

# LCPS

المركز اللبناني للدراسات  
The Lebanese Center  
for Policy Studies

Policy Paper

## Arab Country Product Space Report **Introduction and Methodology**

Sebastian Bustos & Muhammed Ali Yildirim

FEB 17



Founded in 1989, the Lebanese Center for Policy Studies is a Beirut-based independent, non-partisan think tank whose mission is to produce and advocate policies that improve good governance in fields such as oil and gas, economic development, public finance, and decentralization.

This research was funded by the International Development Research Center



**IDRC | CRDI**

International Development Research Centre  
Centre de recherches pour le développement international

**Canada** 

Copyright© 2017  
The Lebanese Center  
for Policy Studies  
Designed by Polypod  
Executed by Dolly Harouny

Sadat Tower, Tenth Floor  
P.O.B 55-215, Leon Street,  
Ras Beirut, Lebanon

T: + 961 1 79 93 01  
F: + 961 1 79 93 02  
[info@lcps-lebanon.org](mailto:info@lcps-lebanon.org)  
[www.lcps-lebanon.org](http://www.lcps-lebanon.org)

# Arab Country Product Space Report

# Introduction and Methodology

## Sebastian Bustos

Sebastian Bustos is a PhD candidate in public policy at Harvard University and a doctoral fellow at the Center for International Development at Harvard University. Bustos's research focuses on how international competition impacts local markets and the benefits and challenges presented by multinational corporations operating in developing countries. He holds a master's degree in public administration and international development from Harvard University and a BS in economics and business from the University of Chile.

## Muhammed Ali Yildirim

Muhammed A. Yildirim is an assistant professor of economics at Koç University in Istanbul and an associate at the Center for International Development at Harvard University. He specializes in studying network and spillover effects in a multitude of research areas including industrial policy, international trade, productivity, and economic growth. Yildirim is a co-author of *The Atlas of Economic Complexity* and contributes to the back end of the associated website, which contains millions of data visualizations on the industrial structure of various countries. He obtained his PhD in applied physics from Harvard University and BS degree from the California Institute of Technology.

## Introduction

During the decades prior to the Arab spring, Arab countries on average had one of the lowest per capita growth rates compared to other parts of the world. Although countries in the region registered higher growth rates at the beginning of the century, this economic growth did not bring about a structural transformation required to diversify Arab economies. The challenge remains not only in how to ensure these economies grow but also to ensure that growth is inclusive and sustainable.

The general objective of the report is to lay out the strategic options Arab countries have in undertaking structural transformation. Using the product space methodology, the existing capabilities of selected economies are analyzed by locating products they currently export, which assists in determining which path they should follow to produce more sophisticated and strategic products.

The countries selected for this report can be separated between resource-rich and non-resource rich Arab countries in terms of their main exports. As will be seen later in figure 1, seven of the twelve countries are large oil producers, with natural resources accounting for more than 90% of those seven countries' exports. On the other hand, non-resource countries, while having a more diversified economy, rely primarily on low-technology exports and raw and processed natural resources.

Given their different positions within the product space, studied countries can also be examined in a number of ways. Those countries afflicted by war in recent years (Iraq, Syria, and Libya) have seen their industrial capacity decimated and therefore have few remaining competitive industries. Oil is the only major export for all three countries, and given that the production of this natural resource is in the periphery, the methodology recommends building production capability in less complex products, such as those in the textiles and foodstuffs clusters. Of the three, Iraq's situation is slightly less dire, as the recommendations also include mechanical and chemical products, which are close to its oil industry, of a higher complexity, and would open up further diversification possibilities in the future.

For countries whose export basket is dominated by oil the outlook is more positive than those struck by war. Nevertheless, the great dependency on oil, a product that has few linkages to other products in the space, makes moving toward more complex production difficult. Despite this limitation, the path for development is largely in the chemical, plastics, and mechanical cluster. In most cases, taking advantage of the petrochemical sector would be key, as it would entail utilizing large oil industry capabilities already available in the country.

Countries such as Egypt, Jordan, and Lebanon have a somewhat brighter outlook, as their diversified product space puts them in a better position to further increase the complexity of their respective industrial capabilities. For these countries, the suggested path could also entail making products in the chemical, mechanical, and plastics sectors, a result that is relatively easier to attain with the productive knowledge available in the country.

Finally, possible export destinations are explored. In general, trade within the Middle East is healthy in terms of the share of products exported to nearby countries. Additionally, for most countries, Asia is an important market and the study finds that Africa has increased its relevance in recent years. According to data gathered on the studied countries, Europe (particularly Germany and France) receives less than its fair share of exports. Therefore, measures to support and facilitate exports to this region should be considered.

## I Methodology

### a Economic growth in light of complexities

This report attempts to map future opportunities and missed chances of countries using a methodology that takes into account the productive structure of studied countries. Using this methodology, the study goes beyond the usual one-size-fits-all approaches that only focus on the global quality measures of a given country but fail to integrate each country's specific economic conditions to pinpoint areas of growth. Hence, recommendations can be made based on the capabilities of countries by taking into account a given country's current context and the nature of industries present in the country.

Until recently, research on growth strategies has not focused on diagnosing constraints to a country's economic growth that could help formulate solutions suited to each nation's specific economic and institutional conditions. Implicitly, the policy search concentrated on global recommendations that maximized the universality of evidence in favor of particular interventions but overlooked a country's context, productive structure, or culture. This problem becomes even more challenging when considering jump-starting growth in countries that primarily rely on natural resources and have lacked or are lacking the necessary productive base. Current macroeconomic descriptions underestimate the complexity of the world by describing countries based on a few aggregate factors. Aggregation has gone wild to the extent that current descriptions of countries' economies assume states as different as Iraq and Japan have similar economic components, without adequately taking into account the complexities in each case study.

There can be little comprehensive understanding when complexities are neglected. Aristotle assumed over 2,000 years ago that the world

was made of earth, water, wind, and fire, whereas, in the modern age their economic equivalents are assumed to be land, capital, human capital, and labor. A major leap in our understating of the world came when we moved beyond the Aristotelean view and realized that all material in the world is formed of atoms. Analogously, it is necessary to identify the building blocks of economies and understand economic growth in light of these blocks. Here, what should be focused on is how a product is made, brought into the market, and delivered to consumers. For instance, delivering fresh roses from a field to an individual requires financing the operation, obtaining rose seeds, planting roses, a specialist optimizing the vegetation environment, harvesting roses at the right time, moving roses through a cold-storage logistic chain, and delivering them to flower shops. All these operations can also be broken down further and each of these tasks requires some knowledge specific to an individual or machine. Many of these capabilities can be used in other products. For instance, a cold-storage chain can also be used in the fresh produce industry, roads and shipping facilities can be used in many different industries, etc. Hence, the atoms of production are a set of capabilities and productive knowhow that are used to produce a good (or a service).

Given this atomistic view of production, how can economic growth be triggered? It is important to note that countries do not grow rich by making more of the same. Rather, they change what they produce by moving into new and more productive activities. To be more specific, countries do not jump from making coffee beans to making airplanes in one step. Countries need to gradually build capabilities and knowhow to move into an ever-expanding set of new and more sophisticated products. Sophistication cannot be achieved by simply acquiring more raw materials, capital, and labor. Instead, countries should increase their capabilities and productive knowledge base. Productive knowledge and capabilities can then be used to make products. Productive knowledge and capabilities are distributed in society, whether among individuals or in institutions. The products that a country makes are indicators of the knowledge that is embedded and has accumulated in them. Therefore, economic growth is related to the accumulation of capabilities and productive knowledge.

Countries hold productive knowledge to make goods or deliver services, while products or industries differ in how much productive knowledge is required to make them. There are products such as medical imaging devices or space shuttles that require vast amounts of knowledge. By contrast, harvesting sesame seeds requires much less. Hence, most sophisticated products will be produced by countries holding a vast amount of productive knowledge. The distributed productive knowledge in a country is termed economic complexity.

