WPS8117

Policy Research Working Paper 8117

On the Predictability of Growth

Matthieu Cristelli
Andrea Tacchella
Masud Cader
Kirstin Roster
Luciano Pietronero

WORLD BANK GROUP
IFC Country Economics & Engagement
June 2017

Policy Research Working Paper 8117

Abstract

A country's productive structure and competitiveness are harbingers of growth. Growth is a dynamic process based on capabilities that are difficult to define and measure across countries. This paper uses a global measure of fitness (or complexity-weighted diversity of production) as a method to explore a country's relative growth potential. The analysis finds that there are two types of growth, predictable or laminar, and unpredictable. This classification is used to create a selection mechanism (the Selective Predictability Scheme), defining future growth trajectories for similar countries, and compares projected long-term, five-year forecasts with traditional methods used by the International Monetary Fund. The analysis finds that production structure is a good long-term predictor of growth, with prediction performance falling off for countries not yet in the laminar classification.

This paper is a product of the IFC Country Economics & Engagement. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at mcader@ifc.org.

On the Predictability of Growth

How Industrial Structure Predicts Long-Term Growth

Matthieu Cristelli, Andrea Tacchella, Masud Cader, Kirstin Roster,
Luciano Pietronero

JEL classification: O11, O14, O47, O50, O57, E17, F14, F43, F47, C14, C40
Keywords: growth projections, long-term forecast, growth predictability, growth potential, economic complexity, productive structure, country competitiveness, capabilities, diversification.

1. Introduction

1.1 How to select the past to model future growth: A long-standing question and a concrete answer

Growth is the aggregate result of uncountable interactions among economic actors occurring at different temporal and geographical scales and its modeling still represents one of the main challenges of economics. Modeling the conditions leading countries to increase competitiveness and enabling country growth,[1] in fact, has a pivotal role. On one hand the understanding of the growth process allows to provide the suitable tools to design the most effective and inclusive economic policies with the ultimate goal to foster broad economic growth. On the other hand, economic practice is rooted in the concept of expectation and consistent modeling is essential to provide consistent forecasts.

In this work we aim to show that countries' productive structures, and more specifically cross-country differences in productive structure, are a good mid-long term predictor of economic performance. Specifically, on the basis of a recently proposed complexity weighted production diversity measure - fitness [Tacchella, 2012] - we can define the fitness-GDP *per capita* plane which allows to select in the past the closest comparators for the country we aim to forecast.

The definition of this new space carries two pivotal conceptual points. Our forecasting scheme is, conceptually speaking, rooted in what is called, with the jargon of dynamical systems, the *method of the analogues* [Lorenz, 1969]. Let us suppose we want to forecast the evolution trajectory of an event of this system a number of periods ahead and we do not know the rules of evolution of the system (i.e., the underlying equations). Provided some degree of stationarity of the system[2] and provided a way to measure the distance between events in their space of evolution, we can devise, in principle, an alternative approach to obtain a data-driven assessment of this projection: use the past to model the future. We can look for the closest events to the event we observed in the past and use their evolution to model the projection we want to assess. However, this approach is often unfeasible because it faces the so-called *curse of dimensionality* [Bellman, 1957; Cecconi et al., 2012]: either the system has an extremely low number of dimensions (i.e., 2-3) or we need an exponentially growing past statistics to be successful. The latter condition is usually hardly achievable, especially for economic data.

We need therefore an extra step to model the economic future with the economic past in a scientific and successful way because the development of countries is the result of uncountable interactions and factors at different levels. Is there an economic level which comprehensively and inherently is the result of all these factors? The fitness (and in general the branch of economics named economic complexity) [Hausmann et al. 2007; Hidalgo et al., 2009; Tacchella et al., 2012] aims to be a framework to provide a positive

---

[1] In this work we refer to growth on the long run. Throughout the text we will refer to it simply as growth.

[2] The stationarity is a subtle point which is hardly assessable. However, we can pragmatically deal with this point the other way around. Until this strategy works, we can argue that, to some extent, stationarity is at work and tracking the forecasting power of the strategy in time allows to have insights into this feature.

and concrete answer to this challenge. The level encoding all the factors and their interaction is identified as the economic output of a country evaluated on a competitive basis. In details the cross-country differences of the productive capacities represent the arena to define an aggregate measure of countries' competitiveness. In the literature, several attempts have been proposed to decode this competitiveness, see for instance [Lall, 2001, Hausmann, 2007, Hidalgo, 2009, Tacchella, 2012, Cristelli, 2013] and there are examples of attempts to build an effective low dimensional space in which to embed and predict countries' dynamics, using a competitiveness dimension together with a measure of GDP [McArthur, 2001]. The dynamical part has been treated with linear regression models in order to predict GDP growth [Hidalgo, 2009; Podobnik, 2012]. Such approaches have been shown to be flawed because of the insufficient quality of the competitiveness measures used [Lall, 2001; Cristelli 2013]. Here we show that a much more effective low dimensional space can be built by using fitness together with *per capita* GDP.[3] The validity of such choice can be appreciated a-posteriori, by observing an emerging predictable dynamical structure[4] (we refer to the next sections for further details on this aspect).

To clarify even more, we can interpret the practical implementation of our scheme, named the *Selective Predictability Scheme* (SPS hereinafter) with the jargon of the kernel regression [Nadaraya, 1964; Watson, 1964]. We want to stress that this identification is only formal and it is made to explain the practical SPS specifications. The kernel regression identification alone would not allow why SPS is a scientific framework to model the future with the past to be appreciated.

SPS works as a two-dimension kernel regression, where the two dimensions are the fitness and the GDP *per capita.* The kernel i) selects those past events which are in the neighborhood of the event we want to forecast, i.e., the set of the closest past we previously mentioned and ii) averages the *k*-period ahead displacements of those selected events. This average displacement provides the projection for the evolution trajectory of both the country competitiveness (the fitness) and the country GDP *per capita* (the growth).

The key point and the novelty of this approach is that, in order to build a reliable predictor from past events, we need to properly evaluate the *neighbors* of the future development we want to forecast. The wealth dimension alone is not enough to filter out these neighbors as two countries may have achieved the same level of GDP *per capita* for extremely heterogeneous factors but simultaneously we cannot directly use all these discriminating factors because this task is made unfeasible by the curse of dimensionality. The fitness wants to be a synthetic measure to include in a feasible way those factors and their (complex) interaction by assessing their emerging complexity from the differences of the productive structure across countries.

The forecast scheme we devise is extremely parsimonious in terms of model complexity but, still, it performs comparably (and in some regimes outperforms) with the IMF's 5-year projections. In addition, the scheme allows us to define an empirical measure for the predictability of a country's economic growth depending on the stage of its development, which allows to group countries into predictable and unpredictable types, with such groupings evolving in time. The groupings derived from our scheme can be used to separate IMF projections as well: we find that the prediction performances are higher for

[3] *We will use GDP per capita PPP in current USD.*

[4] *We instead refer to [Tacchella et al., 2012] for a technical discussion on why GDPpc and not GDP is the proper counterpart for the Fitness dimension.*

predictable type of events also in this case. This underpins that the predictability dimension quantifies a fundamental feature of a country's economic status, and is not a concept arising only in our forecasting scheme.

It is also worth to stress that in this work we are dealing with a forecasting problem, i.e. the country productive structure is a predictor for long-term development, and we are not addressing whether countries' productive structure *causes* growth or vice versa (i.e. causation problem). This is a crucial difference to settle in the perspective of the identification problem [Fisher, 1966] which only applies to the latter class of problems. We refer to [Kleinberg, 2015] for a general discussion on this aspect and on the importance of prediction issues in the policy perspective.

The effort to reconcile these approaches with standard economic theory to provide them an economic foundation is beyond the scope of this work. The present paper has the sole goal to discuss how economic modeling and economic practice can benefit from such types of approaches and how they can enhance and complement long-term schemes for growth forecasting by rooting in a data-driven framework the modeling of the selection of growth comparatives.

1.2 Paper organization

The paper is organized as follows:

1. In the remainder of this introductive section, we discuss how the economic complexity framework relates to growth forecasting and modeling, we deepen the discussion concerning its conceptual grounding and the economic implications and we summarize the main result of the paper.
2. In section 2, we detail the general concepts backing the Fitness dimension as a proxy for the country competitiveness. We also briefly discuss the data sets used in this work.
3. In the section 3, we present the main SPS results, namely:
    i. We compare SPS performances with WEO-IMF projections and three limited intelligence univariate models;
    ii. We discuss the meaning of the economic dimension *Predictability* that we define within SPS.

4. In Section 4, we present two case studies, Thailand and Ghana, in order to concretely illustrate how the Predictability dimension reflects a higher/lower degree of coherence of the comparator events' trajectories – the past events we select to be the closest to the event we want to forecast.

5. In Section 5, we provide a detailed technical description of the mathematical specifications of SPS.

6. Section 6 is devoted to a general discussion of how SPS can provide and enhance present economic practice.

1.3 How economic complexity relates to growth modeling

Starting from Adam Smith, economics has witnessed several attempts to tackle economic growth modeling which fall into two major groups: empirical data-driven approaches and theoretical ones. The former are heterogeneous in terms of methods but they all share a common guideline: they try to model growth from the knowledge of the past. Turning towards theoretical approaches, limiting our attention to the last decades and proceeding in a non-exhaustive way, the forefather of modern theoretical approaches to model growth is represented by the Solow model [Solow, 1956] where the aggregate output of an economy is modeled in terms of capital and labor to model productivity. Successive attempts have tried to increasingly describe economic dynamics as an endogenous process. The original Solow model is indeed a purely exogenous model. As examples of this tendency, in [Romer, 1986] Romer proposes to endogenize innovation while in [Aghion, 1990] Aghion and Howitt propose a model to endogenize the Schumpeterian concept of creative destruction.

Growth and in general aggregate economic output are, as mentioned, the results of complex and heterogeneous interactions occurring at different scales. However, the economic mainstream tends to neglect the possible effects of these complex interactions.

For instance, variety of productions, of inputs, of technology etc. are usually reconciled within *de facto* representative product, input, technology, firm, etc. models. In such a way modeling underestimates the role of behaviors that can emerge only at aggregate scales (meso and macro) but cannot be directly related to features of the representative actors/dimensions. The last decade has witnessed the build-up of the awareness that the heterogeneity and complexity of these interactions must be included in the description of economic aggregates features in order to properly deal with their dynamics.

Coherently with this new vision, a number of works have underpinned that countries' productive structure [Hausmann, 2007] and competitiveness are harbingers of growth. Country productive structure, defined on a competitive basis, is proposed as a proxy that carries information about the capabilities owned by a country. Capabilities are country endowments whose combinations define what a country is able to produce and compete on. Thus competitiveness and growth turn out to be a dynamic process based on capabilities' dynamics: the more capabilities a country acquires, the more likely new competitive production variety will be accessible. [5] However, differently from a representative capability framework, the extreme heterogeneity of these endowments makes a direct tackling unfeasible as they are both difficult to define and measure across countries as we have previously mentioned in section 1.1. The stand-off can be overtaken by reversing the usual way to proceed between endowments and economic output, not anymore from the former to the latter but vice-versa: cross-country differences of production structure can be leveraged to assess a country's relative competitiveness.[6] Seminal works in this direction are the indexes proposed in [Lall, 2001] and in [Hausmann, 2007]. A further step is represented by [Hidalgo, 2009] which, differently from previous approaches, explicitly accounts for the heterogeneous networked nature of the production structure. However, as discussed in [Tacchella, 2012] and [Cristelli, 2013], more sophisticated non-linear mathematical specifications are required in order to properly define an index of complexity weighted diversity of production (Fitness) which is consistent with the statistical features of the production network. For a detailed discussion we refer to [Cristelli, 2013] and Section 2.

1.4 More on the curse of dimensionality

A direct selection of the past in order to model the future is unfortunately a problem which is intractable as the state of an economic system, as mentioned previously, is specified, conceptually speaking, by a very large set of endowments and, in practice, by thousands of economic indicators. The time evolution of this system, i.e. the development trajectories, is the result of the interaction among these economic dimensions. These

[5] *It can be shown that the production variety gain (i.e. production diversification) as a function of the number of capabilities is an exponential.*

[6] *Further details on the conceptual grounding of this approach are provided in the following sections.*

dimensions are known under several names in the economic literature with slightly different meaning according to the specific field: determinants of growth in empirical studies, capabilities in trade-related works and endowments for the theory of economic growth.

The evolution of a country's economic development creates a trajectory which technically can be compared to other countries which went through a similar trajectory. While intuitively appealing, this type of comparative thinking is affected by the curse of dimensionality [Bellman, 1957; Cecconi et al., 2012] – i.e., as you measure economic performance in more ways, it becomes difficult to state anything definitive about the similarity of different economies without unrealistically long histories of countries' evolution. Provided we cannot *forge* longer histories of countries for obvious reasons, a natural way to deal with such a scenario is then to define a suitable procedure to reduce the dimensionality of the problem in order to pinpoint the driving dimensions.

Economic literature has proposed a number of strategies to implement such dimensionality reduction. However, most of these strategies share a common feature: they all try to reduce the dimensionality of the space by building an index or any form of compact description as a combination of the endowments (or capabilities, we will use the terms as synonyms throughout the paper). Among the most popular there are direct (linear) combinations of economic variables either serving as regressors [Barro, 1991] or as pillars to build new economic indexes such as the Global Competitiveness Index [WEF, 2016]. More refined approaches treat the problem of the identification of relevant features in a large data set as a statistical optimization problem, treated with techniques ranging from Principal Component Analysis [Jolliffe, 2002], to Self-Organizing Maps [Kohonen, 1982], to other kinds of shallow [Mikolov, 2013] and deep [Bengio, 2007] neural networks.

1.5 Reversing the approach: From economic output to endowments, and the fitness-GDPpc plane

The approach of New Development Economics (i.e. economic complexity) is to reverse this perspective and adopt a different strategy in order to perform the dimension reduction of the space embedding economic dynamics. We do not go from endowments to final output anymore, the information flow goes exactly in the opposite direction. Namely it is the set of produced products that informs on the capabilities of a country and its potential competitiveness in a compact description.

This dimension reduction achieved by reversing the way we proceed between economic output and endowments belongs to the complexity paradigm, specifically, the economic complexity paradigm.[7] The complexity paradigm provides a natural way of thinking to properly deal with adaptive, competitive and heterogeneous systems. In particular, this paradigm provides the playground to define and model the macroscopic and aggregated description of those systems where the interaction heterogeneity among their actors, attributes and activities is a non-negligible element. In our economic setting, actors read as economic actors, specifically countries, attributes as endowments (or capabilities) and activities as economic output, specifically exported products.

[7] *Finding an economic level encoding the relevant endowments and defining a decoding framework to end up with a few new indicators (i.e. few dimensions) is, conceptually speaking, not equivalent to find the 'right' variable (for example to feed a regression model) out of hundreds indicators.*

SPS roots can be therefore traced back in a number of efforts which could be summarized in the fact that what you export matters [Hausmann et al. 2007; Hidalgo et al., 2009; Tacchella et al., 2012]. The fundamental idea backing these approaches is that the differences in capabilities among countries drive and determine the differences in the export basket of countries. More specifically the differences in capabilities are the drivers of qualitative more than quantitative differences in the export basket of countries. Thus the specific diversity of the set of exported products comprehensively encodes all those dimensions which can drive economic competitiveness.

