophistication (as measured by the fitness of country) of these five countries resembles the one of Vietnam exiting the poverty trap at the beginning of the nineties.

As a further consideration, it is well-known that Africa's soil is rich in a number of raw materials ranging from oil to minerals. However, differently from what one would expect, we only spot two countries characterized by a scenario typical of a raw-material exporter like Gulf countries – high GDP *per capita* but extremely low fitness ($\log(F)<-2$). We can draw two conclusions based on this anomaly: on one hand, extensive exploitation of these natural resources is still very marginal in many sub-Saharan countries. On the other hand, differently from Gulf countries, Russia and other typical raw material-based economies, this might shed light on a general inability to hold the capitals which derive from the raw materials. In this second scenario, the chance to use these capitals to ignite a diversification and industrialization process are undermined.

Without surprise, South Africa is the most mature economic system of Africa showing an intermediate level of industrial development. Despite a well-established industrialization process, South Africa's long term scenario suggests the risk of an economy unable to fully develop and of getting stuck in the middle-income trap due to a poor diversification process of its productive system.

We argue that, within our framework, we can provide effective hints to optimize industrialization process, which may accelerate the path of exit from the poverty trap and consolidate the early stages of the industrialization process [Zaccaria et al. (2014)]. As an example, these policy advises can effectively reduce the risk for South Africa to get stuck in the middle-income trap through a smart and efficient planning of the next steps of its industrialization process favoring those sectors, which will provide the highest strategic advantage to South Africa.

The paper is organized as follows: in section 2 we briefly depict the framework of economic complexity, in section 3 we give the mathematical details of the methodology defining the metrics for the fitness of countries and the complexity of products, in section 4 we present the growth scenarios of sub-Saharan countries, in section 5 we discuss the spectroscopy of the emerging sub-Saharan countries and in section 6 we draw a number of conclusions and discuss the policy implications of our analysis. In the appendix A we briefly discuss the datasets and the auxiliary methodologies used.

2 Economic Complexity: an overview

As mentioned, similarly to finance, in our framework the productive capacity *discounts*, in a loose sense of the word, the information encoded in the network of capabilities and in their mutual interaction. In a sense, what a country is able to produce and export is a summary statistic of its network of capabilities.

Specifically, economic complexity refers to a new line of research, which portrays economic determinants of growth as a process of evolution of ecosystems of technologies and industrial capabilities [Tacchella et al. (2012); Cristelli et al. (2013); Tacchella et al. (2013); Caldarelli et al. (2012); Pugliese et al. (2014)]. As a conceptual scheme, the observable bipartite network defined by countries and exported products can be thought of as a contraction of a tripartite network

where the intermediate layer is the capability level. This level acts as a set of hidden variables which represent the intangible assets driving the success and the potential of growth of an economic system. This capability level encodes the information on all the determinants of growth and in this perspective, production encodes all the features of these intangibles. Therefore products can be used to define a metric which quantitatively assesses capabilities even though they can not be directly observed.

This conceptual scheme acts as a classifier among economic dimensions and identifies a hierarchical structure, which marks a strong difference from standard economic approach, especially from the empirical literature of this field. The previously mentioned approaches in standard economics mainly try to find direct proxies for single capabilities, among widely different economic indicators, mostly via a regression approach. On one hand, such an approach suffers from the fact that we cannot achieve a comprehensive description and complete list of these determinants given the elusive definition of capabilities. On the other hand, the non-trivial networked interaction and the heterogeneity of capabilities are completely left out. As a simple example, usually the proposed determinants are combined in a additive way in a regression scheme while, in general, it is very difficult to assess whether the nature of an indicator is additive, multiplicative, mixed or even more complicated. It is also very difficult to determine the relative weights when such indices are built.

Differently, our conceptual scheme proposes an alternative strategy to the direct search for the capabilities and, in general, for the drivers of the growth. Our approach pintpoint the productive system as the level which reflects the complete information about capabilities: the capabilities themselves and their interactions. It is worth noticing that this general framework gives, at the same time, a scientific grounding to the empirical drivers of growth and to the forecasting schemes for economics. As we are going to see, this perspective indeed permits to measure the potential of growth of countries [Cristelli et al. (2014)] and we will use this framework to assess the possible paths of development of sub-Saharan economies.

In practice, this line of reasoning leads to the introduction of non-monetary metrics [Tacchella et al. (2012); Cristelli et al. (2013); Tacchella et al. (2013); Caldarelli et al. (2012)] for country competitiveness (fitness) and product complexity starting from export data of countries, as proxy of the productive system (see Appendix for further details on the dataset). This metrics is the fixed point of an iterative Google-like algorithm [Page et al. (1999)] for the bipartite country-product network defined by the binary country-product matrix. However, the nested structure of the bipartite country-product network calls for a non-linear methodology differently from Google's PageRank. We refer to section 3 for the mathematical details of the metrics.

The fitness is a new economic dimension which quantitatively addresses the complexity and diversity of countries' productive sets. It represents a synthetic measure of the whole network of the interaction among capabilities and consequently allows for quantifying the hidden growth potential of countries by the comparison with monetary figures such as GDP *per capita* [Cristelli et al. (2014)]. This comparison defines the *fitness-income plane* where we observe that country dynamics exhibits strongly heterogeneous patterns of evolution. The flow, in some zones, is found to be laminar while in others a chaotic behavior is instead observed. These two regimes correspond to very different predictability features of the country evolution: in the former regime, we find strong predictable patterns, while the latter scenario is characterized by a very low predictability. In such a heterogeneous framework, new concepts, borrowed from dynamical systems theory, are needed to develop a forecasting scheme rather than choosing a standard regression approach where the fitness or a derived variable would be used as a regressor. The use of a regression approach would inherently and implicitly assume a homogenous response framework in strong contrast with the intrinsic observed heterogeneity of country evolution.