In other words, complex economies are ones that can bring vast amounts of productive knowledge together to generate a diverse mix of knowledge-intensive products. Conversely, simpler economies can only produce fewer and simpler products based on their limited productive knowledge.

One important caveat is that the concepts of productive knowledge and capabilities embedded in a country are mostly abstract and immeasurable. Hence, economic complexity measurements can only be made indirectly. Economic complexity is expressed in products that a country makes. For instance, producing an airplane would indicate that a country has gathered all the requisite knowledge to make it. On the other hand, productive knowledge does not survive long when not used in a productive process. Thus, the economic complexity of a country is revealed through the products it makes.

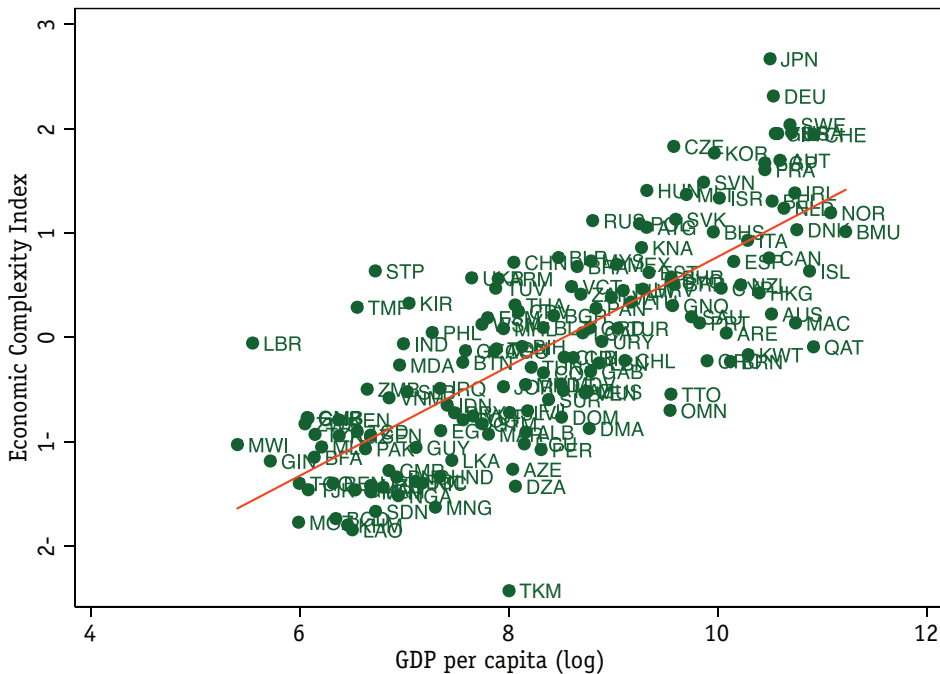
With this insight, attention can shift from countries to products. To simplify matters, it is best to use a simple analogy to the game of Word Warp. In this game, individuals try to write as many words as possible based on the letters they have in hand. Suppose that each type of productive knowledge is a letter and each product is a word composed of these letters. Like the game of Word Warp, each country holds a set of 'letters', with many copies of each letter, and tries to make words out of letters. The challenge in a game of Word Warp is identifying the number of letters in a player's hand by looking at the words that all players write. There are some letters (like the letter E) that go into many words and there are some (like Q) that go into very few. Players who have more letters can obviously write more words. So, to the first approximation, the number of different words that a player writes (i.e., diversity of products) would be indicative of the number of letters in a player's hand. This approach gives the same weight to each word. But in reality, longer words will be written by few players. Hence, by using the ubiquity of the words as a weight, it is possible to have a better measure of the length of letters in a player's hand. But the ubiquity of a word will be distorted if the word uses an uncommon letter (such as X). One can distinguish these words from others by taking into account how many different words the player can write. If a rare word is mostly written by the players who can only write a few words, then this word uses an uncommon letter. But if a rare word is mostly written by players who can also write many other words, then this is indicative of the requirement of many different letters to write this word. Hence, the ubiquity of a word can be used to gather information about the diversity of words that a player can write. In turn, it can also be used to gather information on the frequency of a word appearing by taking into account the number of players that can write it. This mechanism can be repeated as many

times as is necessary and the ultimate result is the measure of letters in each player's hand.

Similarly, more complex products will be less common because only countries that have all requisite knowledge will be able to make them. Products that require little knowledge should be more ubiquitous and vice versa. This study defines the diversification of a country as the number of different products that it can make and the ubiquity of a product as the number of countries that can make that product. Since countries with a larger variety of productive knowledge will be able to make more and more unique products, they will be more diversified. By contrast, products that require more productive knowledge will be made only in countries that have all requisite pieces, meaning they will be less ubiquitous. This implies that there should be a negative relationship between the diversification of a country and the average ubiquity of its products. Diversity and ubiquity are, respectively, crude approximations of the variety of productive knowledge available in a country, or required by a product. Both mappings are affected by the existence of rare products. To address this problem, diversity can be used to correct information carried by ubiquity and ubiquity can be used to correct information carried by diversity (See page 15 'Measuring Economic Complexity' for mathematical implementation). This 'correction' can be repeated an infinite number of times. The process converges after a few iterations and represents the quantitative measures of complexity. For countries, this is referred to as the Economic Complexity Index (ECI) and the corresponding measure for products is the Product Complexity Index (PCI). The PCI is a number unique to each product that captures how much productive knowledge the product requires. This measure is related to the concept of ubiquity. ECI is a number unique to each country that measures the amount of productive knowledge contained in a country. Countries with a high ECI have well diversified economies exporting, on average, high-PCI products.



Figure 1 Economic Complexity Index (ECI) vs GDP per capita



Note Authors' calculation based on the UN COMTRADE, CEPII BAGI database. GDP per capita figures are from the World Development Indicators (WDI) from the World Bank Database.

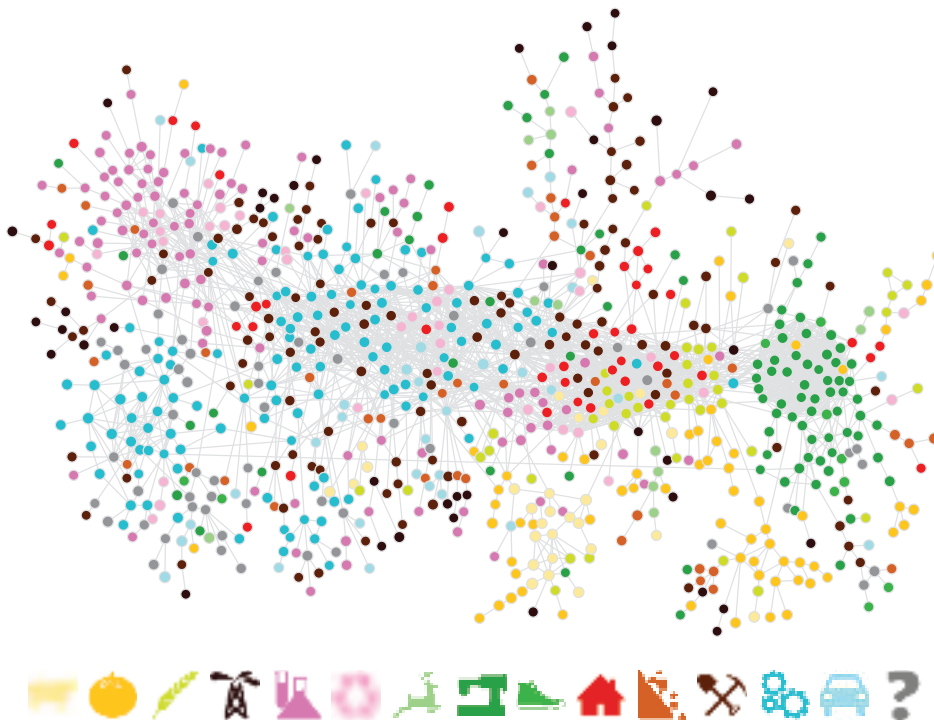
Figure 1 shows the relationship between ECI and GDP per capita for most countries in the world. Strikingly, there is a 76% correlation between ECI and GDP per capita. Furthermore, ECI is predictive of future growth. Moreover, in the Atlas of Economic Complexity, the authors also illustrate that ECI captures orthogonal information on the institutional quality of a country, the educational measures of a country, and the business environment in a country.