Fitness is then both a measure of the relative uniqueness of a country's production as well as the diversity of its production capabilities. By taking these two aspects into account we can define a fitness measure, and the competitiveness dynamics of country development which yields insights on development traps, selecting comparator countries, and projected growth.

To define this fitness measure, we follow the mathematical specifications proposed in [Tacchella et al., 2012] and in [Cristelli et al., 2013]. These specifications are consistently defined with respect to the constraints set by the statistical features of the export basket diversity differently from previous attempts in this field [Cristelli et al., 2013] (we refer to Section 2 for further details).

The resulting 2-dimensional space, the fitness-GDPpc plane, is the space in which we select the right comparatives to model growth projections. We find indeed that the notion of closeness is highly effective to select those comparatives in the past which are informative on the future we want to estimate. This backs and strengthens the effectiveness of our dimension reduction scheme. Moreover, the *regularity* of the economic evolution in this plane turns to be a proxy for the *predictability* of the economic growth. On this account it introduces the concept of heterogeneity of the growth predictability and the predictability dimension, which we discuss in Section 5. This dimension is proven to be effective not only for SPS projections but also for other sources of long-term growth projections, such as the ones provided by the IMF. Growth predictability as measured by the SPS is then a standalone piece of information which provides an assessment of the confidence of the projected growth rates underpinning a real feature of economic systems specific to the stage of development of a country and regardless of the source of the growth projections.

1.6 Summary of the main results

The comparison with traditional methods used by the IMF shows that the SPS model and in general a country's productive structure is a good and parsimonious predictor of growth in the long run.[8] This work bears three major results:

2. **Production structure as measured by the Fitness dimension can be used to model growth over long term horizons** (5+ years) by defining, together with an intensive measure of wealth, an effective dynamical space for the SPS. SPS defines an operative and generalizable procedure to select comparator countries and trajectories to model future GDP growth projections.

---

[8] *It is worth noticing that the specifications here proposed for the SPS are the minimal ones. This implies that there exists a significant space of development for the model, although its forecasting accuracy in the present version is already is already comparable or even higher than model embedding significantly more intelligence.*

3. **The SPS provides an accuracy comparable to IMF forecasts, while being characterized by a much lower model complexity.** Here we compare our results with the available long-term GDP growth projections available through the World Economic Outlook by the IMF with three different accuracy metrics. SPS is a unique parameter-free model, while IMF models have specifications and assumptions tailored to single nations.

4. **The SPS allows definition of a novel economic dimension which measures the *predicability* of economic growth** (and of development at large). This dimension is shown to be applicable to IMF projections as well, defining a country segmentation for which the predictive accuracy of both IMF and SPS projections is a-posteriori shown to be significantly improved.

The implications for economic practice are threefold: **i)** the Fitness dimension and New Development Economics (i.e. Economic Complexity) define a model which couples simplicity and parsimony in terms of assumptions and whose performances are comparable with the state of the art on the same time horizon (5 years); **ii)** the SPS provides growth estimates on time horizons even longer than the state of the art and keeps a comparable level of reliability of the one achieved in a 5-year time horizon; **iii)** the framework offers a scientific foundation for selecting the proper past comparatives to provide a forecast of the future and simultaneously for assessing the expected accuracy of such forecast. This *predictability* assessment provides a criterion to assess the confidence of growth estimates that mostly depends on the status of the country itself, as identified by its Fitness and per capita GDP (GDP_{pc}).

2. Measuring country competitiveness: The Fitness of countries

Fitness synthetically measures differences in country competitiveness from cross-country output differences, specifically from differences in terms of number and complexity (or sophistication) of classes of products.

The complex and heterogeneous structure of the network of interactions between economic actors and economic activities/products plays a special role in order to underpin the main features of economic outputs (even at aggregate level such as the Gross Domestic Product, GDP). The features of this country specific network are the results of latent economic dimensions we have to deal with in order to segment the past in relation to our final aim to find the appropriate comparators to model the future. This is traditionally rephrased and illustrated by stating that the economic output can be seen as the result of country specific multipartite (for the sake of simplicity, tripartite) networks where the three composing layers are the economic actors (here countries), the capabilities and the economic activities (or products) as shown in Fig. 1. This picture clarifies the special role played by the network due to economic actors and economic activities/products as it turns out to be the bipartite projection of this tripartite structure which encodes information we want to access.

It also follows that the mathematical specification defining the measure of country competitiveness becomes a crucial point, far from being a simple exemplification of the narrative. They must conversely be driven by the features of the structure of the bipartite structure of actors and activities which, at country scale, becomes the bipartite network of

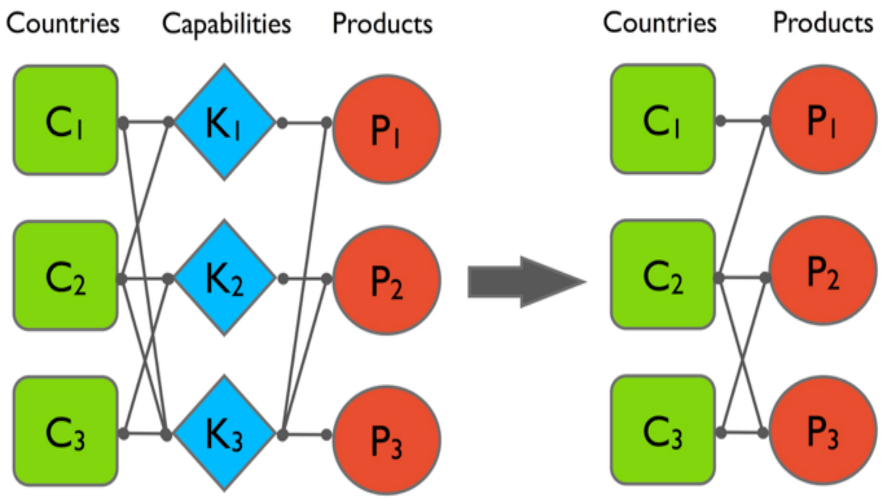

Fig. 1: Cross-country endowment (blue diamonds) differences define the relative competitiveness of countries. Cross-country output differences in terms of number and nature of products synthetically encode this relative competitiveness because the bipartite network defined by countries (green squares) and economic output and activities – in this specific example, the products (red circles) – is the projection of the tripartite network countries-endowments-products. This tripartite modeling is intractable via a direct approach being the result of heterogeneous economic dimensions and hardly measurable interactions. Conversely bilateral trade data are a robust proxy for the bipartite network country-product.

countries-products. Conceptually speaking, the measure of intangible features here discussed is similar to the effort of a class of methods in network science[9] which aims at inferring features of the network nodes starting from the node connections (i.e. from the topology). However, the analogy holds only at a conceptual level as the bipartite nature of the export network and the features of this network itself call for specifications which are not simply an extension of a linear ranking algorithm to the bipartite case as proposed in [Hidalgo et al., 2009].

We follow the commonly accepted procedure to build a proxy for the bipartite structure of countries-products starting from country trade flows, i.e. export [Hidalgo et al, 2009]. We refer to Appendix B.1 for the motivations of this choice.

As mentioned, we are interested in qualitative differences (i.e. what) rather than quantitative (i.e. volumes) differences. This implies the definition of a filter to make the raw export volumes binary quantities. We again leverage a standard and widely accepted indicator to perform this task: the *Revealed Competitive Advantage* (RCA) index proposed originally by Balassa in [Balassa, 1965]. We end up with a binary matrix whose entries M_{cp} are 1 if the country c has a revealed comparative advantage in product p larger than 1, they are 0 otherwise (see Appendix B.1 for the specifications of the RCA matrix and further details about the RCA interpretation).

The visualization of the matrix M (see Fig. 2) reveals a clear statistical feature of nestedness. Adapting the ecological definition of a nested system to economics, an economic system is said to be nested when specialist (non-diversified) countries tend to produce a subset of products which is also made by generalist (diversified) countries; on the other hand only generalist (ubiquitous) products are produced by specialist countries. Nestedness is thought to be one of the principal signatures of complex ecosystems, where actors compete for finite resources.[10] The nested structure sets strong constraints on the mathematical specifications of the algorithm which intends to measure country competitiveness.

2.1 Mathematical specifications of the Fitness dimension

The nested structure of the country-product bipartite network as shown in Fig. 2 sets non-trivial constraints on the amount of information carried by the network edges in order to assess country competitiveness. Let us consider four specific cases to illustrate why different links of the country-product network carry different information:

i. a product is exported by a largely diversified country;
ii. a product is exported by a poorly diversified country;
iii. a country exports a highly nonexclusive product;
iv. a country exports a highly exclusive product.

It turns out that statements i) and iii) carry little information in order to determine the level of sophistication (complexity hereinafter) of a product and the country competitiveness (Fitness) respectively. Diversified countries (e.g. Germany) being competitive on a wide

---

[9] The eigenvector centrality measure and following evolution, the PageRank [Page et al, 1999] is likely the most known in this class of approaches.

[10] In support of this, similar nested structures are also observed in ecological systems. Ecological and economic systems have strong analogies as they are both adaptive, evolutionary systems with limited resources. We refer to [Dominguer-Garcia et al, 2015] where the Fitness and Complexity algorithm has been shown to be insightful to assess and rank the importance of species in terms of the ecosystem stability.

range of products carries little information on their complexity. The limit case of a country exporting all products (even though this limit cannot practically occur due to RCA) would not provide information at all on product complexity. Similarly, statement iii) provides little information to assess the country fitness as non-exclusive products tend to be produced by all countries. Again, if all countries exported a product (even though RCA prevents to observe this limit case), this product would not convey information at all about the Fitness of countries.

Conversely both statements ii) and iv) have a leading role in assessing product complexity and country fitness respectively. Statement ii) is conveying crucial information that that product is produced in a country which is able to be competitive only on non-exclusive products and statement iv) underpins a country able to export a product which only very few diversified countries are able to compete on.

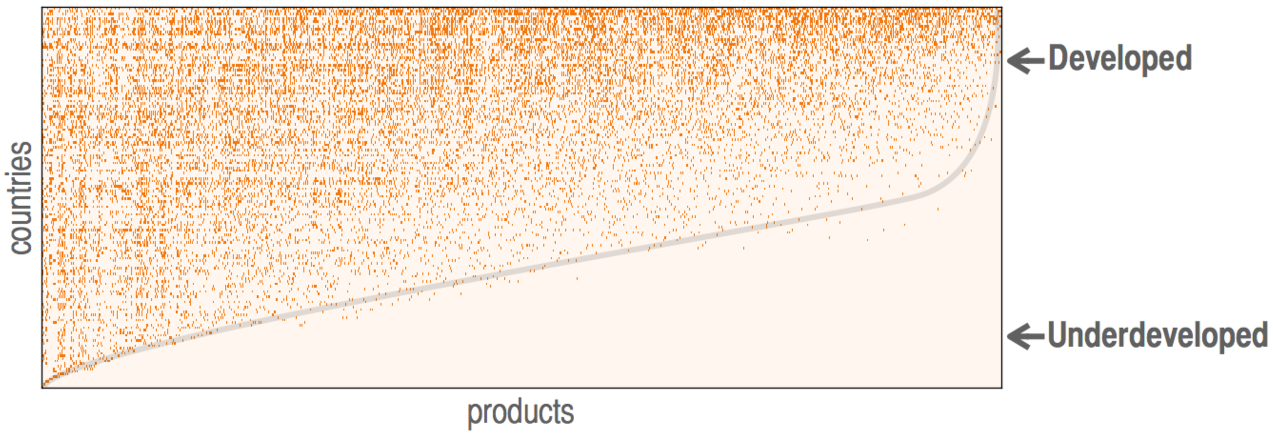

Fig. 2: The nested structure of the binary bipartite network defined by countries and exported products. Rows represent country export baskets and columns instead specify products. A dark orange dot means that $M_{cp}=1$ while light orange means $M_{cp}=0$. Rows and columns are rearranged according to Fitness and Complexity dimensions defined in Section 2.2. A system is said to be nested when specialized actors (i.e. countries) tend to produce only a set of products which is also made by generalist or diversified countries. Source [Cristelli et al., 2013].

To sum up, the nested structure of the bipartite network countries-products, on one hand, calls for a leading role of the diversity of a country production to underpin its competitiveness. On the other hand, nestedness also implies that the complexity of a product must be non-linearly related to this competitiveness dimension (the Fitness) of countries and specifically this relation must be dominated by the less fit exporter. In other words, the Complexity cannot be a simple average of the Fitness of its exporters.

By combining these two arguments, we obtain the algorithm firstly proposed in [Tacchella et al., 2012] which consists in self-consistent non-linear coupled equations for the Fitness of countries - the competitiveness dimension for countries we were looking for - and the Complexity of products.[11] In Eq. 1 we report the specifications of this iterative scheme.

[11] *We refer to [Cristelli et al., 2013] for a detailed description of the differences of these specifications with previous attempts.*

$$
\text{Iteration } (n) = \begin{cases} \tilde{F}_c^{(n)} = \sum_p M_{cp} Q_p^{(n-1)} & F_c^{(n)} = \dfrac{\tilde{F}_c^{(n)}}{\langle \tilde{F}_c^{(n)} \rangle} \\ \\ \tilde{Q}_p^{(n-1)} = \frac{1}{\sum_c \frac{M_{cp}}{F_c^{(n-2)}}} & Q_p^{(n-1)} = \dfrac{\tilde{Q}_p^{(n-1)}}{\langle \tilde{Q}_p^{(n-1)} \rangle} \end{cases} \tag{1}
$$

The fixed point of this map[12] operatively defines the measure for country competitiveness (i.e., the Fitness dimension) and the product Complexity. In Eq. 1, $F_c^{(n)}$ and $Q_p^{(n)}$ are respectively the Fitness of a country and the Complexity of a product at *n-th* iteration of the algorithm and M_{cp} are the entries of the previously defined binary matrix *M.* The symbol $\langle \cdot \rangle$ denotes the average, in the case of $F_c^{(n)}$, over all Fitness values and, in the case of $Q_p^{(n)}$ over all complexities values. We remind the reader that the binary $C \times P$ matrix *M* defines the topology of the bipartite network whose nodes are countries and exported products (*C* and *P* are, respectively, the number of countries and the number of products).

It is worth noticing that, at each step of the algorithm, the Fitness is the diversity of a country weighted by the Complexity of products while the Complexity of products is the harmonic mean of the Fitness of the countries exporting that product up to the normalization factor. On one hand this implies, consistently with the nested structure of the matrix *M,* that the lowest Fitness value among those countries exporting a product is an upper bound for the Complexity of a product. On the other hand, it also means that the larger the number of countries exporting a product, the lower will be the Complexity.

In [Cristelli et al., 2013; Pugliese et al., 2014; Wu et al., 2016] the reader can find extended discussion on the convergence properties of Eq. 1.

2.3 Data sets

The International Trade Statistics Database is made publicly available by United Nations and accessible via UN Comtrade website.[13] In this specific work we will use a dataset derived from UN Comtrade raw data: the BACI data set released by CEPII. The BACI dataset is the result of reconciliation procedure of the UN Comtrade which essentially fixes the inconsistencies between import and export flows, see [Gaulier et al., 2010] for details.