The empirical finding of a high level of heterogeneity indeed suggests a strong similarity with the issue of forecasting the evolution of a dynamical system knowing only its past evolution [Lorenz (1969a, 1969b); Cecconi et al. (2012); Cristelli et al. (2014)].
In summary, Economic Complexity provides a new framework for forecasting the evolution of a dynamical system as in the case of weather dynamics [Lorenz (1969a, 1969b); Cristelli et al. (2014); Sugihara et al. (2012)].

3 Definition of the metrics: the fitness of countries and the complexity of products

The conceptual grounding of our metrics is similar to Google's PageRank methodology which allows to define a quantitative metrics to rank the importance of the nodes of the World Wide Web by extracting information from the network. But the mathematical framework is very different due to the nested structure[1] of the bipartite country-product network as shown in Fig. 1 where we report the country-product matrix in the year 2004 rearranged according to our metrics .

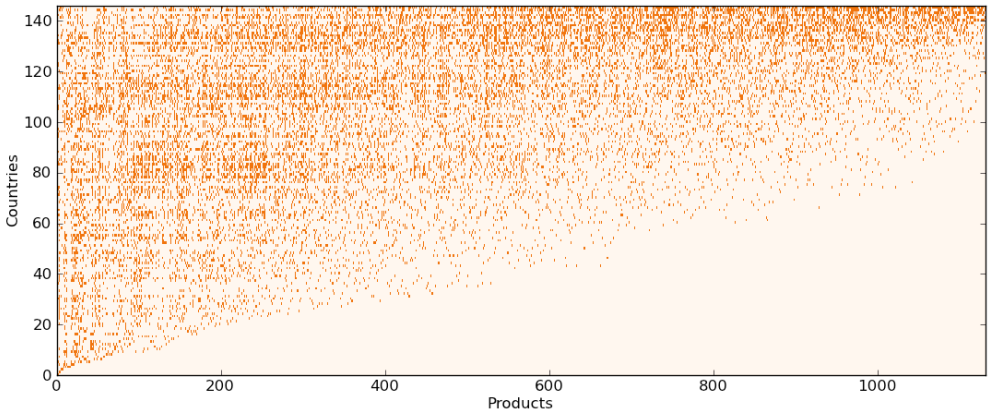

Figure 1: The nested country-product matrix for the year 2004.

The entries of this matrix can assume only two values, 0 or 1 depending on whether a country is a relevant producer of a product (i.e. whether the country has a Revealed Comparative Advantage (RCA) greater than 1, see Appendix for further details on how export flows in USD are made binary). The nestedness of this matrix defines the structure of the suitable mathematical framework to deal with the information content of the matrix: the fitness of countries and the complexity of products must be non-linearly coupled. We indeed get little insight into the capabilities required by a product if we only know that a highly diversified country is a producer of that product. On the other hand, we obtain insight into the upper bound for the complexity of a product if a poorly diversified country (and likely poorly fit) is one of its exporters. Due to the nested structure of the matrix, these countries tend to produce only the ubiquitous products made

[1] by nested we mean that specialists only interact with generalists, i.e. non-ubiquitous products are produced by highly-diversified countries and specialist countries export only ubiquitous products

by almost all countries and likely requiring fewer (or simpler) capabilities. Consequently, a product has a high level of complexity if it is exported by only a small number of countries with a high fitness and a country has a high level of fitness if it exports a large number of products, appropriately weighted by their complexities.

Our mathematical specification must therefore reflect the following properties:

- the fitness of a country is measured in terms of diversity of the productive set weighted by the complexity of products;
- the more countries are able to export a product, the lower is the complexity of the product;
- the upper bound of the complexity of a product must be dominated by the exporters with the lowest fitness.

The simplest mathematical translation of these concepts are the following self-consistent iterative coupled equations for the fitness F_c of countries and the complexity Q_p of products:

$$\begin{cases} \tilde{F}_c^{(n)} = \sum_p M_{cp} Q_p^{(n-1)} \\ \tilde{Q}_p^{(n)} = \frac{1}{\sum_c M_{cp} \frac{1}{F_c^{(n-1)}}} \end{cases} \rightarrow \begin{cases} F_c^{(n)} = \frac{\tilde{F}_c^{(n)}}{\langle \tilde{F}_c^{(n)} \rangle_c} \\ Q_p^{(n)} = \frac{\tilde{Q}_p^{(n)}}{\langle \tilde{Q}_p^{(n)} \rangle_p} \end{cases} \qquad (1)$$

where M_{cp} are the entries of the binary country-product matrix. This iterative method is made of two steps at each iteration: we first compute the intermediate variables $\tilde{F}_c^{(n)}$ and $\tilde{Q}_p^{(n)}$ and then normalize them. The initial conditions are $\tilde{Q}_p^{(0)}=1 \ \forall p$ and $\tilde{F}_c^{(0)}=1 \ \forall c$ but the method has been extensively tested and it has been numerically shown that the point of convergence of the map is independent on initial conditions [Cristelli et al. (2013)]. The metrics for the country fitness and the product complexity is defined as the fixed point of this process and numerically corresponds to the iteration at which the map reaches convergence [Pugliese et al. (2014)].