After quantifying the economic complexity of countries and products, it is possible to address how complexity evolves. Countries accumulate capabilities and productive knowledge at different rates and some places have advanced further in this process than others. After all, making more complex products requires accumulating new productive knowledge, which is costly to acquire and transfer. These capabilities and productive knowledge cannot be obtained just by reading books or downloading blueprints, but also need to be learned through action, which creates a major obstacle for the diffusion process. Industries require capabilities and productive knowledge to exist, whereas productive knowledge will be absent if there are no industries using it, in turn creating a 'chicken and egg' problem. For instance, there will not be any watchmakers if there is not a watch industry in a country. On the other hand, a watch industry will not be

established in a country if there are no watchmakers. Diversification requires coordination of the development of a new industry with the accumulation of missing productive knowledge that it requires. Consequently, accumulation of productive knowledge is often slow and is a path-dependent process. It is easier for countries to move into industries that share capabilities and productive knowledge with industries already present in a country, since this will require building less additional capabilities and acquiring less productive knowledge to make the leap.

The process by which diversification occurs is rife with market failure(s) and is indicative of a very rich structure that makes it predictable. The ability to make things evolve by moving from the current set of goods to others that are nearby is captured by the concept of the 'product space'. While the idea of the product space has been shown to contain useful information about the set of possibilities that each country or location offers, what the policy world is missing is a new set of tools that can be used to make sense of a country in detail and inform its policy process. The product space has helped make explicit something that economics had left out: The products that you make now condition the products you can make in the future. Interestingly, more-sophisticated products are located at the center of the product space, whereas less sophisticated products occupy a less-connected periphery. Empirically, countries move through the product space by developing goods close to those they currently produce. Hence, countries that make products in the periphery of the product space face greater development challenges since they produce goods that do not open up new development opportunities as easily as for countries with more nearby goods. This helps explain why poor countries have trouble developing more competitive exports and fail to converge on the income levels of rich countries. The core-periphery structure of the product space is manifested in its network depiction (figure 2). In this figure, each node represents products and the products are connected if the probability of being co-exported is high (see page 15 for mathematical details). The nodes are colored according the product category that they belong to.

Figure 2 The product space



Note Authors' calculation using HS4-level trade data from United Nations COMTRADE. Products are colored according to the communities that they belong according to the legend above.

In order to explain the usage of the product space, it is useful to think of the following analogy. Consider a product as a tree and the set of all products as a forest. A country is composed of a collection of firms, or 'monkeys', that live on different trees and exploit products on the trees that they occupy. The process of growth implies populating the forest by moving from a poorer part of the forest, where trees have little fruit, to "fruitier" trees that are further away. Countries that have firms in the periphery of the product space should jump long distances, since they will need to redeploy specific human, physical, and institutional capital toward goods that are different from those currently under production. Traditional growth theory assumes there is always a tree within reach, hence, the structure of the forest is unimportant. However, if a forest is heterogeneous, with some dense areas and other more-deserted ones, and if monkeys can jump only limited distances, then monkeys may be unable to move through the forest effectively. If this is the case, analogously, the structure of this space and a country's orientation within it become of great importance to the development of countries.

The product space demonstrates that a country's future exports are significantly more likely to be products that are close by in the network to products that a country was previously exporting, while it is quite

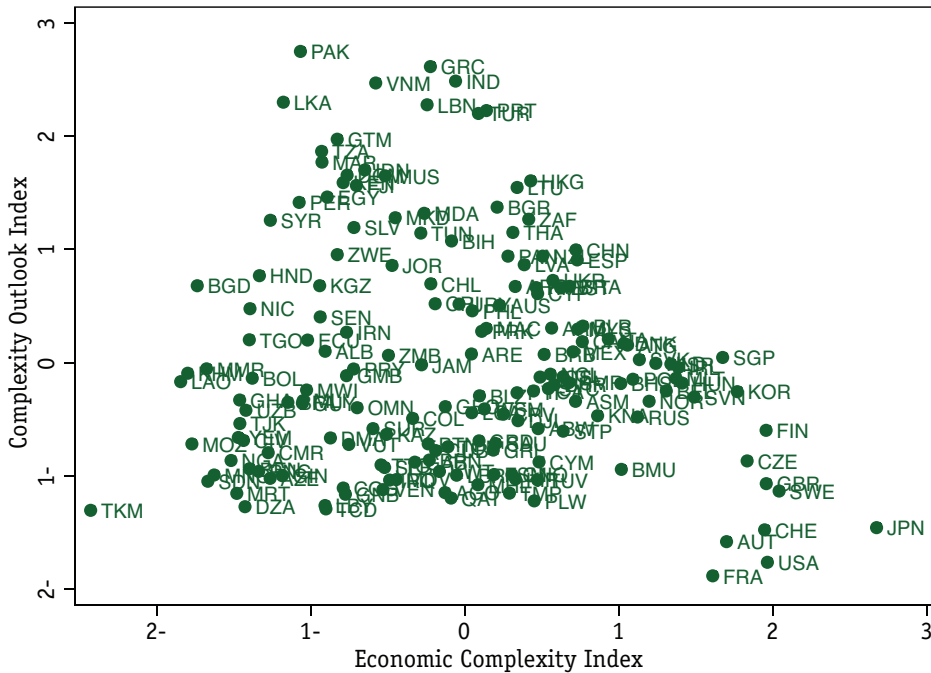
difficult for production to shift to products far away. Hence, policies to promote large jumps are more challenging. According to our findings, a lack of connectedness and the position of countries in the product space help explain difficulties faced by countries trying to converge on the income levels of rich countries, meaning they may not be able to upgrade their productive structure because high-value added products are too far away from where they stand in the product space. Yet, it is precisely these long jumps to new, more sophisticated products that generate further future possibilities for development and growth. These findings have a broad set of implications for international development, since they imply that the development of nations is path dependent and that the industrial opportunities faced by each country are deeply constrained by their current productive structure.

Proximity on the product space is a measure of relatedness between two products. However, countries have a productive base in many products. Hence, it is necessary to aggregate proximities to produce a distance measure between a country and a product. To do this, the product-centric view must be taken into account. Each product has a similarity to another product, and some are made by a country and some are not. Therefore, by adding all proximities from the products that are made by a country and dividing them by all the proximities to a given product, a measure of how similar a country's productive base is to that product is obtained. This is referred to as the 'density' measure. Density is a country-product measure. The inverse of the density is a measure of distance from a country to a product. In a recent paper, Hausmann, Hidalgo, Stock, and Yildirim showed that countries actually diversify to high-density products, which highlights the importance of the density measure for diversification.

So far, two measures of complexity (ECI and PCI) and two measures based on the product space (proximity between products and distance between country and a product) have been introduced. These measures can be combined to offer a better understanding of the evolution of countries' productive capacities. For instance, not all products are equal. Some have higher complexity and some have lower complexity. A country that is closer to higher complexity products nearby in the product space has many opportunities to increase its complexity. Therefore, the open parts of the product forest of a country are an important indicator of its future complexity. For each country, 'Complexity Outlook Index' (COI) is calculated as the PCI-weighted sum of the distance to all products that the country is not currently making. Higher COI implies larger opportunities. There is not a clear relationship between COI and ECI (figure 3) but COI levels for low- and high-complexity countries are usually low, showing that both of these types of countries have less opportunities. Middle-ECI countries, how-

ever, have different levels of COI, some being in a more advantageous position than others.

Figure 3 Economic Complexity Index (ECI) vs. Complexity Outlook Index (COI)



Note Authors' calculation based on the UN COMTRADE, CEPII BACI database.