The latest release of the BACI data set spans the period 1995-2014. This data set provides all yearly trade volumes, expressed in current USD, between pairs of countries broken down at the product level. Products are available up to the 6-digit level of the Harmonized System classification. In this specific work, we will use products aggregated

[12] *The fixed point of a mathematical map f is simply an element of the domain of the function which is mapped into itself. In formula given the function y=f(x), x=p is said to be a fixed point of f if and only if f(p)=p. As a concrete example, economic equilibria are usually fixed points. In this specific case, the fixed point of the Fitness and Complexity map is said to be an attractive fixed point as there is a non-empty subset of the domain of the map for which the iterated sequence x, f(x), f(f(x)), f(f(f(x))), … converges to the fixed point. Furthermore, for this specific map we numerically show that the attractive fixed point is also unique in the subdomain of the map which is economically meaningful (see [Cristelli et al, 2013]).*

[13] *https://comtrade.un.org*

at 4-digit level (HS2007). At 4-digit level, there are approximately 1,150 products exported by at least one country (we discard those products for which all trade flows are zero).

Eventually we filter out countries according to a population and total export volume threshold. This procedure selects approximately 140 countries which account for more than 95% of world GDP.

The source of all the remaining economic dimensions used in this work is World Bank Open Data platform,[14] except when explicitly stated otherwise.

[14] *http://data.worldbank.org*

3 Results: Testing SPS performances and measuring growth predictability

3.1 SPS model in a nutshell

As mentioned in the previous sections, SPS is a scheme providing a set of criteria to select examples from the past to model future growth estimates and in particular, it provides an operative answer to which past should be used to forecast which future.

The evolution of a country's fitness over time defines a trajectory in the fitness-GDP plane. Let's call this the development trajectory. The challenge is to then define the space such that similar trajectories represent comparable development events, which form the basis of similar development futures. This is the role of the SPS.

Country economic states are defined in the Fitness-GDPc plane by *2x1* vectors representing the logarithm of the Fitness and the logarithm of the GDPpc PPP. We the define the set of comparators for an economic state we want to forecast *D*-period ahead as the set defined by the countries which were in the past in the neighborhood of this economic state in the Fitness-GDPpc plane (we refer again to Section 5 for all the details and conditions to filter out comparators in order to avoid biases or country over representations). Provided this set, we average the *D*-period ahead trajectories of the comparators, the starting point is the first year a comparator country is found in the neighborhood of the event to forecast. Let us make an example, we want to forecast the growth of Vietnam from 2014 to 2019 and, for instance, we find that Mexico in 1995 was in the neighborhood of Vietnam in 2014 and Indonesia in 2005. Mexico's trajectory from 1995-2000 and Indonesia's one from 2005 to 2010 will be averaged in order to model the evolution of Vietnam in the period 2014-2019. The comparator set yields an average displacement in the Fitness-GDPpc plane which represents our projections for the event we want to forecast. This implies that SPS provides both a forecast for the GDPpc and the Fitness since SPS provides an assessment of the *D*-period ahead position in the Fitness-GDPpc of the country to forecast. For the sake of comparison with other source of GDPpc projection, we convert this result into an annualized growth rate.

3.2 Overview of IMF's growth modeling

In order to stress the huge parsimony of the SPS in terms of parameters, variables and assumptions, we provide a brief overview of the modeling underlying IMF's growth projections as conceptual benchmark for the remainder of this section. The IMF publishes GDP growth projections in the annual World Economic Outlook. There is no global unified methodology: the projections are computed country by country and are subsequently homogenized and aggregated through a multi-step process with feedbacks among different teams, schematized in Fig. 3 Reading from the IMF website[15]:

"*The IMF's World Economic Outlook uses a "bottom-up" approach in producing its forecasts; that is, country teams within the IMF generate projections for individual countries. These are then aggregated, and through a series of iterations where the aggregates feed back into individual countries' forecasts, forecasts converge to the projections reported in the WEO.*

Because forecasts are made by the individual country teams, the methodology can vary from country to country and series to series depending on many factors."

15 https://www.imf.org/external/pubs/ft/weo/faq.htm#q1g

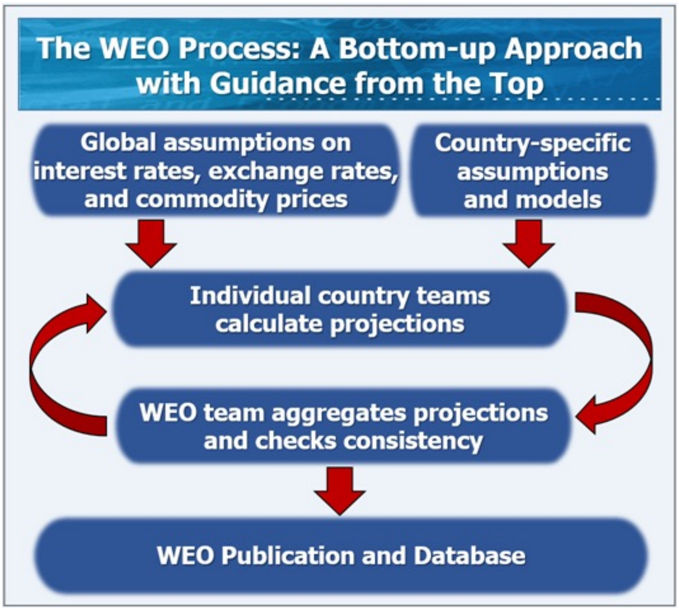

Fig. 3: A scheme representing broadly the IMF methodology for GDP forecasting. (Source: IMF website)

The IMF projections are based on a number of precise assumptions, as stated in [IMF, 2016]. Those assumptions range from exchange rates, to a precise estimate of the oil prices in 2016 and 2017, to interbank rates for several different currencies. On top of these global assumptions there are country-specific assumptions on the continuity of national fiscal and monetary policies.

This results in a very hard to grasp global picture of the models used by the IMF. It is not directly possible to estimate the number of parameters and assumptions used for the forecasting as well as the aggregation procedure since they are inherently country specific.

3.3 Other benchmark models

We also compare our predictive power against three simple country specific univariate time series models. Simple zero/limited intelligence models have been shown to have a non-negligible predictive power in the long run (see for instance [Pritchett et al, 2014] and [Kraay et al., 1999]) and they represent an important benchmark.

Model 1: we use, as a predictor of the annualized *D*-period ahead growth rate, the annualized growth rate observed in the last *D*-period, in formula:

$$g_{c,t,t+D} = g_{c,t-D,t}$$

where $g_{c,t,s}$ denotes the annualized growth rate of the GDPpc of country *c* in the period of time from *t* to *s.*

Model 2: we model the evolution process for the logarithm of GDPpc of country *c* at time *t* (denoted with $y_{c,t}$) as a country specific AR(1) process:

$$y_{c,t} = \gamma_c + \rho_c y_{c,t-1}$$

Model 3: as for model 2 we use a country specific AR(1) process for the evolution process of the logarithm of GDPpc of country *c* at time *t* (denoted with $y_{c,t}$) and we include a trend term linearly related to the time:

$$y_{c,t} = \gamma_c + \delta_c t + \rho_c y_{c,t-1}$$

Model 1 represents the simplest specifications for the argument proposed in [Pritchett et al, 2014] while model 2 and 3 are the simplest specifications for the country specific models proposed in [Kraay et al., 1999].[16]

3.4 Comparison of SPS with IMF's projections and other benchmark models

In order to assess the accuracy of the growth projections we compare the annualized forecasted growth rates with the corresponding realized annualized growth rates over the same time window. Hereinafter we will refer to annualized growth rates as simply growth rates for the sake of simplicity. We benchmark SPS and IMF projections accuracy using three metrics:

- **Pearson Correlation (PC)** coefficient between projected growth and actual growth rates;
- **Mean Absolute Error (MAE):** the average of absolute errors, where the error is defined as the difference between forecasted rates g_i^{proj} and actual growth rates g_i^{actu}, in formula

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|g_i^{proj} - g_i^{actu}\right|$$

- **Root Mean Squared Error (RMSE):** the square root of the average squared errors, where the error is defined as the difference between forecasted rates g_i^{proj} and actual growth rates g_i^{actu}, in formula

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}\left(g_i^{proj} - g_i^{actu}\right)^2\right]^{\frac{1}{2}}$$

While the last two metrics are similar, the RMSE gives a greater importance to large errors. In this way we can also detect the nature of the errors and distinguish scenarios which are on average similar but show differences in terms of the behavior of the outliers of the event distribution. In the strict sense, correlation is not a performance indicator. It rather provides directional information of the error dispersion, we might have an almost vanishing MAE and simultaneously zero correlation or perfect correlation and a non-zero MAE.

IMF projections are provided in the World Economic Outlook on a yearly basis and historical data are publicly available on WEO website.[17] Long term projections (4-5 years) are available only after 2007. We therefore test IMF performances on three time windows 2008-2013, 2009-2014, 2010-2015.

---

[16] *In [Hidalgo et al., 2009] and in [Hausmann et al, 2014] the authors relate complexity index ECI to growth and propose a regressive model which, iterated forward, can be leveraged to forecast growth. In the web page http://atlas.cid.harvard.edu/rankings/growth-predictions/ the authors release their 10 years forecast from 2004 to 2014. This implies that a limited comparison would be, in principle possible, for the period 2004-2014 and 2005-2015 but unfortunately they do not specify whether the forecasted growth refers to nominal or real GDP (or GDPpc). We cannot therefore determine which is the proper actual growth to benchmark the projections based on this model. For these reasons we decided not to include these projections in our comparison.*
In addition, a direct comparison with the present scheme cannot be performed as the dynamics in the ECI-GDPpc plane is observed essentially chaotic everywhere.

[17] *http://www.imf.org/external/ns/cs.aspx?id=28*

SPS estimates are available over six periods: 2005-2010, 2006-2011, 2007-2012, 2008-2013, 2009-2014, 2010-2015. In both cases, we compare the growth rate projections of GDPpc PPP.
For the model 2 and model 3, we estimate the parameters using the available data up to time *t-1* where *t* represents the starting time of the period we want to forecast. We measure the performance over the same six periods used for SPS. In order to have a setting as similar as SPS, we train models 2 and 3 using data from 1995 only.

Let us first consider a setting in which all available countries are considered ('All'). In Table 1 we show the results of the comparison for the accuracy of SPS and IMF growth projections on a 5 year-time horizon in terms of correlation.[18] In all the tables two specifications for SPS are provided. The main difference between these two specifications is that 'SPS + trend', differently from 'SPS', explicitly accounts for an autoregressive term for past growth trends.[19] In section 5 we discuss the motivation of the two specifications and in next section we show that they are substantially equivalent in the predictable regime reinforcing the validity of the selected comparators. In terms of directional information, we find that the IMF performs approximately 10% better than SPS.

	'All'		
	SPS	**SPS + trend**	**IMF**
PC (p-value)	0.30 (<0.00001)	0.37 (<0.00001)	0.42 (<0.00001)
N. Obs.	758	758	384

Table 1: Comparison of SPS (two specifications) and IMF performances with respect to Pearson correlation. We color in green the scheme with higher correlation.

Conversely, as shown in Table 2, we observe a reversed scenario when we consider the accuracy of the projections. Both SPS specifications outperform IMF in terms of MAE and RMSE. Moreover, the difference of the typical error size between IMF and SPS is higher, in relative terms, for RMSE than MAE (SPS is up to 15% and 3% more accurate than IMF on the basis of RMSE and MAE respectively). This has a twofold implication: i) both SPS specifications are more accurate in terms of magnitude of errors and ii) they make smaller extreme errors, i.e. the dispersion of the forecast error distribution is smaller for SPS than for IMF.

It is worth noticing that, although Fitness is an export and manufacturing driven dimension, the statistical features of SPS' errors are marginally dependent on GDP sector composition. For instance, SPS approximately performs on average the same error for countries driven by manufacturing or by services and the same finding applies as a function of the GDP share due to exports. We refer to Appendix A.2 for extensive discussions on this point.

---

[18] *For each row of the provided tables we highlight in green the method with the best performance. For MAE and RMSE the lower the value, the higher is the forecasting power. This means that MAE=0 (RMSE=0) correspond to an error-free forecast. For correlation (PC) the information provided is only directional and higher or slower correlation does not imply necessarily higher or slower accuracy in terms of projection errors.*

[19] *In this way, we can deal with the known empirical facts that past growth rates, explain approximately 10% of future growth rates variance [Pritchett et al, 2014].*

	'All'					
	SPS	**SPS + trend**	**IMF**	**Model 1**	**Model 2**	**Model 3**
MAE (CI)	2.01 (1.91,2.11)	1.97 (1.87,2.07)	2.03 (1.85,2.22)	2.71 (2.52,2.89)	4.62 (4.26,5.03)	5.42 (5.02,5.87)
Accuracy gain % (ref. IMF)	**1.0%**	**3.0%**	0.0%	**-33%**	**-129%**	**-167%**
RMSE (CI)	2.64 (2.49,2.80)	2.60 (2.45,2.75)	3.05 (2.64,3.48)	3.79 (3.50,4.09)	7.07 (6.20,8.03)	8.10 (6.87,9.45)
Accuracy gain % (ref. IMF)	**13.4%**	**14.8%**	0.0%	**-26%**	**-131%**	**-166%**
N. Obs.	758	758	384	763	760	760

Table 2: Comparison of SPS (two specifications), IMF, Model 1-2-3 performances with respect to MAE and RMSE. The confidence interval CI is estimated via bootstrapping and values reported correspond to 95% CL. We color in green the scheme with the highest accuracy.

In table 2 we also show the predictive power for Model 1-2-3 and, although Model 1 performs surprisingly well, their accuracy is significantly lower than SPS and IMF projections. We can conclude that SPS and IMF provide a significant improvement of the accuracy when compared with county specific limited intelligence models.

In summary, SPS and IMF behave overall similarly on the basis of the three metrics but:

- SPS uses less heterogeneous data than IMF and is more parsimonious in terms of parameters and still achieves similar performance
- SPS performs slightly better than IMF in both RMSE and MAE terms.

In next section we discuss how SPS can extract a further dimension, the Predictability, and how it can be used to refine the IMF forecast.

3.5 The predictability of growth

The dynamics in the Fitness-GDPpc allows defining a dimension, the *Predictability* hereinafter, which essentially measures how regular the flow of the economic evolution is (see Section 5 for the mathematical specifications). As shown in Figs. 6, 8 and 9, the two regimes (P and UnP) emerge. Dividing the countries on the basis of this variable illustrates if there are systematic differences of the SPS performances on the two subsets. As a first test, we say that a country belongs to the *predictable* regime if it has $ln(f) > -1$ and to the *unpredictable* regime otherwise (see Section 5 for a discussion of this threshold). As shown from the first two rows of the last column of Table 3 and 4, approximately 60% of the original events fall into the *predictable* regime and the remaining 40% in the *unpredictable* regime.

3.5.1 Segmenting countries on the basis of Predictability: SPS cases

Let us first consider the two SPS specifications in order to discuss whether the Predictability defines a non-trivial segmentation of countries. If Predictability dimension is informative on the degree of growth predictability, we should observe that accuracy of 'P' regime is higher than the one of the overall case ('All') and that the 'All' case has a higher accuracy than the one measured for countries belonging to the 'UnP' regime.