4 The heterogeneous dynamics of economic complexity: sub-Saharan growth scenarios

The main tool, within this novel framework, to assess growth scenarios is the fitness-income plane: a plane where we report the fitness versus the GDP *per capita* of countries. As previously argued, the comparison of monetary figures with non-monetary information can reveal non-trivial hidden potential of country development (we refer to [Cristelli et al. (2014)] for a detailed description of this scheme). We now discuss the results of a comparative analysis of sub-Saharan countries with respect to the development path of Asian emergent countries. We recall that the dynamics in this plane shows two main regimes: a laminar regime characterized by a high degree of predictability where the past evolution is a good predictor for future scenarios and a chaotic one where predictability drops and fitness is no more the main driver for understanding the development of countries.

In Fig. 2 we report the dynamics of evolution of 46 sub-Saharan countries (see Appendix A for the country list) from 1995 to 2012 and the trajectories of the Asian tigers, namely Vietnam (VNM), Thailand (THA), Malaysia (MYS), South Korea (KOR) and Singapore (SGP). The

dashed black line is the expected level of GDP *per capita* given the level of fitness[2]. Consequently a country below this line is not reflecting, in terms of income, the complexity of its productive system as measured by the fitness, while countries above the line have a level of income which can not be explained in terms of the complexity of their productive systems[3]. The red line instead schematically represents the zone of transition between the laminar regime with high predictability (right part of the plane) and the chaotic regime with low predictability (left part of the plane).

[2] This line is the result of the minimization of the Euclidean distance from the line weighted by the country GDP

[3] This last scenario does not necessarily imply economic instability because the origin of the wealth not explainable in terms of fitness might be due to exogenous sources such as services (not covered in our dataset) or raw materials. Many of the countries in the top left corner are indeed Gulf countries.

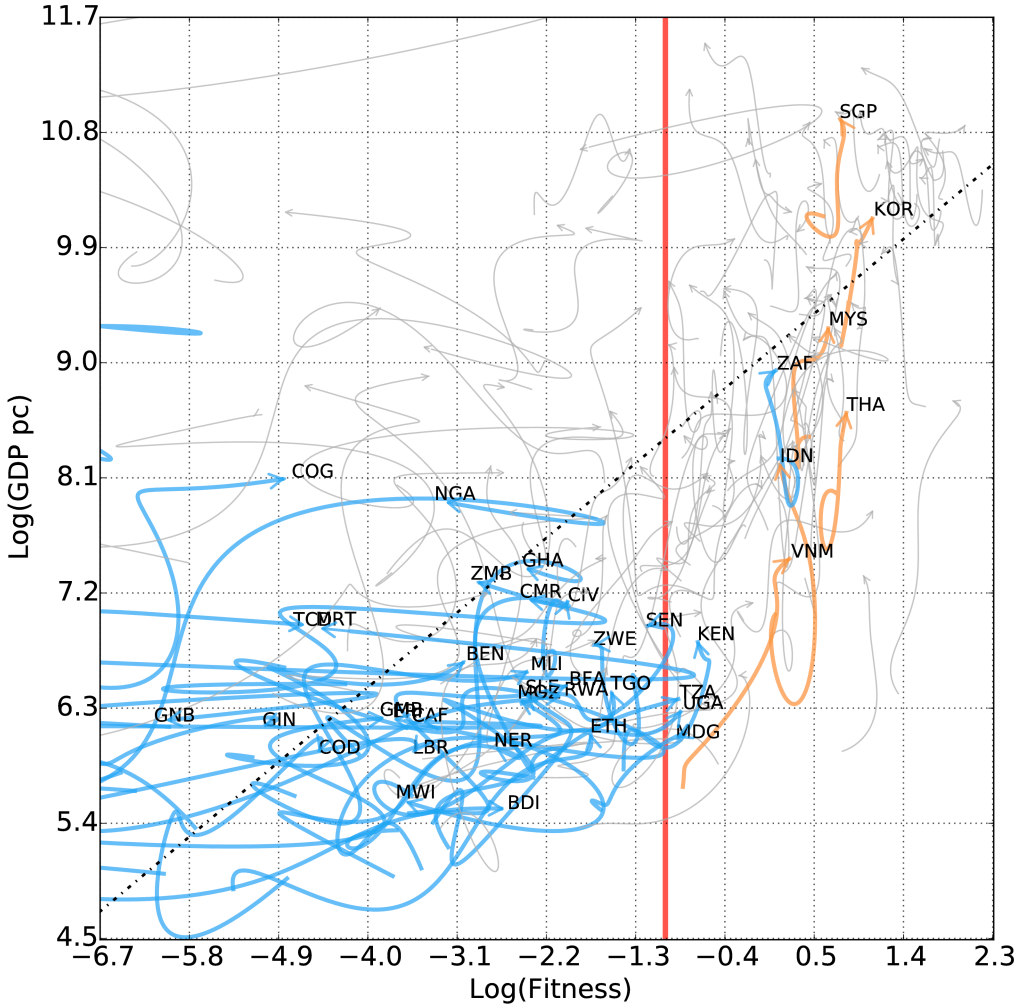

Figure 2: *The dynamics of country evolution from 1995 to 2012 in the fitness-income plane for sub-Saharan countries (light blue trajectories) and five Asian tigers (orange trajectories). The black dashed line represents the expected level of GDP per capita given the level of fitness, while the red vertical line defines the transition zone between the chaotic regime (left part of the plane) and the laminar one (right part). South Africa, as expected, appears in a completely different economic regime being the most mature economy of the southern part of Africa. Among the remaining countries, Senegal, Kenya, Tanzania, Madagascar and Uganda seem to have exited from the chaotic regime and from the poverty trap and are entering into a potential flow comparable with the position of Vietnam and Indonesia at the beginning of the nineties.*

Most sub-Saharan countries are concentrated (and stuck) in the bottom left corner of the chaotic regime which we identify as the poverty trap. In this region, capabilities-driven arguments

(and consequently the fitness) compete with several exogenous factors such as civil wars, corruption, natural disasters, famine, epidemics, corruption, which dump and prevent the start of any effective industrialization process.