When countries diversify, they can increase their immediate complexity by targeting higher PCI products. But another strategy is to diversify production, taking a country into more favorable positions in the product space. These products can be seen as strategic bets for countries and can be identified by quantifying the change in the Complexity Outlook Index if a product is added to the product basket of a country. The 'Complexity Outlook Gain' (COG) measure is defined as the change in COI and can be calculated for every product currently not made by a country. High COG products bring countries to the more abundant and rewarding parts of the product space and increase opportunities for future diversification.

In all, complexity concepts and measures are novel tools to reveal diversifying opportunities to trigger economic growth. This report is based solely on these measures. In the next section, these complexity measures will be formally defined.

## b Which country makes what products?

### 1 Data source

The primary data used to map the product space is international trade data. Here, two related versions are employed: Data at the Harmonized System four-digit classification level (HS4 data) compiled by the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) and data at the Standard International Trade Classification four-digit level (SITC4) published by the United Nations. Both provide information on exports from more than 200 countries at the product level. The advantage of HS4 data is that it is more granular (disaggregated into 1,240 different products) than the SITC data (774 different products). However, the SITC data covers a longer time period (1964-2010).

There are a number of drawbacks to the data. First, the data covers exports of goods and not aggregate production. Thus, it is not able to capture the productive knowledge used in the production of non-traded goods or services. In the case of Rwanda, goods exports represent only 5% of GDP. Thus, the analysis of exports will not be able to fully capture the knowledge that exists in Rwanda to the extent that it is expressed only in the production of non-traded goods or services. However, since the aim of this example is to identify opportunities for the diversification of Rwanda's exports, looking at export data is a natural choice. Moreover, the fact that certain locally produced goods are not exported suggests that a country may not yet be very efficient or competitive at producing them.

Second, countries may also export products they do not make. While many countries' customs offices clean data from re-exports, not all do so with a high degree of accuracy. To circumvent this issue in this study's analysis, a country is required to have a significant presence in a product to assume that it makes it.

Finally, the data includes only goods, not services. This is an important drawback, as services are becoming a rising share of international trade and service exports are expected to make a significant contribution to overall export growth. Unfortunately, there are no international datasets on services comparable to the one that exists for goods and since existing services data is not sufficiently granular, it is not yet possible to integrate it into the product space. As a result, the discussion in the remainder of this paper will be largely limited to the exploration of goods exports.

### 2 Identification of simultaneous production of exports using RCA

Given that this study uses export data, it is necessary to distinguish between re-exports and products that a country is actually making. According to trade theory, countries will make products that they have a comparative advantage in. However, it is impossible to measure

absolute comparative advantage without observing the efficiencies of countries in making products. Therefore, Revealed Comparative Advantage (RCA) is employed in this study, which is an index used to calculate the relative advantage a country has in the export of a certain good. Balassa's definition of RCA indicates that a country has an RCA larger than one in a product if it exports more than its 'fair share,' or a share that is equal to or greater than the share of total world trade that the product represents.

For example, in 2010, soybeans represented 0.35% of world trade with exports of \$42 billion. Of this total, Brazil exported nearly \$11 billion. Since Brazil's total exports for that year were \$140 billion, soybeans accounted for 7.8% of Brazil's exports. Because  $7.8/0.35 = 22$ , Brazil exports 22 times its "fair share" of soybean exports, meaning it is possible to infer that Brazil has a high revealed comparative advantage in soybeans.

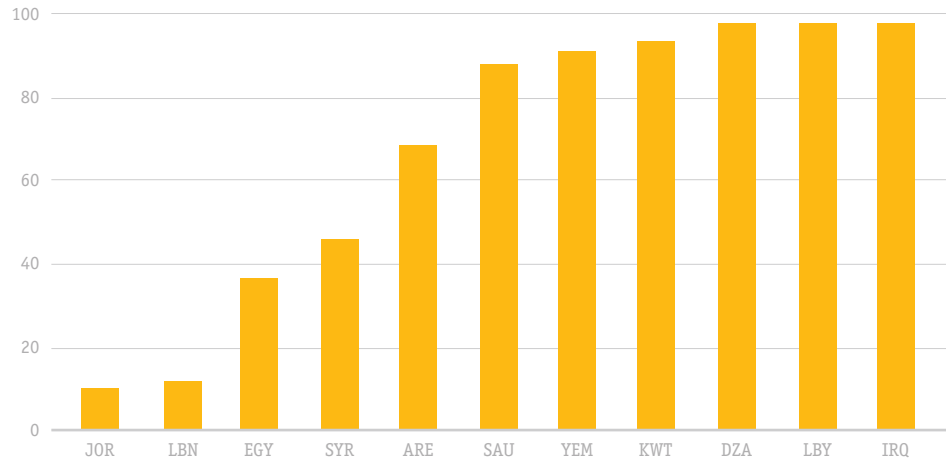
Formally, if  $\text{exp}_{cp}$  represents the exports of product  $p$  by country  $c$ , the RCA that country  $c$  has in product  $p$  can be expressed as

$$\text{RCA}_{cp} = \frac{\text{exp}_{cp} / \sum_c \text{exp}_{cp}}{\sum_p \text{exp}_{cp} / \sum_c \sum_p \text{exp}_{cp}} \quad \text{Equation 1}$$

RCA, however, is sensitive to price changes. Moreover, if a country's export basket is predominated by several products, then it becomes harder for any other product to show higher RCA values even though the country makes it. For instance, Saudi Arabia is an oil exporter. Since oil dominates Saudi Arabia's exports, it is unlikely and difficult for any other product to have large  $\text{exp}_{cp} / \sum_p \text{exp}_{cp}$ .

In the 2011 and 2013 atlas, a country was categorized as producing a given product if the share of that product in the country's export basket was higher than the share of that product in world trade exports. But with that method, a sharp increase in the price of a commodity affects the presence of other products. For instance, a rise in the price of oil would increase the total exports of a country like Saudi Arabia, which also makes dates. Even if Saudi Arabia continues to make the same amount of dates, due to fluctuations in the price of oil, dates may end up being absent in Saudi Arabia's complexity calculations. This is particularly important for countries in this report as Arab countries are intensive exporters of natural resources, mostly fuel. Figure 4 shows the impotence of natural resources in countries' export baskets, the presence of which might mask the productive base of the country.

Figure 4 Share of natural resources in total export basket



Note Authors' calculation using the World Development Indicators (WDI) from the World Bank Database. Share of Natural Resources correspond to the sum of fuel, ores, and metals exports as a share of total merchandise exports, averaged over 2009 and 2012. For United Arab Emirates, data is taken from UN COMTRADE.

This report seeks to improve the methodology to better capture the capabilities and economic complexity of countries that are rich in natural resources. Since this study primarily focuses on oil exporting countries, it was considered advantageous to calculate RCAs to specially address the needs of these countries. Instead of using a pure RCA measure, this study uses a two-stage RCA, where in the second stage products that really dominate the country's exports are removed. Briefly, here is the procedure that is employed:

- Calculate  $RCA^1$  (Regular Balassa's RCA)
- Identify the products which have high  $RCA^1$  (e.g.,  $RCA^1 > 5$ )
- Take these products out of a country export basket momentarily
- Recalculate  $RCA^2$  (with the updated export basket)
- Define RCA as  $RCA^1$  for the products excluded in step 3 and  $RCA^2$  for all other products

This definition of RCA is applied to all countries (oil and non-oil producers). This procedure should affect countries that have extremely concentrated export baskets and assist in revealing capacities of countries that are oil producers. This measure can be used to construct a matrix that connects each country to the products that it makes. This matrix is the starting point of all other complexity calculations. Entries in the matrix are 1 if country  $c$  exports product  $p$  with an RCA greater than 1, and 0 otherwise. Formally,  $M_{cp}$  matrix is defined:



$$M_{cp} = \begin{cases} 1 & \text{RCA}_{cp} \geq 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{Equation 2}$$

$M_{cp}$  is the matrix summarizing which country makes what, and is used to construct the product space and measures of economic complexity for countries and products.