	'All'		'P' – Laminar/Predictable		'UnP' – Chaotic/Unpredictable	
	SPS	**SPS + trend**	**SPS**	**SPS + trend**	**SPS**	**SPS + trend**
PC (p-value)	0.30 (<0.00001)	0.37 (<0.00001)	0.41 (<0.00001)	0.38 (<0.00001)	0.21 (0.00012)	0.42 (<0.00001)
N. Obs.	758	758	429	429	329	329

Table 3: Comparison of SPS (two specifications) with respect to Pearson correlation for countries belonging to the *Predictable* regime (*ln(Fitness)* > *-1*) and to the *Unpredictable* regime (*ln(Fitness)* < *-1*). We report in columns 'All' the results of Table 1 for the sake of comparison.

In Table 3 we report SPS comparison in terms of correlation. For SPS without trend specifications the expected ordering of the correlation magnitude is observed: 'P' is more correlated than 'All' case and 'All' case is more correlated than 'UnP' regime. We observe a 35% correlation gain for SPS without trend (see Section 5). In the case SPS + trend, the correlation is instead substantially left unchanged (as we will see in Table 4 the same behavior is observed for MAE and RMSE: the relative accuracy gain is higher for SPS than for SPS + trend). We discuss in Section 3.5.3 the conceptual implications for the nature of the comparative events of this different behavior.

	'All'		'P' – Laminar/Predictable		'UnP' – Chaotic/Unpredictable	
	SPS	**SPS + trend**	**SPS**	**SPS + trend**	**SPS**	**SPS + trend**
MAE (CI)	2.01 (1.91,2.11)	1.97 (1.87,2.07)	1.81 (1.69,1.93)	1.94 (1.81,2.06)	2.26 (2.09,2.45)	2.08 (1.92,2.25)
Accuracy gain % (ref. 'All')	0.0%	0.0%	**10.0%**	**1.5%**	**-12.4%**	**-5.6%**
RMSE (CI)	2.64 (2.49,2.80)	2.60 (2.45,2.75)	2.33 (2.19,2.58)	2.51 (2.35,2.67)	3.00 (2.72,3.28)	2.75 (2.50,3.01)
Accuracy gain % (ref. 'All')	0.0%	0.0%	**11.7%**	**3.5%**	**-13.6%**	**-5.8%**
N. Obs.	758	758	429	429	329	329

Table 4: Comparison of SPS (two specifications) performances with respect to MAE and RMSE. The confidence interval CI is estimated via bootstrapping and values reported corresponds to the interval defining 95% CL. We color in green the scheme with the highest accuracy. We report in columns 'All' the results of Table 2 for the sake of comparison.

In Table 4 we show for both SPS specifications the comparison in terms of forecasting accuracy. Filtering countries on the basis of the splitting 'P' and 'UnP' yields a significant improvement of SPS forecasting accuracy when we consider MAE and RMSE metrics. In detail, the accuracy for predictable countries 'P' is larger than accuracy for 'P' and 'UnP' countries combined together (i.e. 'All' columns): from 1.5% to 10% more accurate for MAE and from 3.5% to about 12% for RMSE. 'All' case yields results which are from 5.6% to 12.4% more accurate than 'UnP' countries for MAE and from 5.8% to 13.6% for RMSE. We therefore find that 'P' accuracy is always higher than 'All' case accuracy and the accuracy of 'UnP' is always lower than the accuracy of the 'All' case.

The direct comparison of 'UnP' and 'P' shows that 'P' regime is up to 20% more accurate than 'UnP' regime for MAE and up to 22% for RMSE.

3.5.2 Why predictability matters for IMF projections

Let us replicate the analysis of the previous section for IMF projections in order to discuss how the predictability dimension can be leveraged to reduce forecasting errors and enrich the information provided by IMF projections. As shown in Table 5 correlation is left unchanged by the segmentation of countries as in the scenario SPS + trend.

	'All'	'P' – Laminar/Predictable	'UnP' – Chaotic/Unpredictable
	IMF	**IMF**	**IMF**
PC (p-value)	0.42 (<0.00001)	0.42 (<0.00001)	0.41 (<0.00001)
N. Obs.	384	213	171

Table 5: Comparison of IMF with respect to Pearson correlation for countries belonging to the *Predictable* regime (*ln(Fitness) > -1*) and to the *Unpredictable* regime (*ln(Fitness) < -1*). We report in columns 'All' the results of Table 1 for the sake of comparison.

In terms of accuracy, the 'P'/'UnP' country segmentation surprisingly delivers results which are similar to the scenario we observe for SPS. IMF projections restricted on 'P' regime are more accurate than the case considering all countries and the 'All' case is more accurate than the 'UnP' regime. 'P' regime for IMF is 19% and 33% more accurate than 'UnP' in terms of MAE and RMSE respectively.

	'All'	'P' – Laminar/Predictable	'UnP' – Chaotic/Unpredictable
	IMF	**IMF**	**IMF**
MAE (CI)	2.03 (1.85,2.22)	1.87 (1.70,2.06)	2.22 (1.87,2.58)
Accuracy gain % (ref. 'All')	0.0%	**7.9%**	**-9.4%**
RMSE (CI)	3.05 (2.64,3.48)	2.46 (2.22,2.68)	3.66 (2.88,4.37)
Accuracy gain % (ref. 'All')	0.0%	**19.3%**	**-20%**
N. Obs.	384	213	171

Table 6: Comparison of IMF performances with respect to MAE and RMSE. The confidence interval CI is estimated via bootstrapping and values reported corresponds to the interval defining 95% CL. We report in columns 'All' the results of Table 2 for the sake of comparison.

Accuracy increase is then observed not only for SPS but also for IMF. While for SPS this observation is coherent with the definition of Predictability, this is *a priori* unexpected for IMF. Even more surprisingly, the performance gain is, in relative terms, even higher for IMF than for SPS (both cases) in terms of RMSE (IMF gain is about 20% while SPS gain is at most about 12%). Filtering countries on the basis of Predictability therefore reduces the average size of IMF projection errors and it significantly reduces the number of projections producing large errors. This means that, even if IMF growth modeling is completely different, the *Predictability* dimension is capturing a real feature of the countries (and consequently of the economic regime) pooled in those areas of the plane where the *Predictability* is found to be high.

As a general comment, while SPS and IMF perform similarly when all available countries are considered (IMF outperforms SPS in terms of correlation but SPS outperforms IMF in terms of MAE and RSME), on this reduced set of *predictable* countries SPS (the specification without the trend component) matches or outperforms IMF accuracy in all cases.

In Appendix A.1 we provide a table summarizing all the results provided in Table 1-6. In Appendix A.5 we show the results of Table 6 for Model 1-2-3. For the three univariate models, we observe much more limited relative differences between the three scenarios 'P', 'UnP' and 'All countries'. This finding is unsurprising because the three univariate models leverage the past of the country itself which might be very far – i.e. in a different regime - in the fitness-GDPpc plane from the trajectory we want to forecast. For these reasons, we do not expect that the 'P' and 'UnP' classification should affect the forecasting performance as for SPS or IMF.

3.5.3 Conceptual implications of the Predictability for SPS comparators and IMF projections

Let us now discuss the different behavior of SPS and SPS + trend when we restrict our analysis to the regime 'P' and the implications of this on the nature of the comparators SPS select. For the sake of clarity of this section, the difference between the two specifications consists in the fact that SPS does not directly account for an autoregressive component in the growth estimates (i.e. in the set of past comparators of a country we exclude those provided by the country itself) while 'SPS + trend' specifications explicitly account for an autoregressive term for past growth trends as discussed in Section 5 where we provide further details for both specifications.

By comparing SPS and SPS + trends performance gains as in Tables 3 and 4, we find that filtering countries by predictability improves in relative terms SPS accuracy more than 'SPS + trend' one and makes SPS outperforming 'SPS with trend' and consequently reverses the ranking observed in Table 1 and 2. By coupling these observations with the specifications' differences, we can conclude that in the *predictable* regime, the past trend component is very similar for all comparatives. They are not only good comparatives for the future trajectory but they are also good comparators in terms of past behavior. In other words, by selecting predictable countries we are selecting countries in a regime where there exists a typical future trajectory of evolution, a typical past growth trend and the comparators selected by SPS provide good estimates of both aspects.

Conversely in the other regime and in the case of with available countries, although the closeness in the plane is still informative, this closeness is only partial and the trend components measured by the autoregressive term of country past growth is not represented as well as in the *predictable regime* by comparatives' dynamics. For this reason, in the 'UnP' regime the growth trend complements much better the simple SPS forecasting, providing a greater improvement to all the metrics considered with respect to the 'P' case.

The evidence that for countries in the regime 'P' the closeness is concretely pinpointing a closeness of economic states, points in the direction of underlying typical trajectories of development in this regime and the fitness-GDPpc plane and SPS are revealing those patterns. This perspective then explains why filtering countries by predictability also increases IMF performances. SPS is signaling that trajectories of growth exist, are less heterogeneous than in 'UnP'. Therefore the *signal-to-noise* ratio is more favorable for all modeling including IMF ones and not only for SPS.

3.6 SPS accuracy beyond 5 years

'Growth is devilishly hard to predict' is the opening of a 2016 editorial that appeared in the Economist's columns [The Economist, 2016] and the scarcity of systematic and cross-country sources of long term growth projections witnesses the inherent challenges of this task. SPS intends to lay the foundation of a playground to provide growth insights beyond the 5-year time horizon. The possibility within SPS to *naturally* discuss longer time horizon crucially relies on the model's parameter parsimony and on the extreme homogeneity of the leveraged sources of data: a proxy of the production structure. This permits to avoid entering a number of macroeconomic estimates underlying more traditional growth modeling scheme (e.g. IMF). These estimates are crucial to capture short term variability of growth but become an intrinsic limit for increasing time horizon: the more there are parameters/variables to estimate, the higher the chance of errors and chains of error for composite estimates. Conversely if there are stages or regimes of growth where countries undergo (on average) recurrent low dimensional patterns of growth on the long run - and our analysis does support the existence of these typical trajectories - the knowledge of these trajectories becomes increasingly the driving strategy to set up modeling of future growth variability.

It is not by chance that the SPS playground is defined in a logarithmic space which is coherent with the idea of smoothing out country specific growth profiles in order to first map these stage specific long-term trends. Even though this perspective sound as alternative to standard modeling, this narrative is instead synergic with standard indicators and with the body of knowledge concerning feedbacks between macroeconomic dimensions and growth. SPS sets a *hierarchical* framework to discuss growth modeling on the long term by disentangling intrinsic growth trajectories due to the development stage of a country and country specific growth profile originated by cross section differences of their macroeconomic dimensions. Cross section differences are here intended at the same point of closeness in the fitness-GDPpc plane and not at the same point of time as they are usually intended.

Let us discuss SPS results for time windows longer than 5 years. We limit our analysis in this section to the simplest specifications of SPS model without considering the SPS+trend specifications.[20] In order to ensure as much as possible a *ceteris paribus* comparison among different time horizons, we restrict the performance analysis of SPS on those projections with starting year in the time interval 2000-2004. The projections of a country for a specific starting year are retained only if they are available for all time horizons we consider, otherwise we discard it.[21] This implies that all the performance indicators we compute in the following are assessed on the same number of observations.

Panels of Fig. 4 and Fig. 5 illustrate SPS projections' performances in terms of PC and MAE, RMSE respectively as a function of the time horizon *D.* All metrics suggest that SPS performances tend to increase (diminishing MAE and RSME) for increasing time windows. As an example, SPS on a 9-year time horizon is about 25% more accurate than SPS with 5-year specifications in terms of MAE and approximately 30% more accurate in

[20] The main constraints of extending SPS on longer time horizon are set by a decreasing statistics of comparatives to train SPS. To model growth projections on a 5 years time horizon we need comparative trajectories of 5 years length and therefore we can select comparators with a starting point up to 2010, for 7 years horizon up to 2008 and for 10 years only up to 2005.

[21] we discuss in Appendix A.4 the performance of SPS as a function of the parameter defining the size of the neighborhood of an event in the Fitness-GDPpc plane in order to select its comparators. In this subsection the value for this parameter is the same we use throughout Section 3.

terms of RMSE. Similar patterns are observed restricting the analysis to countries in the 'P' regime, we simply have downward shifted patterns for MAE and RMSE and upward for PC (there are 329 observations in this case).

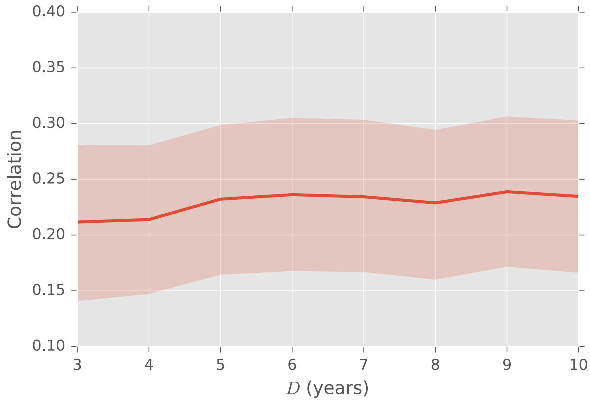

Fig. 4: Correlation between projected and actual growth rates for projection with initial point in the interval 2000-2004 as a function of the forecasting time horizon expressed in years. CI is estimated via bootstrapping. All points are estimated with 622 observations.

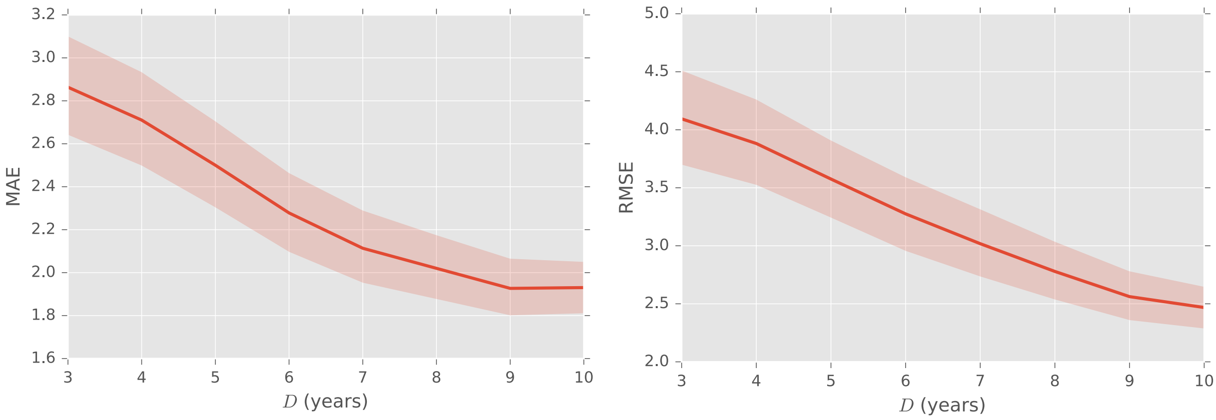

Fig. 5: MAE (left panel) and RMSE (right panel) of SPS projected growth rates for projection with initial point in the interval 2000-2004 as a function of the forecasting time horizon expressed in years. CI is estimated via bootstrapping. All points are estimated with 622 observations.

3.7 Summary of SPS key findings

To sum up, this section provides several key findings:

- The SPS allows to introduce the concept of heterogeneity of the degree of predictability of an economic system. Countries are economic systems whose predictability depends on the development stage itself. This finding is also consistent with the observation that the probability of occurrences of wars, conflicts, and all major sources of economic volatility are dependent on the development maturity of a country.