Only Senegal, Kenya, Tanzania, Madagascar and Uganda seem nowadays to be undergoing the first steps of industrialization and the consequent exit from the chaotic regime and, in the meantime, from the poverty trap. The present position of these countries is compatible with a scenario in which these five countries will be able to enter into the stream followed by Asian emerging countries like Vietnam and Indonesia in the last two decades. Madagascar deserves a special remark: this country has undergone a severe political instability in the last 5 years and it would not be identified by standard methods as an economic systems in the early phases of industrialization. However, here we find that the country is entering into the laminar regime where exogenues instabilities are no more the main driver of the economic evolution. In addition, the economic is not only due to a diversfied export basket, as shown in Figs. 4 and 5 and Table 1 but also to a general trend of increasing production of high complexity products.

We also observe that a small increase of the complexity of the productive systems of countries in the poverty trap but close enough to the transition line (approximately $\log(F)=-2\sim-1$) may trigger a significant transformation of these economies and societies. Countries like Ethiopia, Togo might represent the second wave of african countries undergoing an industrialization and development process driven by the fitness.

The *selective predictability scheme* [Cristelli et al. (2014)] allows for an estimate of the growth rate of these five emerging countries in the period 2012-2022. This scheme is inspired by the so-called *method of analogues* [Lorenz (1969a)] developed within the domain of dynamical systems and the grounding idea of this approach is that, in the laminar regime, the past evolution of other countries which have passed through the same region of the plane is a provides a good forecast for the future evolution of countries now in that region. The more the dynamics is laminar, the higher is the predictability of this scheme. In 2012 Kenya, Tanzania, Madagascar and Uganda stay in the transition zone between the chaotic and the laminar regime characterized by an intermediate degree of predictability. We estimate that Kenya, Tanzania, and Uganda will have a nominal compound annual growth rate (CAGR) in the period 2012-2022 in the range 11-13% while Madagascar in the range 8-10%. The estimation of this CAGR are made in terms of US current dollars. Even if inflation and population growth are not included in our forecast, we observe that the CAGR of these four countries is well above the one forecasted for the world ($CAGR_{world}\approx 6-7\%$).