## c Complexity variables

### 1 Measuring economic complexity

The complexity of a product is measured by using the method of reflections (Hidalgo and Hausmann PNAS 2009).  $M_{cp}$  is employed as a matrix that is 1 if country  $c$  produces product  $p$ , and 0 otherwise. Diversity and ubiquity can be measured simply by summing over the rows or columns of that matrix. Formally:

$$\text{Diversity} = k_{c,0} = \sum_p M_{cp} \quad \text{Equation 3}$$

$$\text{Ubiquity} = k_{p,0} = \sum_c M_{cp} \quad \text{Equation 4}$$

To generate a more accurate measure of the number of capabilities available in a country, or required by a product, it is necessary to correct the information that diversity and ubiquity carry by using each one to correct the other. For countries, this requires calculating the average ubiquity of products that it exports, the average diversity of countries that make those products, and so forth. For products, this requires calculating the average diversity of countries that make them and the average ubiquity of the other products that these countries make. This can be expressed by the recursion relation:

$$k_{c,N} = \frac{1}{k_{c,N}} \sum_p M_{cp} k_{p,N-1} \quad \text{Equation 5}$$

$$k_{p,N} = \frac{1}{k_{p,N}} \sum_c M_{cp} k_{c,N-1} \quad \text{Equation 6}$$

Then Equation 6 is replaced with Equation 5 to obtain:

$$\begin{aligned}
 k_{c,N} &= \frac{1}{k_{c,o}} \sum_p M_{cp} \frac{1}{k_{p,o}} \sum_{c'} M_{c'p} k_{c',N-2} \\
 &= \sum_{c'} k_{c',N-2} \sum_p \frac{M_{cp} M_{c'p}}{k_{c,o} k_{p,o}} \\
 &= \sum_{c'} k_{c',N-2} \tilde{M}_{cc'} k_{c',N-2} \quad \text{where } \tilde{M}_{cc'} = \sum_p \frac{M_{cp} M_{c'p}}{K_{c,o} K_{p,o}} \quad \text{Equation 7}
 \end{aligned}$$

It is interesting to identify cases where the distribution remains constant. This is satisfied when  $k_{c,N} \propto k_{c,N-2}$ . This relationship is satisfied if  $k_{c,N}$  is an eigenvector of  $\tilde{M}_{cc'}$ . The largest eigenvector of the  $\tilde{M}_{cc'}$  matrix is a vector of ones and does not capture variety among countries. This study looks, instead, for the eigenvector associated with the second-largest eigenvalue. This is the eigenvector that captures the largest amount of variance in the system and is, thus, the measure employed in this study of economic complexity. Thus, ECI and PCI are defined as:

ECI = eigenvector associated with the second-largest eigenvalue  
of  $\tilde{M}_{cc'}$   
 PCI = eigenvector associated with the second-largest eigenvalue  
of  $\tilde{M}_{pp'}$

The difference between the method used in Hidalgo and Hausmann (2009) and the Atlas of Economic Complexity is that in the latter the method of reflections is implemented by solving the eigenvector associated with countries and products. Hidalgo and Hausmann used an iterative process that is more time consuming compared to the eigenvector solution.

## 2 Measuring proximity - product relatedness

Empirically, countries move through the product space by developing goods close to those they currently produce. But countries do not make just one product: They make a certain number of them. If two products require very similar productive knowledge, they will either be simultaneously present or absent in most countries. For instance, if producing olives require knowledge similar to that required for tangerines, but different from that required by computers, then for most countries producing olives, tangerine production will also most

likely be observed, but the same will not apply to computers. So the probability that pairs of products are co-produced by countries carries information about how similar these products are. If two goods require roughly the same knowledge, they will be produced by the same countries. Hence, it is possible to identify a similarity measure between products based on the probability of the co-appearance of products.

Data can be used to calculate the probability of co-exporting every possible pair of products. The (symmetric) proximity between two products is defined as:

$$\phi_{p,p'} = \frac{\sum_c M_{cp} M_{cp'}}{\max \{ \sum_c M_{cp}, \sum_c M_{cp'} \}} \quad \text{Equation 8}$$

### 3 Measuring how close a country is to a product (Density)

Proximity measures the similarity between a pair of products, therefore another measure is needed to quantify the distance between products that a country makes and each of the products that it does not. This measure is referred to as distance and is defined as the sum of proximities connecting a new good  $p$  to all the products that country  $c$  is not currently exporting. This distance is normalized by dividing it by the sum of proximities between all products and product  $p$ . In other words, distance is the weighted proportion of products connected to good  $p$  that country  $c$  is not exporting. The weights are given by proximities. If country  $c$  exports most of the goods connected to product  $p$ , then the distance will be short, close to zero. But, if country  $c$  only exports a small proportion of the products that are related to product  $p$ , then the distance will be large (close to 1). How close country  $c$  is to product  $p$  is determined by measuring the presence of a country in the neighborhood of a product. This measure is referred to as density and it is formally defined as:

$$d_{c,p} = \frac{\sum_c M_{cp'} \phi_{p,p'}}{\sum_{p'} \phi_{p,p'}} \quad \text{Equation 9}$$

### 4 Measuring how well a country is located in the product space

Distance gives an idea of how far each product is from one another given a country's current mix of exports. Yet, it would be useful to have a holistic measure of the opportunities presented by a country's position in the product space. Countries that make products which are relatively complex, given their current level of income, tend to grow faster. Hence, it makes sense to include not only the distance to products, but also their complexity. Some countries may be located

near few, poorly connected, and relatively simple products, while others may have a rich unexploited neighborhood of highly connected or complex products. This means that countries differ not just in what they make but in what their opportunities are. This can be understood as the value of the option to move into other products. Hence, to quantify the ‘opportunity value’ of a country’s unexploited prospects, the level of complexity of products that are not being made are weighted by how close these products are to the country’s current export suite.

The density measure can be aggregated to assess how well-positioned a product is in the product space. This measure is referred to as the complexity outlook of a country. Formally:

$$\text{complexity outlook}_c = \sum_{p'} (d_{c,p'} \times (1 - M_{cp'}) \times PCI_{p'}) \quad \text{Equation 10}$$

where PCI is the Product Complexity Index of product  $p$ . The term ensures that only the products that a country is not currently producing are counted. Higher opportunity value implies being in the vicinity of more products and/or of products that are more complex.

## 5 Measuring how the addition of a product would affect a country’s outlook

Opportunity value can be used to calculate the potential benefit to a country if it were to make a particular new product. This is called the ‘opportunity gain’ that country  $c$  would obtain from making product  $p$ . This is calculated as the change in opportunity value that would come as a consequence of developing product  $p$ . Opportunity gain quantifies the contribution of a new product in terms of opening up doors to more and more complex products. Formally, opportunity gain can be written as:

$$\text{complexity outlook gain}_{cp} = \left[ \frac{\sum_{p'} \varphi_{p,p'} (1 - M_{cp'}) PCI_{p'}}{\sum_{p''} \varphi_{p'',p'}} \right] \quad \text{Equation 11}$$

## 6 Complexity variables and economic growth

To analyze the impact of the Economic Complexity Index (ECI) and Complexity Outlook Index (COI) on future economic growth, four regressions are estimated where the dependent variable is the annualized growth rate of GDP per capita for the periods 1978-1988, 1988-1998, and 1998-2008. This study makes use of the SITC4 Rev2 trade dataset that covers data from the 1970s and 1980s, which is not

possible using the CEPII HS4 dataset used for the other sections of the report. The other variables used in the regression, like GDP and natural resources, come from World Development Indicators. Liberia was excluded for the 1988 sample and Zimbabwe for the 1998 sample because they were extreme outliers.