- The concept of predictability naturally arises from this scheme, and it can be used to assess the confidence of growth projections, regardless of the modeling used, as in the case of IMF. This dimension is then used to enhance the value carried by those projections by segmenting countries on the basis of the estimate reliability.
- The closeness of comparators in the predictable regime points in the direction of concrete closeness of economic states, in other words, in the direction of the existence of typical and less heterogeneous trajectories of development with respect to other regimes. It means that a predictable regime selects countries which offer more favorable conditions to modeling, including the IMF one.
- SPS modeling is more parsimonious than IMF in terms of parameters and leverages less heterogeneous economic indicators. SPS and IMF perform in a comparable way when all countries are considered and SPS outperforms IMF in the predictable regime.

4. Case Scenarios

This section illustrates SPS results for 2019 for Thailand and Ghana, two countries which fall into the predictable and unpredictable regimes respectively. Considering these cases shows the importance of the Predictability dimension as a tool to estimate the forecasting power of growth models such as SPS.

As shown in section 3.5.1, SPS forecasts are very accurate for countries like Thailand, where growth dynamics follow predictable pathways laid out by comparator countries. These predictable countries are located in the red areas of the Fitness-GDPpc plane (see Fig. 6).[22] Between 2014 and 2019, Thailand is expected to follow the well-bounded pathway of comparators like Denmark or Portugal in the late nineties, and increase its per capita GDP by 5.4 annually (see Table 7).

Growth is more difficult to estimate for countries in the unpredictable (blue) areas of the Fitness-GDP plane, since per capita GDP is influenced by numerous exogenous economic indicators whose dynamics are difficult to model even conceptually. Ghana, for example, belongs to the set of unpredictable countries, where models like SPS or IMF's forecast are less accurate in determining how a country's income will change. While Ghana's per capita GDP is also expected to increase by over 5 percent annually, its comparators follow very heterogeneous trajectories. They do not move in a similar direction across the Fitness-GDPpc plane as is the case with Thailand (Figure 6). Because of this, Ghana's anticipated rise in per capita GDP in 2019 is less certain.

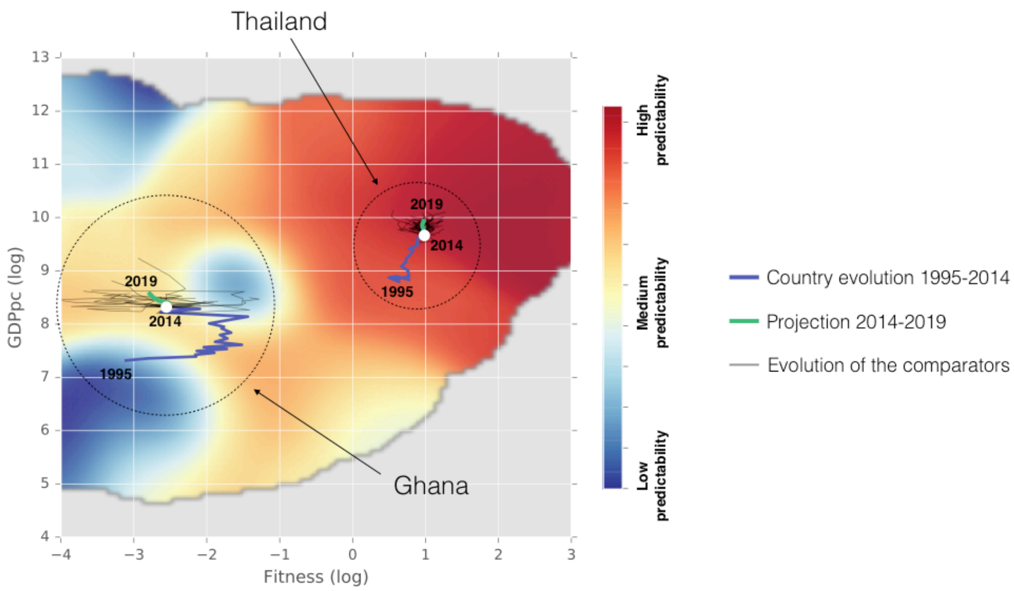

Fig. 6: Ghana's per capita GDP is more difficult to predict than that of Thailand, which is illustrated by their position in the Fitness-GDPpc plane. Thailand's comparators follow a fairly uniform trajectory, while Ghana's comparators are more heterogeneous.

[22] *We refer to Section 5.1.2 (Step 3) for the specifications and a technical discussion of the predictability dimension.*

	Thailand	Ghana
Per capita GDP (PPP), Prediction for 2019	$20,523	$5,274
Per capita GDP (PPP) in 2014	$15,776	$4,102
Predicted CAGR in per capita GDP	5.40%	5.15%
Predictability	1.1	0.5

Table 7: Forecasts and predictability for Thailand and Ghana

While the expected growth rates are similar for Thailand and Ghana, the countries are differentiated by the Predictability dimension. Thailand's predictability has improved from 0.6 in 1995 to 1.1 in 2014 (see figure 7). Following its expected development trajectory (into the darker red area of the Fitness-GDPpc plane) will further increase its predictability and may place it on par with some of its current comparators, such as Hungary or Malaysia. Ghana displays less certain behavior: The trajectories of its comparator set are quite heterogeneous and predictability is low. The progression of Ghana's trajectory will in turn determine the predictability of any future forecasts. It is unclear whether Ghana's per capita GDP will become more or less difficult to estimate. Ghana currently sits at the periphery between the unpredictable and predictable zones of the fitness-GDPpc plane. Depending on changes in its Fitness and per capita GDP, Ghana can move to a more (red) or less (blue) predictable position. The volatility of Ghana's historical predictability (figure 7) illustrates that such switches between more and less predictable regimes are not unprecedented. Countries in different predictability regime are structurally different as, for instance, shown in Table 8 where we report the typical dependence on natural resources of Thailand and Ghana's comparators as a percentage of GDP. Thailand's comparators are countries with a marginal dependence on low complexity primary products as natural resources while Ghana's comparators are economies strongly driven by natural resources (on average 10% of their GDPs).

	Thailand's comparators	Ghana's comparators
Total natural resource rents (% of GDP)	Average: 1.9 Median: 0.8	Average: 11.9 Median: 10.0
Fitness	Average: 2.1 Median: 1.9	Average: 0.1 Median: 0.1

Table 8: Forecasts and predictability for Thailand and Ghana

For the set of unpredictable economies, it is important to consider country-specific exogenous determinants of growth. Ghana, for example, is reliant on natural resource extraction, and changes in commodity prices can affect its per capita GDP. The Predictability dimension can benefit both the SPS and IMF models, as it confirms the predictive power of growth forecasts for laminar countries and highlights where predictive limitations exist. It enables a clear segmentation of predictable and unpredictable countries without requiring specifications for the countless economic indicators that influence country dynamics (see sections 1.1 and 1.4). It also captures predictability changes over time, as countries move between predictable and unpredictable regimes. In this way, Predictability provides information on which countries can be modeled accurately, and which trajectories are likely influenced by the volatility of exogenous dynamics.

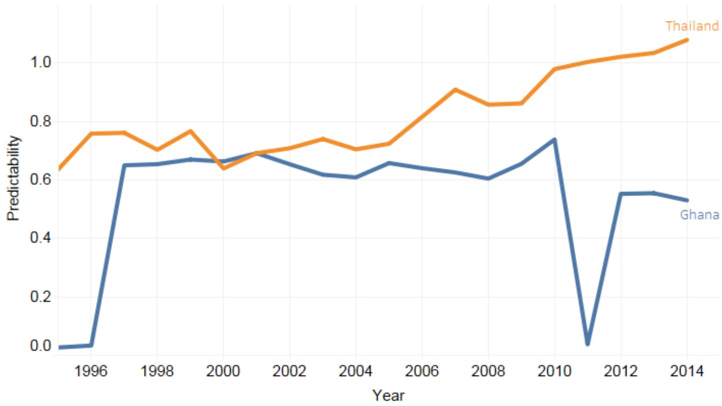

Fig. 7: Predictability of development trajectories for Ghana and Thailand. Thailand's per capita GDP is increasingly easier to estimate, while Ghana's predictability has been volatile. The negative peak in 2011 corresponds to the fact that Ghana enters for a short period in the blue area in the middle of the heatmap discussed in Fig. 6. This region is characterized by an extremely low coherence of the distribution of the comparators.

5. Methods: SPS specifications

In this section, we discuss the SPS mathematical specifications and how the Fitness-GDPpc plane is used to select the comparatives modeling the future we want to estimate.

5.1 Mathematical specifications of SPS

5.1.1 The Fitness-GDPpc plane and the economic dynamics in the plane

Fitness is a measure of country competitiveness and competitiveness is a driver of the differences in growth profiles. More competitive economies are expected to grow faster and consistently than less competitive economies.

The direct comparison defines the Fitness-GDPpc plane [Cristelli et al., 2015], in which it is possible to track simultaneously the evolution of an economic system in terms of endowment competitiveness and per capita monetary performance. The Fitness-GDPpc plane is defined by the natural logarithm of the Fitness and of GDPpc, in such way we can encompass the dynamics of countries over several orders of magnitude (GDPpc ranges approximately from 10^2 USD to 10^5 USD) while Fitness approximately ranges from 0 to 10.

In Fig. 8 we report the local 1-year average displacement of countries in this plane and we visualize the results as a vector field. In detail, we split the Fitness-GDPpc plane according to a grid and we average all 1-year displacements belonging to a box. A 1-year displacement is said to belong to a box if the starting point of this displacement belongs to the box.

This procedure then compares the evolution of economic states (the position of a country in the Fitness-GDPpc plane) which may be *a priori* far in time but are close in the plane, as an example we find that nowadays frontier African countries are potentially on the verge of a sustained industrialization. E.g. Kenya and Uganda are in an economic state in this plane which is close to the position of Vietnam in the early 1990s or, going back further, to the Republic of Korea in the 1960s and 1970s.

The grey band in Fig. 8 is an estimate of the expected level of GDPpc of a country provided its level of Fitness. We refer to Appendix A.3 for the mathematical details of this estimate. Countries below the line have a *per capita* wealth lower than what is expected from Fitness. Conversely for those above, Fitness level does not account for the whole GDPpc. Despite a highly tempting strategy, the signed distance from the Fitness-GDPpc is not trivially correlated with the future growth of countries as witnessed by the smoothed dynamics of Fig. 8.

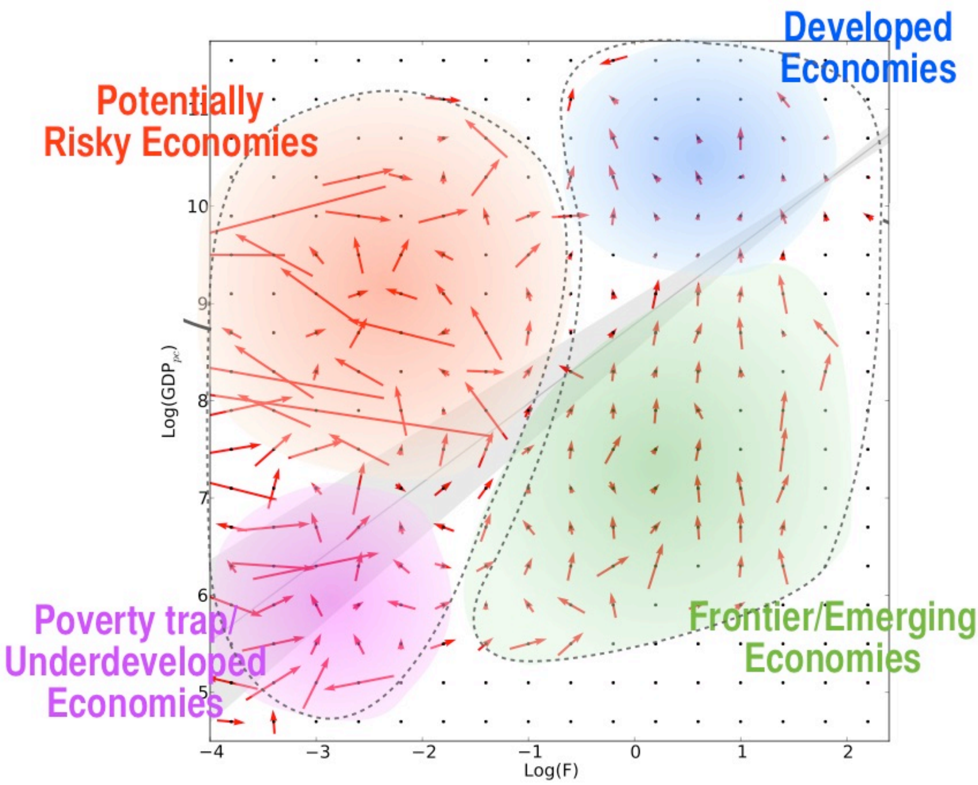

Fig. 8: the heterogeneity of the economic development in the Fitness-GDP plane. The plane is divided into square boxes and arrows are the local 1-year average displacement of countries in each box. Grey band is an estimate of he expected level of GDPpc provided the level of Fitness, see Appendix A.3 for the details of this estimate.

The comparison of economic states allowed by the Fitness dimension yields non-trivial results as witnessed by the fact that the dynamics of countries shows different degree of smoothness. On the right part of Fig. 8 the flow of evolution tends to be *laminar* (or predictable) while in left bottom corner and in left top part the emerging dynamics is much less regular. This leads us to observe that, on average and in the long-term, we expect to find region more predictable than others and where using the behavior of the past will be informative on the future evolution of countries in a similar area of the plane. In this sense the Fitness-GDPpc plane will act as a feature selector in the region where the flow is laminar and regular where close countries tend to evolve on average in a similar way. This heterogeneity of regimes, on one hand, is the main reason why a regressive approach will fail in attempting to forecast growth. Different regions have substantially different dependences on Fitness (and GDPpc) and why instead a non-parametric approach is more fruitful in this context as described in the next section.

Furthermore, if we use for the same comparison the raw diversification of countries (using the notation of Section 2 the diversification of country *c* would be $\sum_p M_{cp}$) instead of Fitness, this would produce a pattern resembling an incoherent or random motion in the diversification-GDPpc plane. This strengthens the non-trivial ordering provided by Fitness. It holds back the concept of closeness induced by this economic dimension as highly informative and effective at reducing the dimensionality embedding the economic dynamics.

5.1.2 SPS: Selecting the right comparatives and definition of the Predictability

The mathematical specifications of the SPS can be described in terms of a 3-step process:

1. Selection of the candidates to be *comparatives* in the Fitness-GDPpc plane;
2. Modeling of growth projections as a result of the average development trajectories derived from the previously selected *comparatives;*
3. Measuring the Predictability of the region the previously selected comparatives belong to as a result of a generalized signal-to-noise ratio.