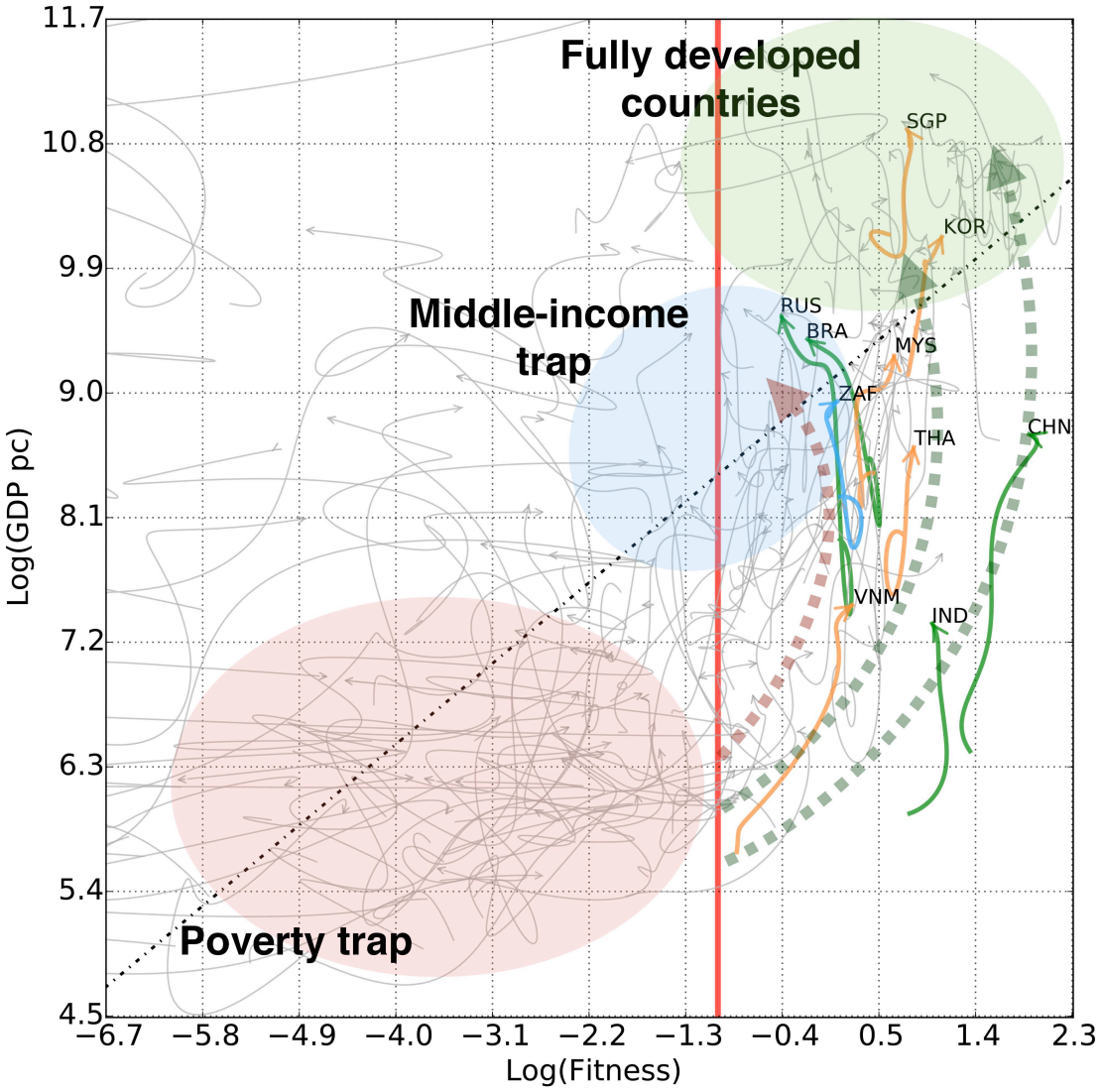

Figure 3: *Comparison of BRIC evolution (green trajectories) in the fitness-income plane from 1995 to 2012 with South Africa (light blue trajectory) and Asian Tigers (orange). Even though South Africa is the most mature African economic system, its trajectory reveals the concrete risk of getting stuck in the middle-income trap due to an incomplete diversification process in favor of a short term income growth. The three regions, namely the poverty trap, the middle-income trap and the developed countries must be viewed qualitatively.. The dashed arrows qualitatively represent the paths of countries which achieve full development (green lines) and of countries which get stuck in the middle-income trap (red line) due to a too quick conversion of their fitness potential into monetary wealth.*

As a further remark, Nigeria and Democratic Republic of Congo reveal a pattern of growth which is typical of raw material exporters like Gulf countries. They have indeed a high level of

GDP *per capita* with respect to other sub-Saharan countries (similar to Indonesia and Vietnam's wealth) but a very low level of sophistication of the productive system.
South Africa deserves a specific analysis since, as our scheme confirms, it emerges as the most mature economic system of the southern part of Africa and, likely, of the continent. South Africa (ZAF) is well-outside the poverty trap and its position is steadily in the laminar regime. The comparison of the evolution of South Africa with the behavior of BRIC countries, namely Brazil, Russia, India and China, is shown in Fig. 3. This reveals non-trivial features of the industrialization process of South Africa. BRIC countries, even though extremely similar in terms of GDP growth, exhibit very different patterns of development. China and, to a less extent, India are undergoing a real diversification of their productive systems. As a result, China and India underwent a delayed increase of the monetary wealth (i.e. GDP per capita) but this allowed for the development of a strong technological ecosystem, which nowadays is competing with western developed countries. Brazil and Russia has instead shifted to short time monetary return rather than expanding their high fitness potential and fueled their growth by means of the primary sector. South Africa's behavior in the period from 1995 to 2012 is close to the one of Brazil and Russia. In qualitative terms, it can be argued that this poor development of the fitness potential seems to lead these countries to an incomplete industrialization process (see Fig. 3) characterized by an income which sets Russia, Brazil and South Africa well outside the poverty trap but below the region occupied by fully developed countries (top right corner).

5 Spectroscopy of the emerging sub-Saharan countries

Spectroscopy is a tool to visualize the structure of the productive system of a country in terms of its products' complexity. Products are ordered according their complexity and, along this axis, we plot the export values expressed in USD. In this way, we are able to pinpoint structural features of the value distribution of a specific country. In fig. 4 we compare the spectroscopy of the five sub-Saharan countries in 2010 with Vietnam's spectroscopy in 1995. On a qualitative basis, the spectroscopies of the productive systems of these five countries show a marked similarity with the spectroscopy of Vietnam in 1995.

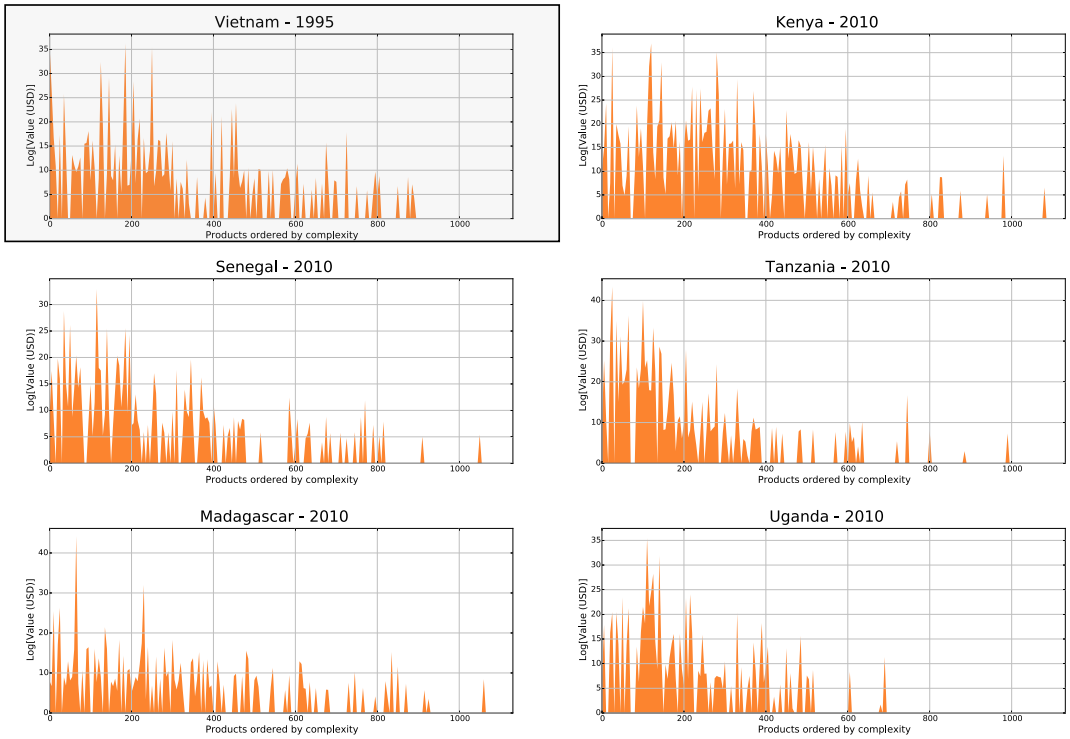

Figure 4: *Spectroscopies of Senegal, Kenya, Tanzania, Madagascar and Uganda in 2010 compared with that of Vietnam in 1995 (top right panel). Export values of product (vertical axis) are ordered according to increasing product complexity (horizontal axis).*

In order to assess the evolution of the spectroscopies in the time interval 1995-2010, in Fig. 5 we evaluate for each country and for each year the box plot of the spectroscopy thought of as a probability distribution.