In the first of these equations neither ECI nor COI is included and the study uses only two control variables: The logarithm of the initial level of GDP per capita in each period and the increase in natural resource exports is in constant dollars as a share of initial GDP. The first variable captures the idea that, other things equal, poorer countries should grow faster than rich countries and catch up. This is known in economic literature as convergence. The second control variable captures the effect on growth of increases in income that come from natural resource exports, which complexity does not explain. In addition, a dummy variable is included for each decade, capturing any common factor affecting all countries during that decade, such as a global boom or a widespread financial crisis. Taken together, these variables account for 29% of the variance in countries' growth rates. This is shown in the first column of table 1.

**Table 1 Growth per capita**

Variables	Annualized growth in GDP pc (by decade) (1978-1988, 1988-1998, 1998-2008)			
	(1)	(2)	(3)	(4)
Initial income per capita, log	-0.001 (0.001)	-0.011*** (0.001)	-0.006*** (0.001)	-0.011*** (0.001)
Increase in net natural resource exports in constant dollars - (as a share of initial GDP)	0.059*** (0.012)	0.065*** (0.009)	0.065*** (0.010)	0.067*** (0.009)
Economic Complexity Index		0.019*** (0.002)		0.014*** (0.002)
Opportunity Value Index			0.012*** (0.002)	
Constant	0.023*** (0.007)	0.097*** (0.010)	0.058*** (0.009)	0.095*** (0.010)
Observations	301	301	301	301
R-squared	0.291	0.472	0.436	0.498
Year FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses \*\*\* p< 0.01, \*\* p< 0.05, \* p< 0.1

In addition to initial income and growth in natural resource exports, the second regression includes the effect of the value of ECI at the beginning of the period. The second column of table 1 shows that ECI is strongly associated with future economic growth. This variable is highly significant both economically and statistically. Its inclusion increases the explanatory power of the equation in column 1 by 66%. A 1-standard deviation increase in ECI is estimated to accelerate annual growth by 1.9%.

In column 3, the Complexity Outlook Index (COI) and the two control variables of column 1 are introduced. It also shows that COI is highly significant, both economically and statistically, raising the explanatory power of the equation by 52% relative to column 1. A 1-standard deviation improvement in COI is associated with a 1.2% increase in growth of GDP per capita.

In column 4, both ECI and COI are introduced into the growth equation. Both variables remain highly significant and the equation as a whole explains half of the variance of ten-year growth over three decades in this report's sample of over 100 countries. The difference between column 4 and column 1 indicates that the ECI and COI jointly increase the regression's  $R^2$  by 21 percentage points.

#### d Measuring trade health and potential

Countries have varying potential for trading with other countries. If a country is good at making a product and if another country is a purchaser of that product, the natural assumption would be that the latter country would import this good from the former country. In reality, there might be inefficiencies in this process. In this section, the study attempts to identify systematic mismatches between countries by comparing the observed trade to the expected trade based on the export share and import size of these countries.

Formally,  $\text{exp}_{c \rightarrow c', p}$  is denoted as the exports of country  $c$  to country  $c'$  in product  $p$ . The size of all exports between countries  $c$  and  $c'$  can be defined as:

$$\text{exp}_{c \rightarrow c'} \equiv \sum_p \text{exp}_{c \rightarrow c', p}$$

Similarly, all exports of product  $p$  by country  $c$ , all imports of country  $c'$  in product  $p$ , and all exports in the world in product  $p$  can be defined by:

$$\text{exp}_{cp} \equiv \sum_{c'} \text{exp}_{c \rightarrow c', p}$$

$$\text{imp}_{c'p} \equiv \sum_c \text{exp}_{c \rightarrow c', p}$$

$$\text{exp}_p \equiv \sum_c \sum_{c'} \text{exp}_{c \rightarrow c', p}$$

respectively. Before calculating the expected trade between countries  $c$  and  $c'$  in product  $p$ , it is also necessary to take into account that some products can only be traded locally (e.g., cement will only travel short distances). So, this should not include products whose worldwide movements are local. The average distance that the product moved in the world can be defined by:

$$\text{dist}_p \equiv \frac{\sum_c \sum_{c'} \text{dist}_{cc'} \text{exp}_{c \rightarrow c', p}}{\sum_c \sum_{c'} \text{exp}_{c \rightarrow c', p}}$$

where  $\text{dist}_{cc'}$  is the distance between countries  $c$  and  $c'$ . With these definitions, products will only be included in this study's calculation with the following properties to measure the trade health between countries  $c$  and  $c'$ :

- 1 *Export Condition*: Country  $c$  should have a comparative advantage in product  $p$ . Formally:

$$\text{RCA}_{cp} > 1$$

where RCA is defined as in equation 2.2.1.

- 2 *Import Condition*: Product  $p$  should be an important part of the imports of country  $c'$ . Akin to the definition of RCA, the following is used:

$$\frac{\text{imp}_{c'p}/\text{exp}_p}{\sum_p \text{imp}_{c'p}/\sum_p \text{exp}_p} > 0.5$$

- 3 *Distance Condition*: The distance between countries should be reasonable for the product to be shipped. To this end, two types of products are defined:
  - **Global products**: Products for which it is assumed that the distance between countries is not an issue. These products are identified as those

in the upper half of the distance distribution (i.e.,  $\text{dist}_p > 4,600\text{km}$ , where 4,600km is the median of the distance distribution of products). All these products are included in our set of products that are used to calculate trade health.

- **Local products:** These are products for which it is assumed that the distance between countries should be taken into account. These products are identified as those in the lower half of the distance distribution. A local product is included in the calculations of trade health if:

$$\text{dist}_p > \text{dist}_{cc'}/2$$

The set  $P_{cc'}$  is defined as the set of products that satisfy all three conditions. For all the products in  $P_{cc'}$ , the expected trade between these two countries is calculated as:

$$\widehat{\text{exp}}_{c \rightarrow c', p} \equiv \frac{\text{exp}_{cp} \text{imp}_{c'p}}{\text{exp}_p}$$

This equation implies that country  $c$  should at least preserve its shares in product  $p$  in the imports of country  $c'$ . Based on this expected value, trade health and potential are defined as:

$$\text{trade health}_{cc'} \equiv \text{median}_{p \in P_{cc'}} \frac{\text{exp}_{c \rightarrow cc', p}}{\widehat{\text{exp}}_{c \rightarrow cc', p}}$$

$$\text{potential}_{cc'} \equiv \sum_{\substack{p \in P_{cc'} \\ \widehat{\text{exp}}_{c \rightarrow c', p} > \text{exp}_{c \rightarrow c', p}}} (\text{exp}_{c \rightarrow c', p} - \text{exp}_{c \rightarrow c', p})$$

Hence, the trade health variable summarizes how healthy the trade relationship between countries  $c$  and  $c'$  is based on the observed trade pattern; whereas the potential measure quantifies how much more trade would occur between countries  $c$  and  $c'$  if the levels of trade increase to expected levels for underperforming products.



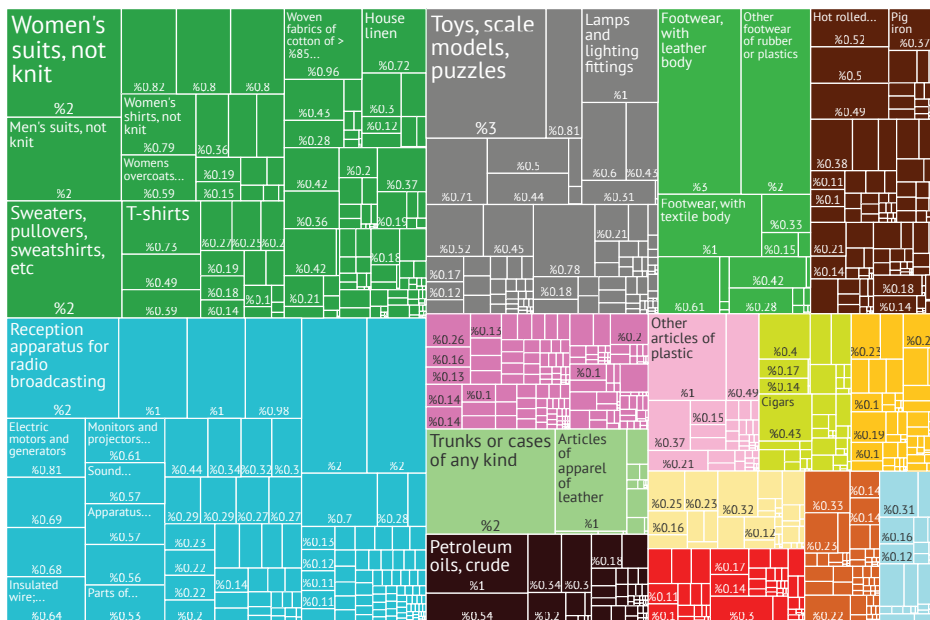
## II China's Economic Transformation

Over the last few decades, China has exhibited an incredibly fast growth rate. From 1992 to 2012, China's GDP per capita increased at a stunning average of 9.5% per year. What is the source of this success? The answer can be found in China's ability to act strategically and seize opportunities by diversifying their production throughout the product space. China has not only increased its output in industries they were already present in, but has been adding new industries with a higher value added to their production. Evidence of this is the fact that China has changed its export basket dramatically.