Provided these specifications, SPS specifications be identified as a two dimensional non-parametric regression (also known as kernel regression [Nadaraya, 1964; Watson, 1964]) where the selected trajectories of the comparatives are playing the role of the sample statistics and the density we want to estimate is calculated only in those points centered at the points in the fitness-GDPpc plane corresponding to the events we want to project. The conceptual basis of the scheme to select comparatives can be found in [Cristelli et al., 2015]. The specific version of the scheme here proposed is a tailored version of the concepts of [Cristelli et al., 2015] to deliver growth projections. SPS as discussed in [Cristelli et al., 2015] is aimed at investigating general features of the flow of the economic dynamics. In this sense, SPS in [Cristelli et al., 2015] can be thought of as a discrete Eulerian specification of the economic flow field. Here instead, as we are interested in modeling the growth projections of specific events, i.e. a country growth over a given time window, we develop a sort of Lagrangian specification of the economic flow dynamics.[23]

Step 1: Comparatives selection

We define the event $e_{c,t}$ as the economic state of the country *c* at time *t.* Considering that our events lie in the Fitness-GDPpc, $e_{c,t}$ is a 2x1 vector representing the logarithm of Fitness and the logarithm GDPpc at time t of country *c*, in formula $e_{c,t} \equiv (f_{c,t}, y_{c,t})$ where for the sake of notation simplicity, in this section we will refer to the logarithm of fitness as simply *f* and to the logarithm of GDPpc as *y.*

Definition: the set of comparatives of a reference event $e_{c,t_0} = (f_{c,t_0}, y_{c,t_0})$ is the set composed of all those events $e_{c',t}$ in fitness-GDPpc plane for which it holds:

- $\| e_{c',t} - e_{ct,_0} \| < r$ where $\|.\|$ denotes the Euclidean distance (i.e. L_2 norm)
- $c \neq c'$
- $t < t_0$

This set turns to be the set of events belonging to the neighborhood of radius *r* of the reference event e_{c,t_0} we are interested in. The parameter *r* is the only parameter of the SPS. The size of this neighborhood is the result of a trade-off between two opposite constraints, on one hand we would like to have the neighborhood as small as possible but, on the other hand, we must have enough samples in the set to allow reliable projections. We discuss in Appendix A.4 the choice of this parameter and the existence

[23] Lagrangian specification of the field is a way of looking at fluid motion where the equation follows an individual fluid parcel as it moves through space and time. Therefore the solution of these specifications represent an individual parcel through time and gives the trajectory of the parcel. This is usually visualized as sitting in a boat and drifting down a river.
Conversely the Eulerian specification of the flow is a way of looking at fluid motion that focuses on specific locations in the space through which the fluid flows as time passes. We therefore divide our dynamics into several box and the specifications define equations for quantity defined in those boxes.

of a range for which the performances of SPS are substantially independent on the value of *r* making the SPS a *de facto* parameter free scheme.

We also include a fourth condition to filter the comparatives of e_{c,t_0}: a country cannot be the source of new comparatives for *n* years after the first selection unless this country exited from the neighborhood defining the comparatives. A natural choice for the value of *n* is the time horizon *D* of the projection we want to build (see Step 2). An example can help to clarify this condition: let us suppose the country *c'*, say Malaysia, in 2000 is selected to be a comparative of the reference country *c*, say Vietnam in 2014 and let us suppose *n=5*. The previously stated condition means that Malaysia cannot be selected from 2001 up to 2005 as a comparative of Vietnam even if the neighborhood condition would be satisfied. Malaysia could be considered as a comparative of Vietnam in 2004 only if Malaysia exited the neighborhood of Vietnam 2014 before 2004 and re-entered into it in 2004. This extra filter is again to avoid biases and overrepresentation of specific countries in the set of comparatives of the reference event. We denote as $\mathcal{C}[e_{c,t_0}]$ the set of comparatives satisfying the four conditions.

Before moving to next section, we want to comment on the first condition, the exclusion of the events arising from the past of the country itself in order to avoid biases due to the inclusion of the country itself.

If self-contribution were not excluded at this stage, we would have projections too autocorrelated with past growth rates in those regions where the set of comparatives is sparse. See Section 5.2 for a simple way to reintroduce in the scheme an autoregressive component to the SPS which is, in the present selection, now completely neglected.

Step 2: Modeling the growth projections

Provided this set of comparatives $\mathcal{C}[e_{c,t_0}]$ and the time horizon *D* of the projection we want to achieve, we have all the ingredients to model the growth estimate on the basis of the comparatives.

Denoting the projected position of the event e_{c,t_0} *D*-period ahead as e_{c,t_0+D}, we have that $e_{c,t_0+D} = e_{c,t_0} + T_{\mathcal{C}}$ where $T_{\mathcal{C}}$ is a 2x1 vector defined as the average[24] of the trajectories of length *D* years and whose starting points are defined by the comparatives in $\mathcal{C}[e_{c,t_0}]$. In formula:

$$T_{\mathcal{C}} = \left[\frac{1}{N_{\mathcal{C}}} \sum_{c',s \,\in\, \mathcal{C}[e_{c,t_0}]} \left(f_{c',s+D} - f_{c',s}\right), \frac{1}{N_{\mathcal{C}}} \sum_{c',s \,\in\, \mathcal{C}[e_{c,t_0}]} \left(y_{c',s+D} - y_{c',s}\right)\right]$$

where $N_{\mathcal{C}}$ denotes the number of events in the set $\mathcal{C}[e_{c,t_0}]$. By denoting the average over the events in $\mathcal{C}[e_{c,t_0}]$ as $\langle\, . \,\rangle$, we can rewrite previous formula in a compact way as:

$$T_{\mathcal{C}} = \left[\langle f_{c',s+D} - f_{c',s}\rangle, \langle y_{c',s+D} - y_{c',s}\rangle\right]$$

Let us exemplify the procedure considering a real case. In Table 9 we report all the comparators selected for Vietnam to model growth from 2014 to 2019. The estimated trajectory of Vietnam in the period 2014-2019 will be then the average of the sample

---

[24] *in this work, we use a simple average but weighted averages are a potential generalization for the growth modeling. Candidates for those weights could be distances suitably defined among countries at sector level or product level.*

trajectory distribution composed of the 53 comparative trajectories selected, namely Brazil from 1995 to 2000, Bulgaria from 1995 to 2000, …, Indonesia from 2000 to 2005, … and finally Tunisia from 2009 to 2014. The comparative countries are extremely heterogeneous on both a geographical basis (only Oceania and North America are not represented in the list) and a size basis (we have for instance Brazil and the Russian Federation on one hand and Albania and Latvia on the other).

The main output of this second step is therefore the average trajectory of this distribution due to the comparatives' trajectories. This average yields our growth projection. Practically we retain only those $\mathcal{C}[e_{c,t_0}]$ which have at least 5 events which represents our minimum statistics threshold. Provided the estimate for e_{c,t_0+D} is straightforward to assess the annualized projected growth rate for the GDPpc as[25]:

$$g_{c,t_0+D,t_0} = (e^{y_{c,t_0+D}-y_{c,t_0}})^{1/D} - 1$$

List of comparative countries selected by SPS for Vietnam to model 2014-2019 Vietnam's growth trajectory					
Brazil, 1995	Mexico, 1995	Kazakhstan, 1996	Indonesia, 2000	Colombia, 2002	Romania, 2005
Bulgaria, 1995	Panama, 1995	Russian Fed., 1996	Jordan, 2000	Panama, 2002	Ukraine, 2006
Croatia, 1995	Poland, 1995	Ukraine, 1996	Romania, 2000	South Africa, 2002	Albania, 2008
Estonia, 1995	Romania, 1995	Panama, 1997	Thailand, 2000	Tunisia, 2003	Belize, 2008
Hungary, 1995	Slovak Rep.., 1995	South Africa, 1997	Turkey, 2000	Bosnia Herz., 2004	Bosnia Herz, 2008
Indonesia, 1995	South Africa, 1995	Latvia, 1998	Belarus, 2001	El Salvador, 2004	Egypt, 2008
Latvia, 1995	Thailand, 1995	Lebanon, 1999	Philippines, 2001	Philippines, 2004	Philippines, 2009
Lithuania, 1995	Turkey, 1995	Brazil, 2000	Russian Fed., 2001	Bulgaria, 2005	Tunisia, 2009
Macedonia, 1995	Belarus, 1996	Bulgaria, 2000	Ukraine, 2001	Indonesia, 2005	

Table 9: list of the 53 comparative countries selected by SPS to model Vietnam growth trajectory from 2014 to 2019.

We can assess the dispersion of the trajectory distribution as well and this leads us to the third step of the SPS, how to measure the *Predictability* dimension.

Step 3: Measuring the Predictability

The Predictability $P[e_{c,t_0}]$ conceptually speaking, is similar to a signal-to-noise ratio or to an information gain/loss ratio. The value is defined as the ratio of the initial dispersion of the comparatives (the dispersion of the starting position σ_0 of the trajectories selected to be the comparatives of e_{c,t_0} and the dispersion σ_{evo} of comparatives after *D* years (the dispersion of the final position of the trajectories selected to be the comparatives of e_{c,t_0}):

$$P[e_{c,t_o}] = \frac{\sigma_0}{\sigma_{evo}}$$

The dispersion σ can be defined in two different ways: i) as the sum of the dispersions along the two dimensions which corresponds to the sum of the standard deviations of the comparators' (logarithms of) Fitness and GDPpc,[26] i.e. the dispersion of the two sample joint probability distributions; or ii) as the square root of the sum of the two squared

[25] *Since y denotes the logarithm of the GDPpc, exp(y_{t+D} -y_t) represents the ratio of the ending value to starting value.*

[26] *In σ_0, the sum of the SD of the initial position of the comparators, In σ_{evo}, the sum of the SD of the position of the comparator countries after D-periods.*

dispersions. The two specifications provide essentially the same results and we adopt the former definition for all the graphs shown in this work.

It is worth noticing that the scheme also allows to measure separately the two terms contributing to the predictability, the predictability along Fitness and the one along GDPpc.

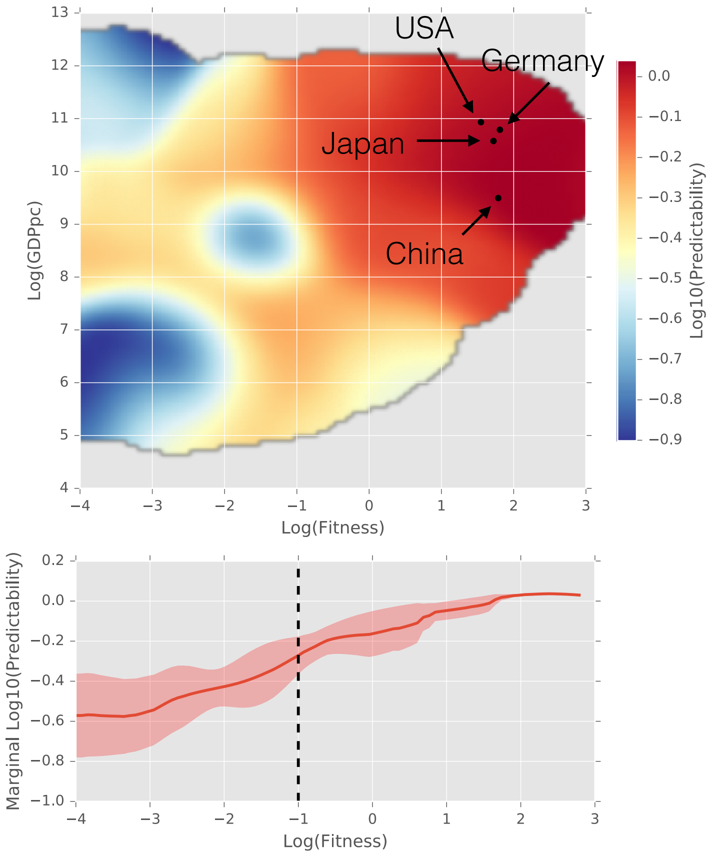

Figure 9: (Top panel) we visualize the heterogeneous pattern of the Predictability. We report the position in the plane of China, Germany, Japan and United States in 2014 as a reference. Interestingly China is entering in one of the most predictable areas of the fitness-GDPpc plane. (Bottom panel) we marginalize the predictability dimension along the Fitness axis. The threshold *ln(Fitness) = -1* we choose corresponds approximately to the 50th percentile (i.e median) of the marginalized predictability.

The predictability *P* is therefore a non-negative number defined in the range [0,+∞). The limit $P \to +\infty$ would correspond to a vanishing dispersion of the evolution of the comparatives after *D* years since σ_0 is, in all practical scenarios, a finite value whose order of magnitude is set by the neighborhood size *r.* This would correspond to a perfect predictability scenario because all the comparatives evolved towards the same economic state in the Fitness-GDPpc plane. Conversely, if $P \to 0$ the predictability is vanishing as this correspond to a scenario of an infinitely dispersed evolution of the comparatives and therefore the knowledge of the past would be useless. The case $P \approx 1$ corresponds to a scenario in which the dispersion of the comparatives after *D* years is similar to the starting one and therefore this value is a divide between a gain (*P>1*) and a loss (*P<1*) of predictive resolution.

In order to visualize the behavior of this dimension, we perform a kernel regression in the Fitness-GDP plane with a Gaussian kernel as shown in Fig. 9 (top panel). The pattern reveals a strong degree of heterogeneity of the Predictability, the color scheme is logarithmic and the above defined ratio covers almost two decades.

The threshold *ln(Fitness) = -1* we use to define predictable 'P' and unpredictable 'UnP' countries is approximately the median (50th percentile) of the Predictability marginalized along Fitness axis as shown in Fig. 9 (bottom panel). The threshold corresponds approximately to $log_{10}(P) = -0.25$, namely a predictability of 0.5 which means that the dispersion of the comparators' evolution is approximately twice the dispersion of the originating point of the comparators' trajectories.

5.2 SPS + trend: Accounting for the autoregressive growth component

The exclusion of the self-contribution of a country as a comparative on one hand avoids biases, especially in those areas of the fitness-GDPpc where the density of comparatives is very low. The exclusion is indeed substantially negligible for a large set of comparatives. An example conversely for which the self-contribution would lead to a growth projection strongly biased by the past is the case of China, where the number of comparatives is in the range of 5 to 10. The scheme would hardly detect the smooth slowdown of China growth rates from double digit rates to current values in the range 6-8%.

However, neglecting totally this autoregressive term means to discard a variable which we know accounts for approximately 10% of the variance of the quantity we want to project. The preferred way to deal with such issue would be to add a third dimension to our scheme, the past growth rates and perform the comparative search in this 3-dimensional space. A point in this new space would be a three-dimensional array specified by fitness, GDPpc and past value of growth of the GPDpc. The neighborhood of an event would be then a sphere of radius *r* and formally the specifications to implement SPS would be similar. However, in such a setup we would deal again with the curse of dimensionality and, as a general trend, we would have much smaller statistics to estimate growth projection. The number of countries for which we would be able to yield a growth estimate would be drastically reduced.

With the priority being the delivery of growth estimates for as many countries as possible, we propose an alternate scenario to reconcile SPS with an explicit term accounting for the autoregressive component of growth.

We propose a simple model, SPS + trend, in the form of a regressive scheme as follows:

$$g^{SPS+trend} = \alpha\, g^{SPS} + \beta_1\, g^{trend\ 1y} + \beta_2\, g^{trend\ 2y}$$

where g^{SPS} is the estimated growth rate as described in the previous section, $g^{trend\ 1y}$ and $g^{trend\ 2y}$ are the growth rates one and two years before the starting time of the period we want to project. As long as a term including recent past trend is provided in the model, we find substantially similar results. The proposed specifications are the ones which best perform (the difference among the different specifications are marginal). β_1 and β_2 are trained on the time window 1995-2005 via a standard OLS optimization. We refer to Section 3.5.2 for the motivation for which the trend term is not needed for modeling growth estimates of countries in the predictable regime.