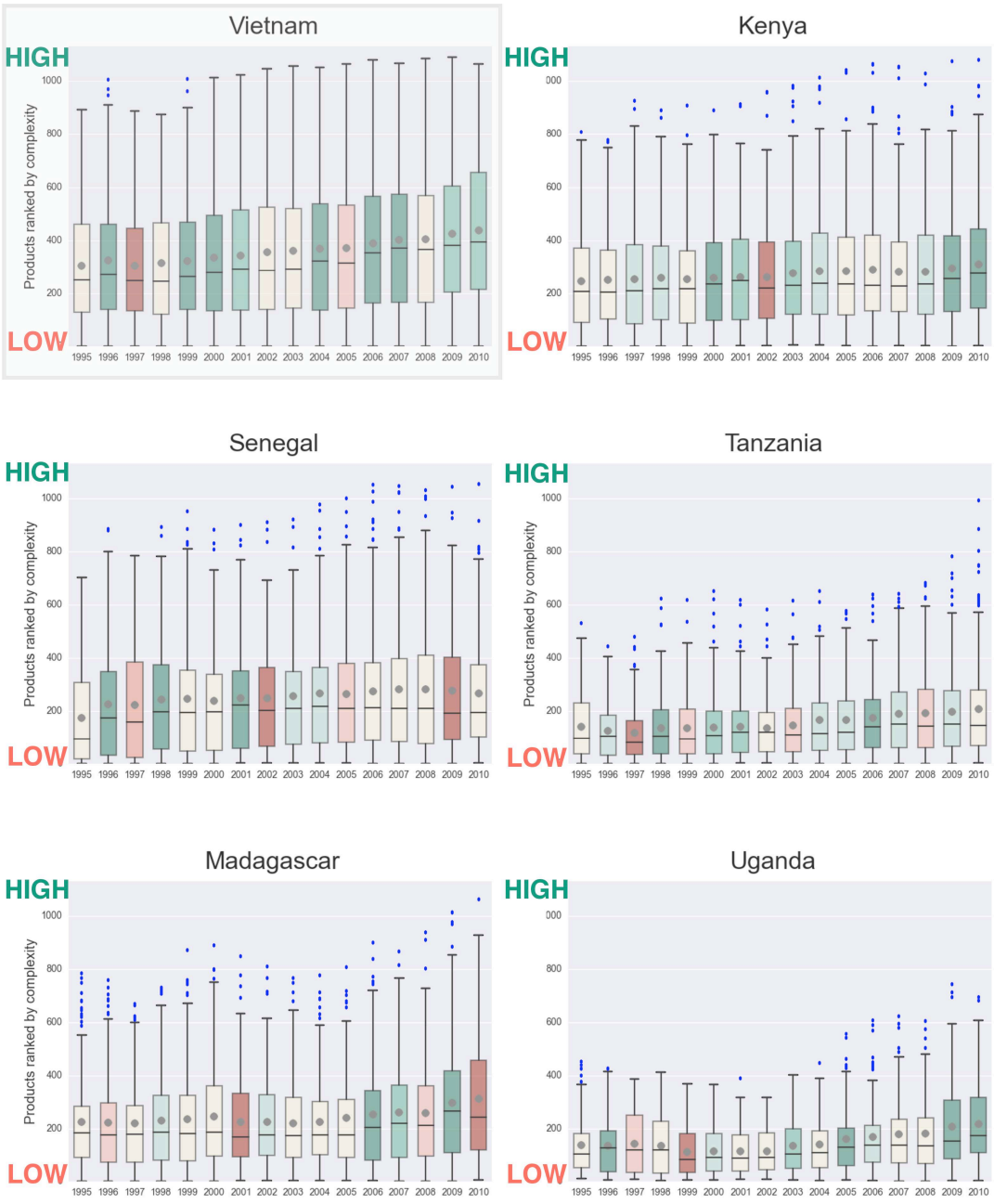

***Figure 5**: Time evolution from 1995 to 2010 of spectroscopy box-plots for Senegal, Kenya, Tanzania, Madagascar and Uganda compared with that of Vietnam (top right panel). The box borders correspond to the first and third quartile of the spectroscopy, the grey dot is the mean while the horizontal black line is the median of the spectroscopy. The color intensity measures the yearly variations of the median. The darker is the box, the higher is the variation (green for positive and red for negative changes). The pale yellow boxes refer to a median change lower than 5 positions.*

The evolution of the spectroscopy supports the scenario of economies in the early stages of the exit from poverty trap which are achieving a stable development regime. This is also witnessed by the box-plot range which is, as a general picture, increasing towards high complex products .
Following [Pugliese et al. (2014)], we can estimate the two contributions to the fitness change: the diversification-driven term and complexity-driven one.

	Variation due to changes in the number of products in the export basket	Variation due to changes in the products' complexity
Senegal	+478%	-75%
Kenya	+158%	+68%
Tanzania	+408%	+218%
Madagascar	+119%	+81%
Uganda	+512%	+247%
South Africa	+0.48%	+3.4%
Vietnam	+156%	+89%

Table 1: *Decomposition of fitness variations in the two components for the time period 1995-2010. The variations are reported as a percentage of the initial fitness value for the time span in question*.