Figure 5 shows the composition of China's exports in 1995 and 2012. Exports increased from \$210 billion in 1995 to over \$2.1 trillion in 2012. Even though textiles—once China's most important export industry—has expanded the volume of exports, machinery, and electronics have taken over as the most important industries, accounting for more than 50% of China's total exports.

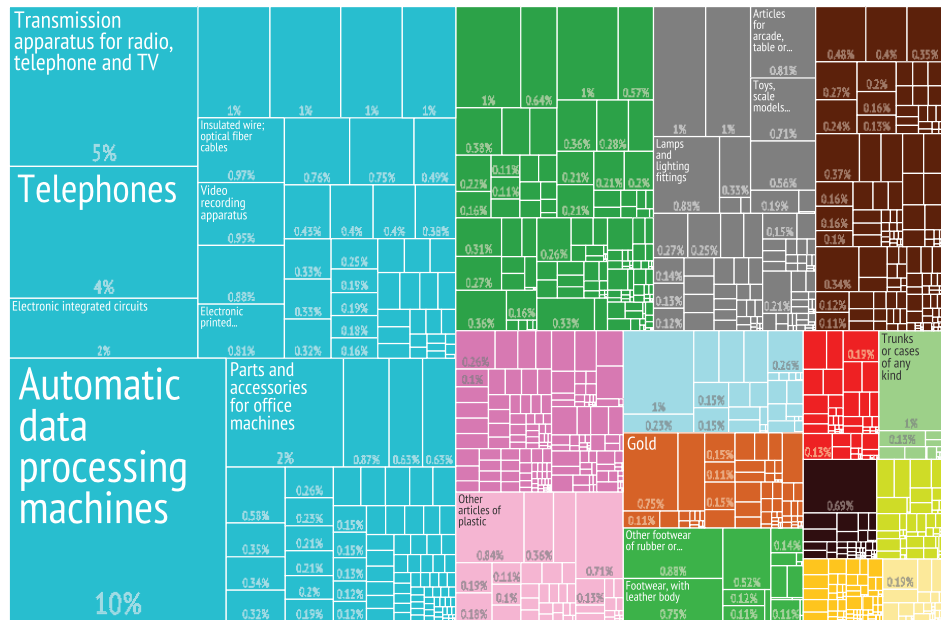
Figure 5 China's Exports

a 1995



\$210B

b 2012



\$211T

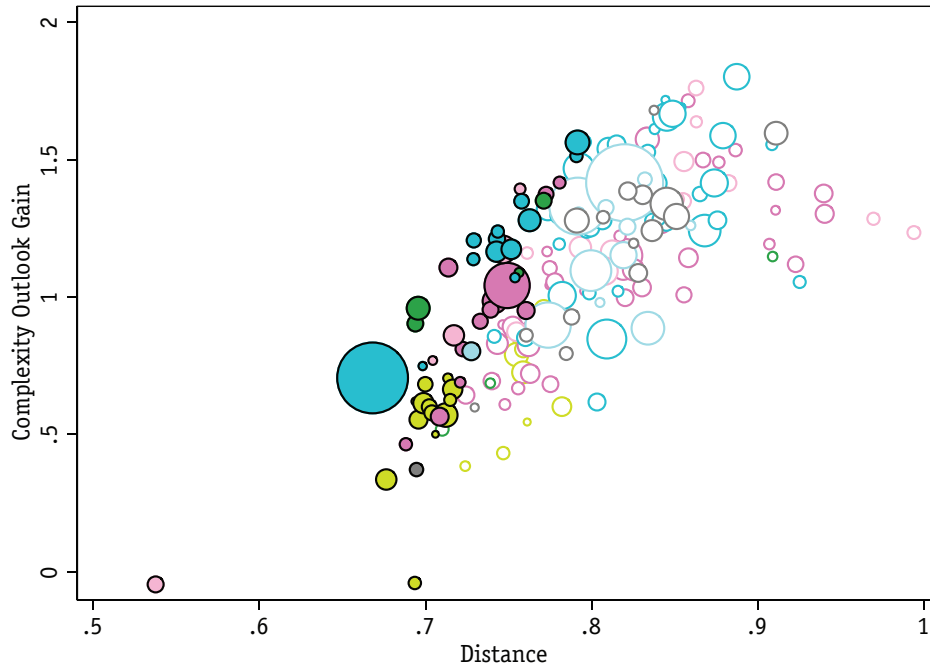
Note Authors' calculation using HS4-level trade data from United Nations COMTRADE. Products are colored according to the communities that they belong according to the above legend.



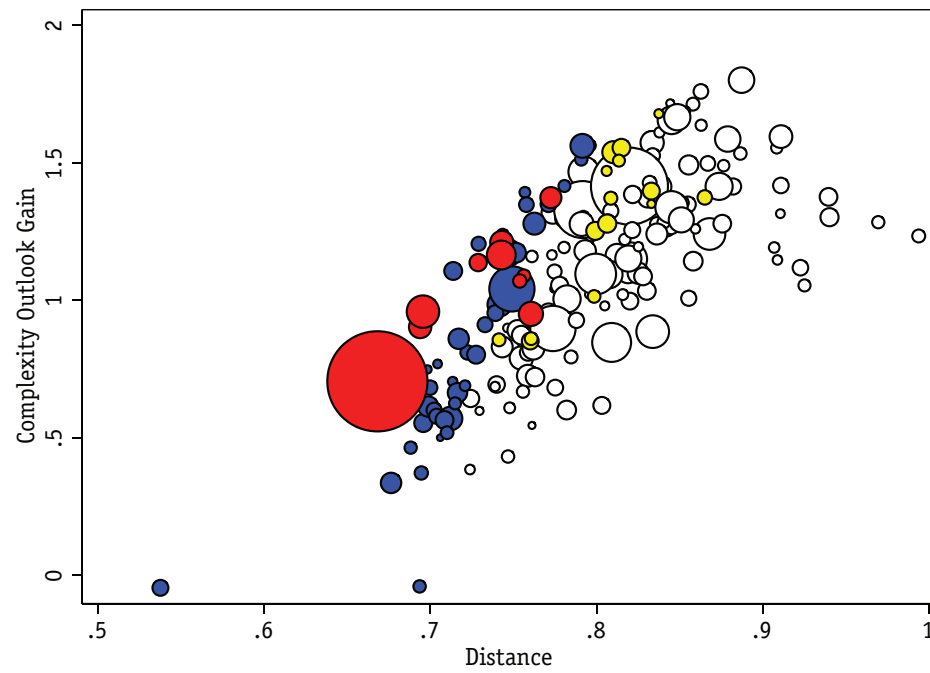
- What are the products our method would have suggested for China in the year 2000? Did China seize those opportunities?
- China became competitive by 2010 in a significant number of the products identified by our method in year 2000. Figures 6b and 6d show in red those products that would have been suggested for targeting and China did achieve  $RCA > 1$  by 2010. In blue products are shown which were suggested to have a fair chance of becoming competitive, but were not seized. In yellow, products are shown which China achieved competitiveness in, in world markets ( $RCA > 1$ ), but were not on the target list. Not surprisingly, most of these later products were nearby the target list of products.
- Because it has developed a presence in many areas of the product space, especially in products that are in a central position of the product space or connected to many other products, it is easier for China to jump to many more products than any other country. According to these calculations, China became more competitive in more industries between 2000 and 2010 than any other country.