6. Discussion and Perspectives

The ability to assess a consistent measure of the competitiveness of a country's productive structure proves to be a crucial element when approaching the complex task of predicting growth. Such measures are traditionally built with a bottom-up approach, i.e. by an informed aggregation of a large number of indicators. This way of measuring competitiveness suffers from many shortcomings, most importantly the heterogeneity of such indicators across different countries and the objective difficulty of defining suitable "rules of sum" that allow to effectively synthesize such indicators into a coherent measure of competitiveness.

In this work, we point out how reversing this approach allows the definition of a measure of competitiveness, the fitness, that is extremely effective in capturing the growth potential of countries within a simple forecasting model. The Fitness is defined as a measure of the outcome of a country's productive structure in terms of diversity and complexity of produced products. This definition provides the advantage of i) relying on a single global dataset of world trade, which is consistent across countries and widely available, and ii) automatically incorporate the "rules of sum" of capabilities by looking directly at the outcome that these capabilities allow.

Being an intensive measure of competitiveness, the Fitness is naturally compared with *per capita* GDP. Countries display a remarkably peculiar dynamical behavior when analyzed with respect to fitness and GDPpc time series, that allow us to define two distinct regimes of growth, a predictable - or laminar - regime, and a unpredictable - or chaotic - regime. In the laminar regime, we observe that countries within a specific region of the fitness-GDPpc plane tend to have similar growth patterns. This allows us to define the Selective Predictability Scheme (SPS), that uses past trends of selected comparator countries, that have been in the same area of the Fitness-GDPpc plane in the past, to predict future growth.

The main advantages of the SPS are three:

- It naturally provides an accurate measure of the forecasts, that is defined as the average predictability of past trajectories in the neighborhood of the country's present position in the fitness-GDPpc plane;
- Such accuracy is not a mere property of the SPS method, but rather an estimate of the actual "predictability" of a country's economy, that implicitly affects the accuracy of other forecasting models as well (e.g. IMF);
- It achieves results in terms of forecasting errors that are comparable with very complex, multi-parameter, country-specific models by using one global model, virtually parameter-free, to forecast every country in the data set.

An approach like the SPS is able to describe and forecast effectively medium-long term growth patterns, as it is based on a fundamental measure of competitiveness, that drives countries over a several-year time horizon. For this reason, it must be complemented with other, more fine grained methods when there is need to forecast on shorter time horizons, or to incorporate informed knowledge about relevant industrial and monetary policies, and international and local scenarios.

References

Aghion, P., and Howitt, P. (1990) *A model of growth through creative destruction*. No. w3223. National Bureau of Economic Research.

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

Barro, R. (1991). *Economic Growth in a Cross Section of Countries.* Quarterly Journal of Economics 102(2), pp. 407–443.

Bellman R. E. (1957). Dynamic programming. Princeton University Press

Bengio, Y., Lamblin, P., Popovici, D., & Larochelle, H. (2007). *Greedy layer-wise training of deep networks*. Advances in neural information processing systems, 19, 153.

Cecconi, F., Cencini, M., Falcioni, M. and 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., and 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., and Pietronero, L. (2015). *The heterogeneous dynamics of economic complexity*. PLoS ONE 10(2): e0117174.

Domínguez-García, V. and Munoz, M. A. (2015). *Ranking species in mutualistic networks*. Scientific Reports 5, Article number: 8182

Fisher, F. M. (1966) *The Identification Problem in Econometrics*. McGraw–Hill Book Co.

Gaulier G. and Zignago S. (2010) *BACI: international trade database at the product-level (the 1994–2007 version)*.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. A. (2014). *The atlas of economic complexity: Mapping paths to prosperity*. Mit Press.

Hausmann, R., Hwang, J. and Rodrik, D. (2007) *What you export matters.* J Econ Growth 12: 1

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

International Monetary Fund (2016). *World Economic Outlook: Subdued Demand: Symptoms and Remedies*. Washington.

Jolliffe, I. (2002). *Principal component analysis.* John Wiley & Sons, Ltd.

Kleinberg, J., Ludwig, J., Mullainathan, S. and Obermeyer, Z. (2015). *Prediction Policy Problems.* The American economic review, 105(5), 491-495

Kohonen, T. (1982). *Self-organized formation of topologically correct feature maps. Biological cybernetics*, 43(1), 59-69.

Kraay, A., and Monokroussos, G. (1999). *Growth forecasts using time series and growth models*. World Bank eLibrary

Lall, S. (2001). *Competitiveness indices and developing countries: an economic evaluation of the global competitiveness repor*t. World development, 29(9), 1501-1525.

Lorenz, E. N. (1969). *Atmospheric predictability as revealed by naturally occurring analogues*. Journal of the Atmospheric sciences, 26(4), 636-646.

McArthur, J. W., & Sachs, J. D. (2001). *The growth competitiveness index: Measuring technological advancement and the stages of development.* The global competitiveness report, 2002, 28-51.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). *Distributed representations of words and phrases and their compositionality*. In Advances in neural information processing systems (pp. 3111-3119).

Nadaraya, E. A. (1964). *On Estimating Regression*. Theory of Probability and its Applications. 9 (1): 141–2. d

Page, L., Brin, S., Motwani, R. and Winograd, T. (1999) *The PageRank citation ranking: bringing order to the web*. Technical Report, 1999.

Podobnik, B., Horvatić, D., Kenett, D. Y. and Stanley, H. E. (2012) *The competitiveness versus the wealth of a country*. Scientific Reports 2 (678)

Pritchett, L. and Summers, L. H. (2014). *Asiaphoria meets regression to the mean.* Technical report, National Bureau of Economic Research.

Pugliese, E., Zaccaria, A., and Pietronero, L. (2014). *On the convergence of the Fitness-Complexity Algorithm*. arXiv preprint arXiv:1410.0249.

Pugliese, E., Chiarotti, G.L., Zaccaria, A., and Pietronero, L. (2015). *Complex Economies have a Lateral Escape from Poverty Trap* arXiv e-print 1511.08622.

Romer, P. M. (1986). *Increasing returns and long-run growth.* The journal of political economy, 1002-1037.

Solow, R. M. (1956) *A contribution to the theory of economic growth*. The quarterly journal of economics, 70(1):65–94,

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., and Pietronero, L. (2012). *A new metrics for countries' Fitness and products' complexity*. Scientific reports, 2, 723

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

The Economist, *A mean feat.* June 9th 2016 http://www.economist.com/node/21685480

Watson, G. S. (1964). *Smooth regression analysis*. Sankhyā: The Indian Journal of Statistics, Series A. 26 (4): 359–372

World Economic Forum (2016), *The Global Competitiveness Report 2015-2016*

Wu, R.-J., Shi, G.-Y., Zhang, Y.-C., and Mariani, M. S. (2016). *The mathematics of nonlinear metrics for nested networks.* arXiv preprint arXiv:1604.06284.

Zaccaria, A., Cristelli, M., Tacchella, A., and Pietronero, L. (2014). *How the taxonomy of products drives the economic development of countrie*s. PLOS ONE 9(12): e113770

Appendix A: Results

Appendix A.1: Comparison SPS vs IMF accuracy (full table)

We provide a table summarizing all the results provided in Table 1-6 of Section 3.

	PC			MAE			RMSE			Observ.		
	All	P	UnP	All	P	UnP	All	P	UnP	All	P	UnP
SPS	0.30	0.42	0.21	2.01	1.81	2.26	2.64	2.33	3.00	758	429	329
SPS + Trend	0.37	0.38	0.42	1.97	1.94	2.08	2.60	2.51	2.75	758	429	329
IMF	0.42	0.42	0.41	2.03	1.87	2.22	3.05	2.46	3.66	384	213	171

Table A.1: Comparison of SPS and IMF performances with respect to three metrics: Pearson Correlation (PC), Mean Absolute Error (MAE) and Root Means Square Error (RMSE). For SPS we propose the results of two specifications of the scheme, the latter explicitly accounting for past growth trends. We color in green in each column the model which best performs. Considering all available countries, columns 'All', in terms of correlation IMF slightly outperforms SPS while in terms of average errors SPS outperforms IMF. When we consider countries only from the predictable regions as measured by the Fitness-income plane (columns 'P'), performances tend to be improved both for IMF and SPS. Results in columns 'UnP' correspond to select only country from the low predicability region.

Appendix A.2: Projection accuracy vs GDP sector composition

SPS relies on Fitness and Fitness is an export driven measure of cross-country diversity relying on trade data accounting mainly for manufacturing. Here we show that these features do not introduce any significant bias: SPS' projection accuracy is essentially independent on the specific country GDP sector composition. We consider six economic indicators for each country[27]: services, manufacturing, industry, agriculture and natural resources rents; all indicators are provided as a percentage of GDP. We also include the export share to address whether there are systematic differences between countries driven by domestic market or by export.

We compare projection error distribution and MAE as defined in Section 3 but we segment countries in four groups for each variable. These four groups are defined by the sector composition quartile a country belongs to. Let us consider for instance manufacturing, countries in the first quartile of manufacturing GDP share are those countries whose share is between 0 and 25th percentile of the manufacturing GDP share distribution. Those in the second quartile have shares between the 25th and 50th percentile

[27] *Source: World Bank Data platform http://data.worldbank.org*

and so on for the third and fourth quartiles. The fourth quartile is composed of countries which are dominated by a specific sector (or are export driven in the case of export share). Conversely the first quartile is composed of countries for which the sector (or the export) is less relevant. For instance, in the time windows we test SPS on, manufacturing accounts on average for 6.9% of GDP for countries in the first quartile and for 22.3% in the fourth quartile. We measure and compare MAE and error distribution for each quartile of each variable. MAE error bars are estimated via bootstrapping while error distributions are compared via a box plot. As shown in the panels in the following, we conclude that SPS accuracy is marginally dependent on the country sector GDP composition and on the size of the domestic market.

As expected for the first quartile in the case of manufacturing and industry, we observe larger errors as Fitness is essentially a blind proxy being those sectors negligible. The remaining quartiles show instead a higher and similar accuracy.

For agriculture, accuracy is slightly decreasing as a function of the quartile and this can be explained in terms of an increasing exposure to the exogenous volatility of food commodities. For services, we have the same pattern observed for manufacturing and industry, larger errors for countries in the first quartile and constant profile for the remaining ones.

Considering natural resources rents, we have a pattern similar to Agriculture and, again, it can be interpreted as a signal of countries with an increasing exposure to the volatility due to commodities.

A slightly upward and expected trend of the accuracy (i.e downward MAE) is observed as a function of the GDP share due to export.

A2.1 Manufacturing

Manufacturing (percentiles)	Average share (% GDP)
0-25	6.9
25-50	12.3
50-75	16.1
75-100	22.3

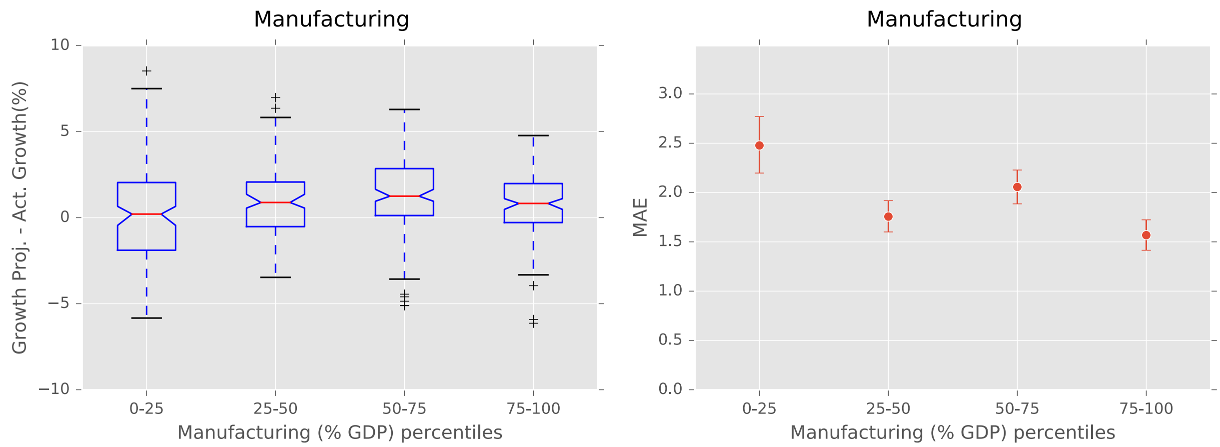

Fig. A.2.1. boxplot comparison of the SPS error distribution (left panel) and mean absolute error (right panel) as a function of manufacturing GDP share quartile.

A2.2 Industry

Industry (percentiles)	Average share (% GDP)
0-25	18.9
25-50	26.0
50-75	31.3
75-100	44.9

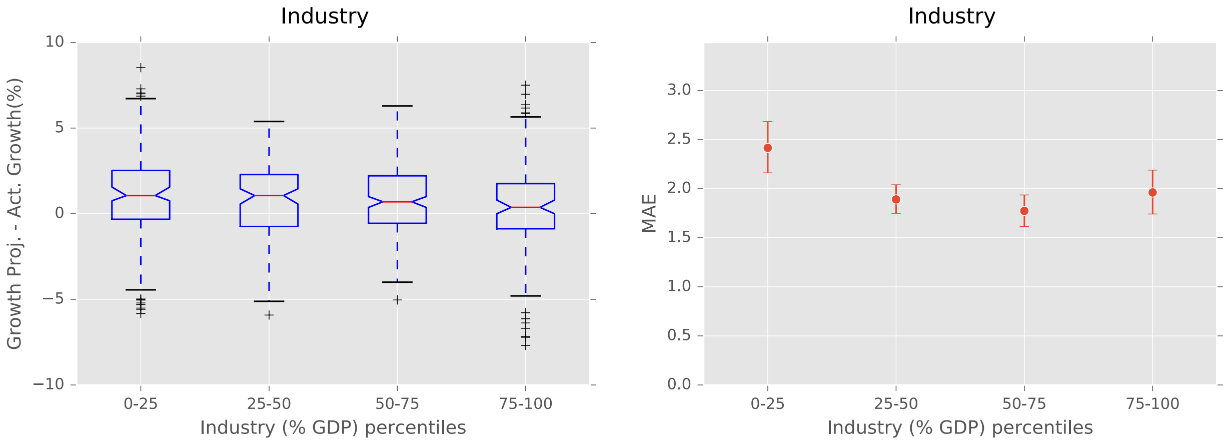

Fig. A.2.2: boxplot comparison of the SPS error distribution (left panel) and mean absolute error (right panel) as a function of industry GDP share quartile.