The five emerging countries are all characterized by both a strong increase in diversification and a strong complexification of the export basket expect for Senegal where the fitness increase is fueled only via diversification. Once again, the results of this analysis confirm a strong analogy between the pattern of development of these countries and that of Vietnam (last row of Table 1).

On the other hand, the analysis confirms the static and stagnant scenario for South Africa whose growth potential is losing momentum.

6 Discussions

Economic complexity is a data-driven framework, which aims at casting economic theory (and economic analysis) in a more scientific approach.. In this paper, we focus on the paths of country development and the determinants of growth and we specifically analyze sub-Saharan countries. Differently from the standard approach, we do not directly try to assess the role of a specific set of drivers of growth via a regression model. In our framework, the drivers of development – in our jargon, capabilities and their network of interaction – play the role of hidden variables whereof information is encoded by the productive system of a country, as proxied by the export basket. We therefore measure the effect of all determinants as a whole by introducing two metrics: the fitness of countries and the complexity of products. The comparison of the information from this new economic dimension for countries and the one coming from standard monetary figures (e.g. GDP per capita) sheds light on the hidden growth potential of countries. The application of this scheme to 46 sub-Saharan countries provide the following scenarios:

- most countries are trapped in a subsistence regime where no relevant signals of sustained development are observed (poverty trap);

- although the majority of these countries have huge raw materials reserves, only two (namely Democratic Republic of Congo and Nigeria) have a pattern similar to other raw-material exporters as Russia and Gulf economies. We argue that it can be due to a still limited exploitation of these resources or to a structural and political inability to hold back the capitals deriving from these activities;
- only Uganda, Senegal, Kenya, Madagascar and Tanzania appear to have started an industrialization process similar, in its early steps, to the one achieved by Asian tigers;
- South Africa is the only mature economic system even though its path towards full development is threatened by the risk of getting stuck in the middle-income trap due to its poor diversification which undermines the growth potential.

Our framework provides direct policy implications and, specifically, suggests policy interventions for optimal long term industrial planning. In particular, using techniques to forecast the most probable evolution of the export basket of countries [Zaccaria et al. (2014)], it is possible to boost growth or, as for South Africa, reduce the risk of entering the middle-income trap by means of a country-specific industrial policy favoring the diversification process on strategic high-complex sectors.

A Materials and Methods

We use data extracted from the BACI dataset [Gaulier & Zignago (2009)]. This dataset is, in turn, grounded on the COMTRADE dataset (freely accessible from the UN Comtrade website[4]) and is the result of a reconciliation procedure extensively discussed in [Gaulier & Zignago (2009)]. This dataset, on a yearly basis, reports trading data about more than 200 countries and 5000 products classified according to a six digit code (categorization: Harmonized System 2007[5]). We have aggregated the dataset at a 4 digits level (the maximum detail corresponds to 6 digits level). Products are 1131 at this level of aggregation.

In order to binarize the raw country-product matrix whereof the entries are the export flow of a given product of a given countries in a specific year expressed in USD, we adopt the so-called Revealed Comparative Advantage (RCA) criterion [Balassa (1965)]. The RCA matrix defines the binary product-country matrix M in the following way: we say that a country c can be considered a relevant producer of a product p if $RCA_{cp}>1$ and therefore the entry M_{cp} is set equal to 1, conversely $M_{cp}=0$ if $RCA_{cp}\leq 1$. The RCA is defined as the ratio between the fraction of a product within the export basket of a country and the fraction of that product over the world export, in formula:

$$RCA_{cp} = \frac{\frac{q_{cp}}{\sum_{p'} q_{cp'}}}{\frac{\sum_{c'} q_{c'p}}{\sum_{c'p'} q_{c'p'}}} \qquad (2)$$

where q_{cp} are the entries of the raw export matrix.

[4] http://comtrade.un.org

[5] http://www.wcoomd.org

All the remaining economic indicators are extracted from the World Bank dataset[6].

We report the 46 sub-Saharan countries and their codes used throughout this report: Angola (AGO), Burundi (BDI), Democratic Republic of the Congo (COD), Rwanda (RWA), Sao Tome and Principe (STP), Cameroon (CMR), Central African Republic (CAF), Chad (TCD), Republic of the Congo (COG), Equatorial Guinea (GNQ), Gabon (GAB), Kenya (KEN), Tanzania (TZA), Uganda (UGA), Eritrea (ERI), Ethiopia (ETH), Somalia (SOM), Botswana (BWA), Comoros (COM), Lesotho (LSO), Madagascar (MDG), Malawi (MWI), Mauritius (MUS), Mozambique (MOZ), Namibia (NAM), Seychelles (SYC), South Africa (ZAF), Swaziland (SWZ), Zambia (ZMB), Zimbabwe (ZWE), Benin (BEN), Mali (MLI), Burkina Faso (BFA), Cape Verde (CPV), Cote d'Ivoire (CIV), Gambia (GMB), Ghana (GHA), Guinea (GIN), Guinea-Bissau (GNB), Liberia (LBR), Mauritania (MRT), Niger (NER), Nigeria (NGA), Senegal (SEN), Sierra Leone (SLE), Togo (TGO).

Acknowledgements

All the authors thank Yaw Ansu, Chief Economist at ACET (African Center for Economic Transformation), for helpful discussions and suggestions. Authors acknowledge CNR Progetto di Interesse "CRISIS LAB" and EU Project nr. 611272 "GROWTHCOM".

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth*, *1*, 385–472.