Figure 6 Strategic bets for China in year 2000

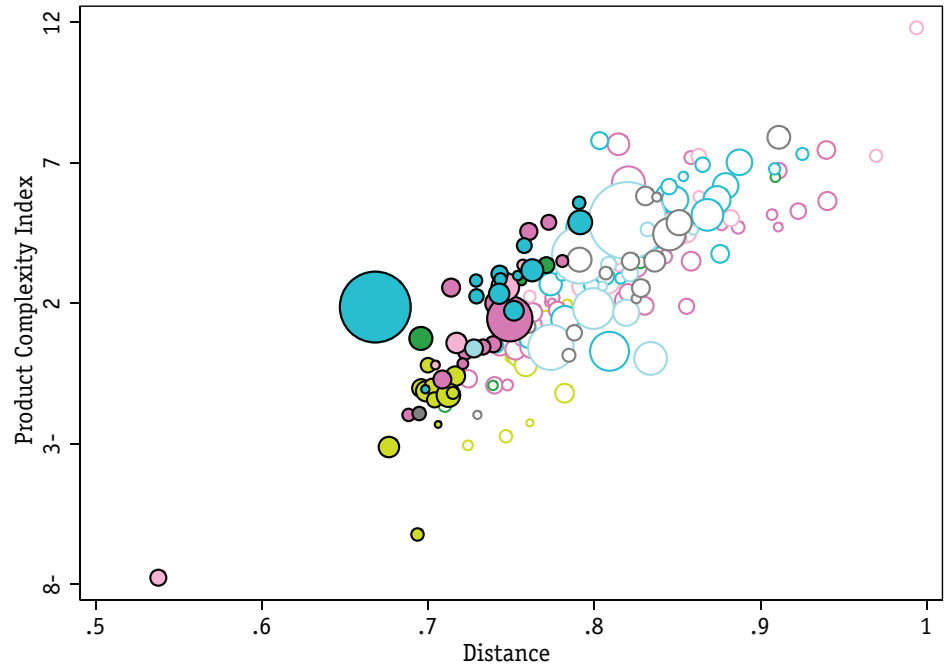
a Opportunity Gain Index



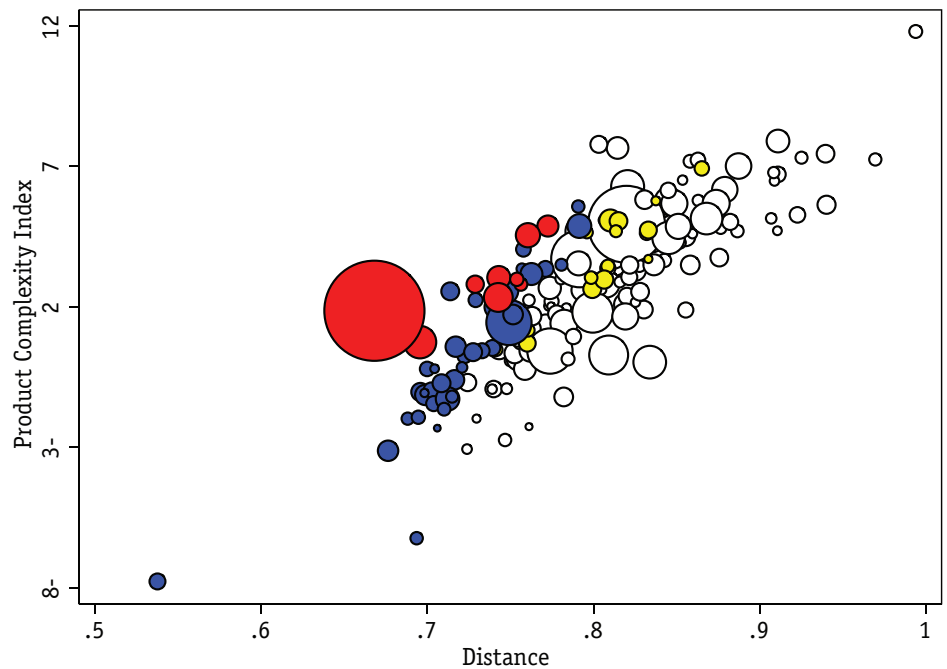
b Opportunity Gain Index



**c** Product Complexity Index



**d** Product Complexity Index



Note Authors' calculation using HS4-level trade data from United Nations COMTRADE. Node size is proportional to world trade. In (a) and (c), red nodes are conquered by China and were also in our target list, blue nodes are not conquered by China and were in our target list. Finally, yellow nodes are conquered but were not in the target list. The nodes are colored according to the communities that they belong to in (b) and (d).

### III Atlas Complexity Rankings and Results in this Report

This report uses the methodology presented in the Atlas of Economic Complexity 2013. There are, however, some differences:

- The atlas and this report use different classifications of international trade data. The atlas uses the Standard International Trade Classification at the four-digit level, revision 2 (SITC4rev2), restricting the number of products to 773 countries and 128 products. This choice was made to allow for tracking the evolution of economic complexity from 1962 to 2010. The SITC4rev2 classification has been the same since 1984. Thus, this classification does not capture improvements in technology and the introduction of new products. Moreover, many customs authorities do not report trade flows using this classification. To provide a better depiction of a country's export basket and its economic complexity, this study uses the Harmonized System Classification using the four-digit level of disaggregation (HS4). The HS4 classification has 1,242 products, but this study excludes products that accounted for less than 2% of world trade, or when a product description suggested that it was waste or a residual. The final list contains 640 products.
- Iraq was excluded from the final list of countries in the atlas because it was a country that had experienced a recent war. In this report, Iraq and other countries that have experienced war or internal turmoil have been included. Hence, the ranking of the atlas was done over a total of 128 countries, while this report uses 129.
- This report modifies the way competitiveness in international markets is measured. In the atlas, a country is recognized as competitive if Balassa's RCA is equal or larger than one. As explained in detail earlier, this is problematic for natural resource-rich countries since large exports of those products could hide a country's competitive presence products other than natural resources. Since many of the countries in this report are intensive exporters of oil, it was decided that the calculation of the RCA be modified.

Table 2 compares the Economic Complexity Index and the ranking as presented in the atlas, for year 2010, and the ECI and ranking using the newest data available for years 2010 and 2012. Most changes are positive, save for countries affected by civil wars.

**Table 2 Comparison between Economic Complexity Index (ECI) of the atlas and this report**

Country	Atlas - 2010		This report - 2010		This report - 2012		Change
	ECI	Ranking	ECI	Ranking	ECI	Ranking	
Algeria	-0.71	95	-0.09	64	-0.83	102	-
Egypt	-0.10	67	-0.63	93	-0.57	87	+
Iraq	n.a.	n.a.	-1.07	110	-1.14	111	-
Jordan	0.36	45	-0.22	70	-0.19	66	+
Kuwait	-0.03	64	-0.02	60	0.21	51	+
Lebanon	0.33	47	0.05	55	0.04	56	-
Libya	-1.24	116	-0.34	78	-1.64	127	-
Saudi Arabia	0.11	59	0.21	51	0.38	45	+
Syria	-0.27	71	-0.80	99	-0.60	88	+
Tunisia	0.24	53	-0.37	80	-0.19	67	+
United Arab Emirates	0.33	48	0.27	48	-0.03	59	-
Yemen	-1.53	121	-1.21	115	-0.73	96	+

Note Authors' calculation using HS4-level trade data from United Nations COMTRADE.

# LCPS SERIES