A2.3 Agriculture

Agriculture (percentiles)	Average share (% GDP)
0-25	1.9
25-50	6.2
50-75	15.4
75-100	32.8

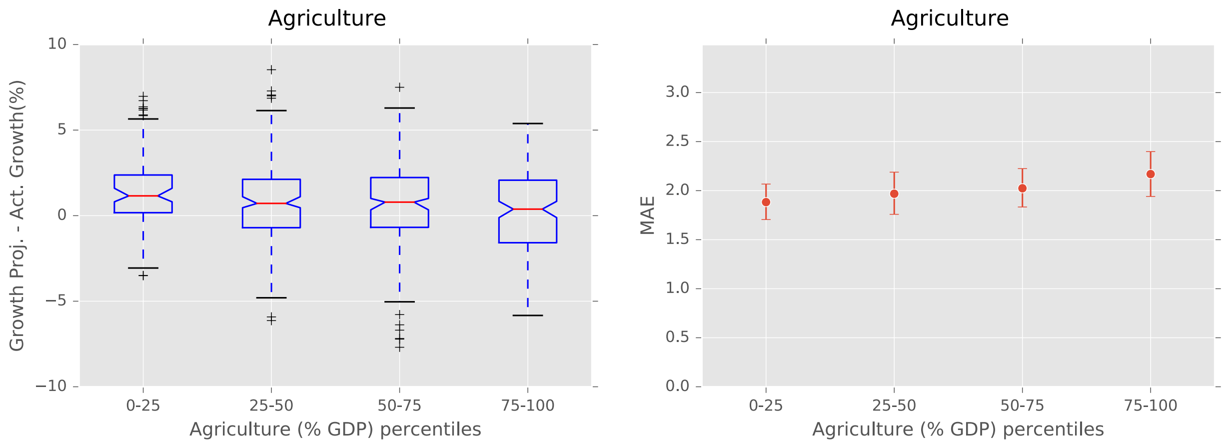

Fig. A.2.3: boxplot comparison of the SPS error distribution (left panel) and mean absolute error (right panel) as a function of agriculture GDP share quartile.

A2.4 Services

Services (percentiles)	Average share (% GDP)
0-25	38.5
25-50	51.4
50-75	60.9
75-100	71.5

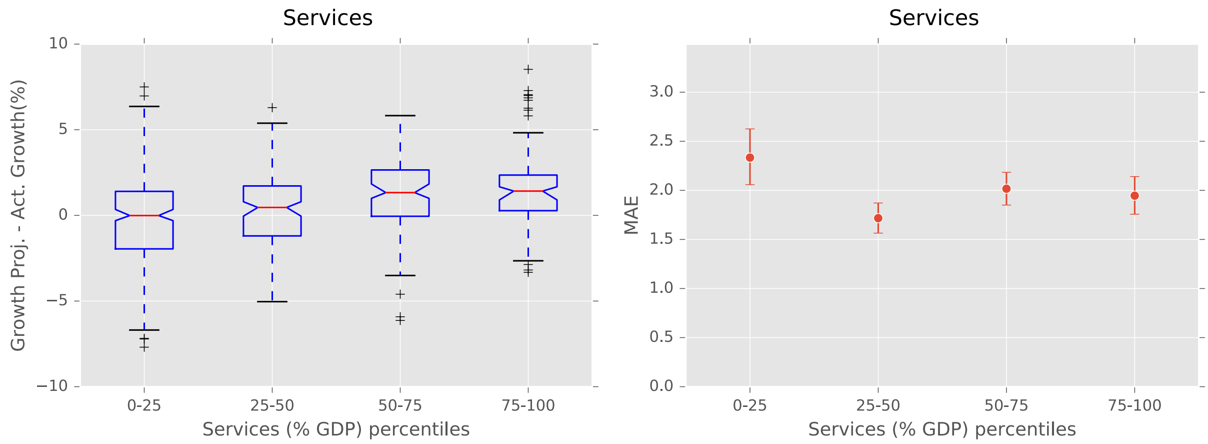

Fig. A.2.4: boxplot comparison of the SPS error distribution (left panel) and mean absolute error (right panel) as a function of services GDP share quartile.

A2.5 Natural resources rents

Natural Resources (percentiles)	Average share (% GDP)
0-25	0.5
25-50	3.1
50-75	9.4
75-100	31.7

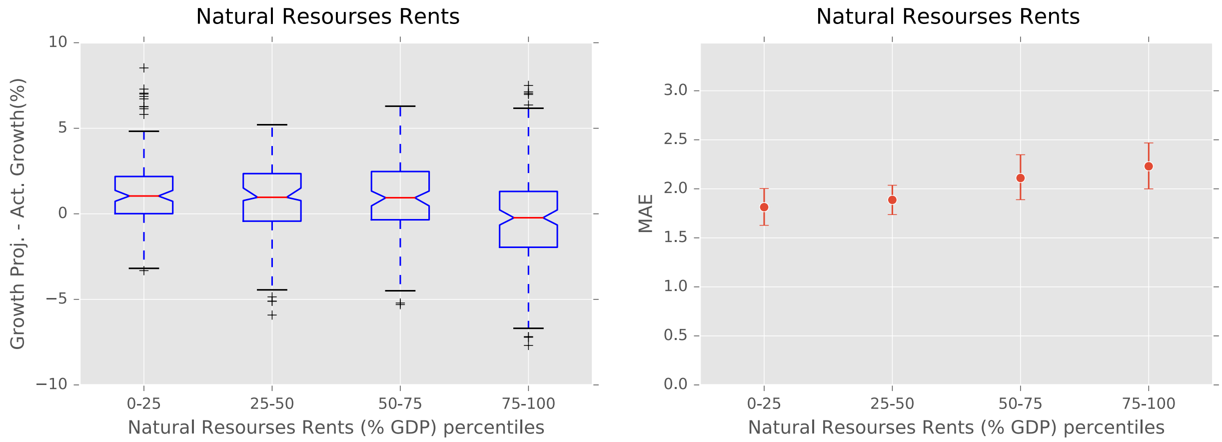

Fig. A.2.5: boxplot comparison of the SPS error distribution (left panel) and mean absolute error (right panel) as a function of natural resources rents GDP share quartile.

A2.6 Export (goods and services)

Goods&Services Exports (percentiles)	Average share (% GDP)
0-25	18.0
25-50	29.7
50-75	43.1
75-100	69.8

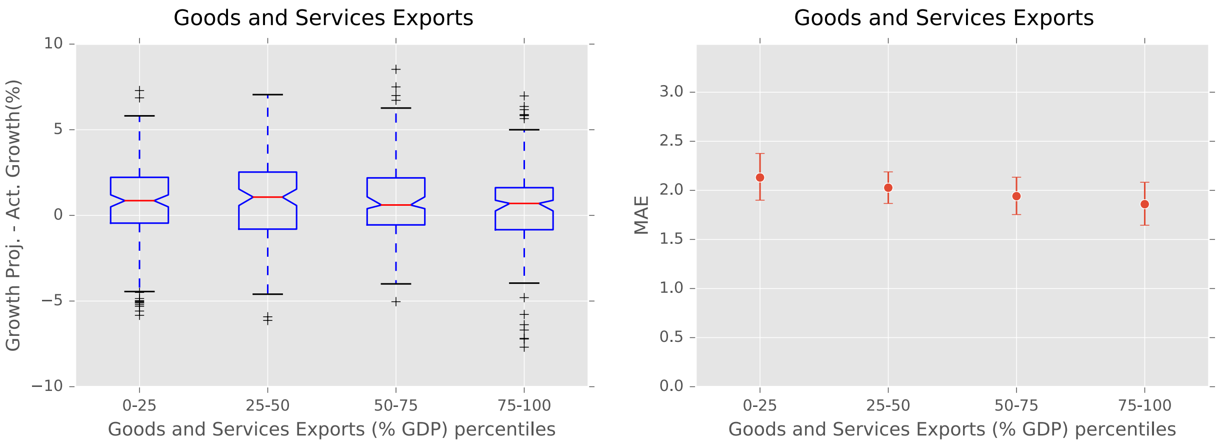

Fig. A.2.6: boxplot comparison of the SPS error distribution (left panel) and mean absolute error (right panel) as a function of export GDP share quartile (goods and services).

Appendix A.3: Fitness-GDPpc relationship

In order to estimate the expected level of GDPpc provided the level of Fitness of a country, we estimate the parameters of a linear relationship between the logarithm of the Fitness and of GDPpc which minimizes the weighted Euclidian distance from the scatter plot points as follows:

$$\begin{cases} a^* \log(F) + b^* \log(GDPpc) + c^* = 0 \\ \{a^*, b^*, c^*\} \quad \to \quad \min_{a,b,c} d \\ d = \sum_c w_c |a \log(F_c) + b \log(GDPpc_c) + c| \end{cases}$$

The weights w_c are the share of a country's GDP with respect to world GDP.

Appendix A.4: Dependence of SPS results on *r*

We replicate the analysis of Section 3.4 in order to test the variability of SPS' results as a function of the parameter *r* setting the size of the neighborhood to select comparators. In order to make a *ceteris paribus* comparison as in Section 3.4 we must further restrict the set of countries we retain for the analysis as we have to enforce that countries must be available for all time horizons and all values of the parameter *r.* We show the results in Fig. A.4.1 and Fig. A.4.2. In terms of accuracy, results profile is essentially the same regardless of the value of *r* while increasing values of *r* appears to marginally enhance the directional correlation. All results shown in this paper are obtained setting *r=0.6*, a value belonging to the range in which SPS' results are essentially independent on the value of this parameter.

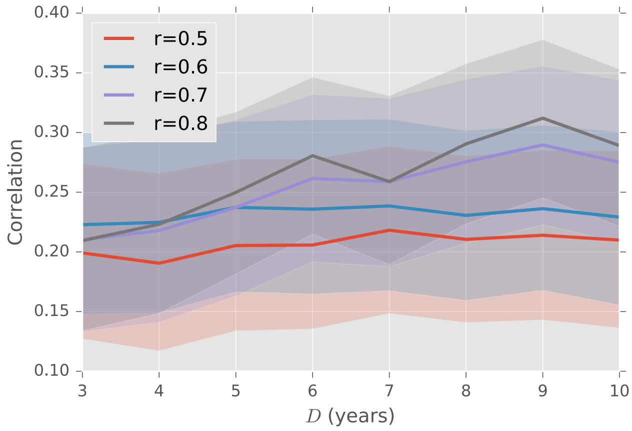

Fig. A.4.1: Correlation between projected and actual growth rates for time windows with starting point in the interval 2000-2004 as a function of the forecasting time horizon expressed in years and of the parameter *r.* CI is estimated via bootstrapping. All points are estimated with 600 observations. Increasing value of *r* slightly increases correlation.

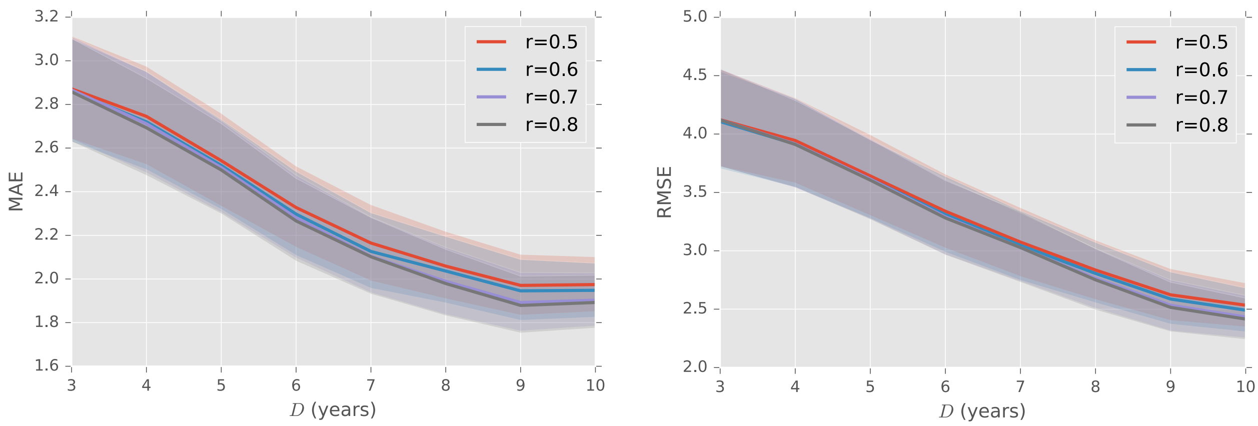

Fig. A.4.2: MAE (left panel) and RMSE (right panel) of SPS projected growth rates for time windows with starting point in the interval 2000-2004 as a function of the forecasting time horizon expressed in years and the parameter *r.* CI is estimated via bootstrapping. All points are estimated with 600 observations. Accuracy turns out to be independent on *r* in the range investigated.

Appendix A.5: Performances of the reference models in the '*P*' and '*UnP*' regime

	'All'	'P' – Laminar/Predictable	'UnP' – Chaotic/Unpredictable
	Model 1	**Model 1**	**Model 1**
MAE (CI)	2.71 (2.52,2.89)	2.83 (2.59,3.07)	2.57 (2.28,2.88)
Accuracy gain % (ref. 'All')	0.0%	**-4.4%**	**5.1%**
RMSE (CI)	3.79 (3.49,4.08)	3.76 (3.45,4.08)	3.82 (3.28,34.35)
Accuracy gain % (ref. 'All')	0.0%	**0.8%**	**-0.8%**
N. Obs.	763	424	339

	'All'	'P' – Laminar/Predictable	'UnP' – Chaotic/Unpredictable
	Model 2	**Model 2**	**Model 2**
MAE (CI)	4.62 (4.25,5.02)	4.57 (4.11,5.06)	4.71 (4.12,5.35)
Accuracy gain % (ref. 'All')	0.0%	**1.1%**	**-1.9%**
RMSE (CI)	7.07 (6.20,8.03)	6.76 (5.90,7.65)	7.44 (5.81,9.19)
Accuracy gain % (ref. 'All')	0.0%	**4.4%**	**-5.2%**
N. Obs.	760	424	336

	'All'	'P' – Laminar/Predictable	'UnP' – Chaotic/Unpredictable
	Model 3	**Model 3**	**Model 3**
MAE (CI)	5.42 (5.02,5.88)	5.63 (5.13,6.19)	5.17 (4.51,5.92)
Accuracy gain % (ref. 'All')	0.0%	**-3.9%**	**4.6%**
RMSE (CI)	8.10 (6.87,9.44)	7.88 (6.67,9.35)	8.36 (6.28,10.8)
Accuracy gain % (ref. 'All')	0.0%	**4.1%**	**-4.8%**
N. Obs.	760	424	336

Appendix B: Methods

Appendix B.1: Exports flows and Specifications of the Revealed Comparative Advantage

Exports are economic outputs endowed with four non-trivial features which are not shared by internal production: i) they are the results of forces shaped by international competition, ii) export data are standardized and homogeneous cross country and cross sector, as a result of the harmonization of customs offices, iii) they are available up to a disaggregate level which is deep enough to pinpoint the heterogeneous structure of productive networks and iv) they are consistently available starting from the 1960s.

The RCA is a non-linear filter which compares two shares: the share of the product with respect to the country export basket and the share of this product with respect to the total volume (in this case with respect to the world GDP due to export). RCA achieves the non-trivial result of filtering out the trivial correlation between country sizes and export volumes. RCA is indeed a relative and multi-scale threshold rather than a flat thresholding procedure.

The entries of the RCA matrix are defined as follows:

$$RCA_{cp} = \frac{\frac{q_{cp}}{\sum_{p'} q_{cp'}}}{\frac{\sum_{c'} q_{c'p}}{\sum_{c'p'} q_{c'p'}}}$$

where q_{cp} is the export of country c of product p expressed in current USD. We define the binary country-product matrix M whose entries M_{cp} are defined as follows:

$$M_{cp} = \begin{cases} 1 & if\ RCA_{cp} \geq 1 \\ 0 & otherwise \end{cases}$$

where RCA_{cp} are the entries of the RCA matrix. The matrix M straightforwardly defines the topology of the bipartite country-product network: an edge exists between a country and a product if (and only if) the corresponding entry of the matrix M is 1.