Aghion, P. & Howitt, P. (1992). A Model of Growth through Creative Destruction. *Econometrica, Econometric Society*, *60*(2), 323–51.

Ansu, Y. (2014). Growth with Depth. Technical report, African Center for Economic Transformation.

Anyanwu, J. C. (2014). Factors affecting economic growth in africa: Are there any lessons from china? *African Development Review*, *26*(3), 468–493.

Ashraf, Q. & Galor, O. (2013). The out of africa hypothesis, human genetic diversity, and comparative economic development. *The American Economic Review*, *103*(1), 1–46.

Auerbach, A. J., Hassett, K. A., & Oliner, S. D. (1993). Reassessing the social returns to equipment investment. Technical report, National Bureau of Economic Research.

Balassa, B. (1965). Trade liberalization and 'revealed' comparative advantage. *Manchester School*, *33*, 99–123.

Beinhocker, E. D. (2006). The origin of wealth. *Harvard Business School Press, Boston*, 254–258.

Bleaney, M. & Nishiyama, A. (2002). Explaining growth: a contest between models. *Journal of Economic Growth*, *7*(1), 43–56.

[6] `http://data.worldbank.org`

Caldarelli, G., Cristelli, M., Gabrielli, A., Pietronero, L., Scala, A., & Tacchella, A. (2012). A network analysis of countries export flows: firm grounds for the building blocks of the economy. *PloS one*, *7*(10), e47278.

Cecconi, F., Cencini, M., Falcioni, M., & Vulpiani, A. (2012). Predicting the future from the past: An old problem from a modern perspective. *Ann. J. Phys.*, *80*, 1001.

Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G., & Pietronero, L. (2013). Measuring the intangibles: A metrics for the economic complexity of countries and products. *PloS one*, *8*(8), e70726.

Cristelli, M., Tacchella, A., & Pietronero, L. (2014). The heterogeneous dynamics of economic complexity. *submitted to PLoS ONE*.

Diamond, J. M. & Ordunio, D. (2005). *Guns, germs, and steel*. National Geographic.

Eichengreen, B., Park, D., & Shin, K. (2013). Growth slowdowns redux: New evidence on the middle-income trap. Technical report, National Bureau of Economic Research.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, *25*(2), 383–417.

Galor, O. & Weil, D. N. (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American economic review*, 806–828.

Gaulier, G. & Zignago, S. (2009). Baci: International trade database at the product-level.

Heckscher, E. F. & Ohlin, B. G. (1991). *Heckscher-Ohlin trade theory*. The MIT Press.

Hidalgo, C. A. & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, *106*(26), 10570–10575.

Kharas, H. & Kohli, H. (2011). What is the middle income trap, why do countries fall into it, and how can it be avoided? *Global Journal of Emerging Market Economies*, *3*(3), 281–289.

Kormendi, R. C. & Meguire, P. G. (1985). Macroeconomic determinants of growth: cross-country evidence. *Journal of Monetary economics*, *16*(2), 141–163.

Lall, S. (2000). The technological structure and performance of developing country manufactured exports, 1985-98. *Oxford development studies*, *28*(3), 337–369.

Leontief, W. (1953). Domestic production and foreign trade; the american capital position re-examined. *Proceedings of the American philosophical Society*, 332–349.

Leontief, W. (1956). Factor proportions and the structure of american trade: further theoretical and empirical analysis. *The Review of Economics and Statistics*, 386–407.

Levine, R. & Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *The American economic review*, 942–963.

Lorenz, E. N. (1969a). Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci*, *26*, 636–646.

Lorenz, E. N. (1969b). Three approaches to atmospheric predictability. *Bull. Am. Meteorol.*, *50*, 345–349.

Murphy, K. M., Shleifer, A., & Vishny, R. W. (1988). Industrialization and the big push.

Ohno, K. (2009). Avoiding the middle-income trap: renovating industrial policy formulation in vietnam. *ASEAN Economic Bulletin*, *26*(1), 25–43.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web.

Pugliese, E., Chiarotti, G. L., Zaccaria, A., & Pietronero, L. (to appear, 2014). Economic complexity as a determinant of the industrialization of countries: the case of india.

Pugliese, E., Zaccaria, A., & Pietronero, L. (2014). On the convergence of the fitness-complexity algorithm. *Working paper*.

Ricardo, D. (1891). *Principles of political economy and taxation*. G. Bell and sons.

Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy*, S71–S102.

Sachs, J. D., Warner, A., Åslund, A., & Fischer, S. (1995). Economic reform and the process of global integration. *Brookings papers on economic activity*, 1–118.

Schumpeter, J. A. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle*, volume 55. Transaction Publishers.

Schumpeter, J. A. (1939). *Business cycles*, volume 1. Cambridge Univ Press.

Smith, A. (1776). The wealth of nations. *New York: The Modern Library*.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, 65–94.

Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *science*, *338*(6106), 496–500.

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., & Pietronero, L. (2012). A new metrics for countries' fitness and products' complexity. *Scientific reports*, *2*.

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., & Pietronero, L. (2013). Economic complexity: conceptual grounding of a new metrics for global competitiveness. *Journal of Economic Dynamics and Control*, *37*(8), 1683–1691.

Tacchella, A., Di Clemente R., G., Gabrielli, A., & Pietronero, L. (2014). The build-up of diversity in complex eco-systems. *In preparation*.

Weber, M. (2002). The protestant ethic and the spirit of capitalism and other writings, edited, translated, and with an introduction and notes by peter baehr and gordon c. wells.

Zaccaria, A., Cristelli, M., Tacchella, A., & Pietronero, L. (2014). How the taxonomy of products drives the economic development of countries. *accepted for publication in PLoS ONE, http://arxiv.org/abs/1408.2138*.