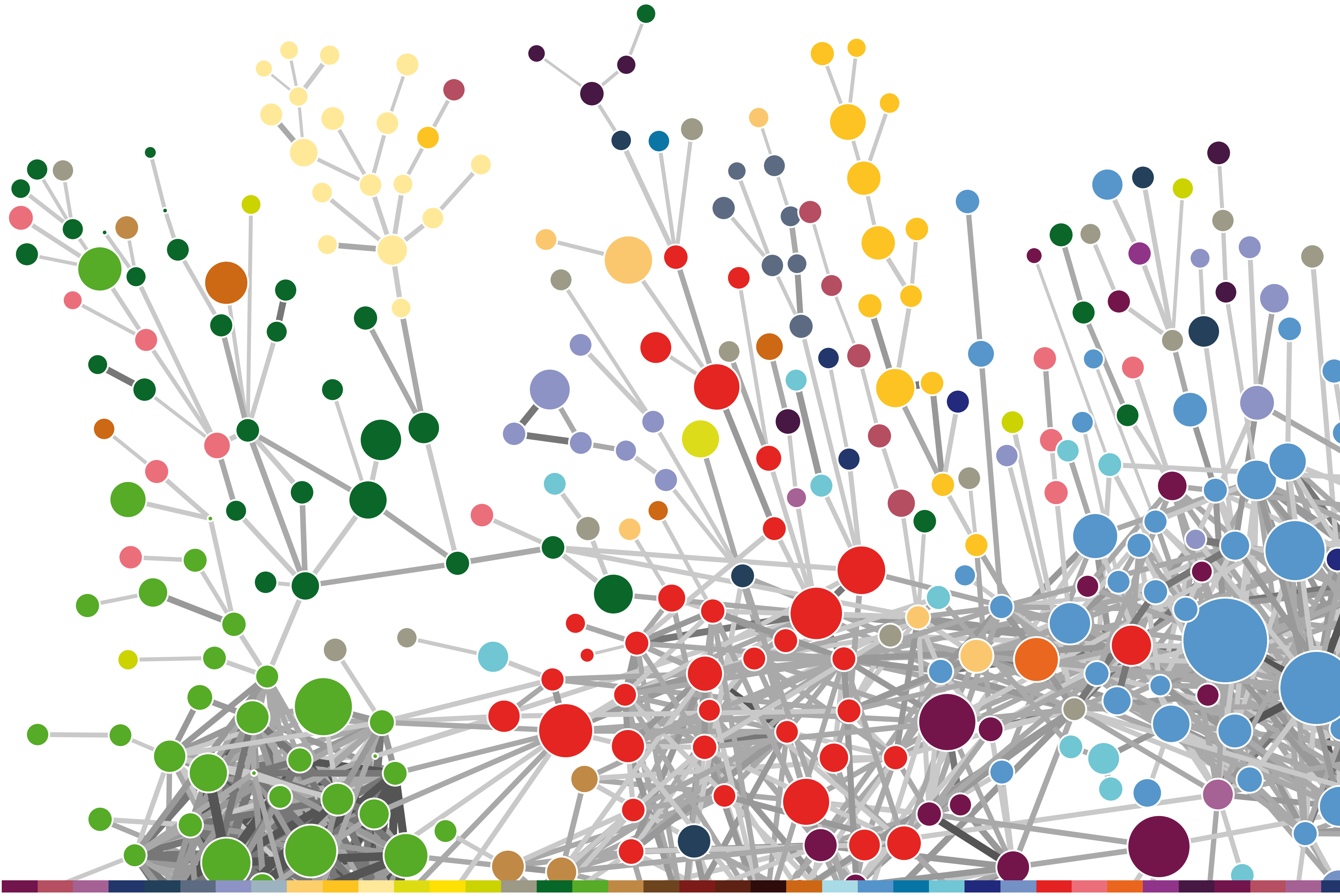
THE ATLAS OF ECONOMIC COMPLEXITY

MAPPING PATHS TO PROSPERITY

Hausmann, Hidalgo et al.



Center for International Development at Harvard University

HARVARD Kennedy School
JOHN F. KENNEDY SCHOOL OF GOVERNMENT

Macro Connections
MIT Media Lab

THE ATLAS OF ECONOMIC COMPLEXITY

MAPPING PATHS TO PROSPERITY

AUTHORS:

Ricardo Hausmann | César A. Hidalgo | Sebastián Bustos | Michele Coscia
Sarah Chung | Juan Jimenez | Alexander Simoes | Muhammed A. Yıldırım

ACKNOWLEDGMENTS

The research on which this Atlas is based began around 2006 with the idea of the product space. In the original paper published in *Science* in 2007, we collaborated with Albert-Laszlo Barabasi and Bailey Klinger. The view of economic development of countries as a process of discovering which products a country can master, a process we called self-discovery, came from joint work with Dani Rodrik and later also with Jason Hwang. We explored different implications of the basic approach in papers with Dany Bahar, Bailey Klinger, Robert Lawrence, Francisco Rodriguez, Dani Rodrik, Charles Sabel, Rodrigo Wagner and Andrés Zahler. Throughout, we received significant feedback and advice from Lant Pritchett, Andrés Velasco and Adrian Wood.

We want to thank the dedicated team that runs Harvard's Center for International Development (CID) for helping bring the Atlas to life: Marcela Escobari, Jennifer Gala, Irene Gandara Jones, Aimee Fox, Adriana Hoyos, Andrea Carranza, Anne Morriss and Catalina Prieto. We are also indebted to the NeCSys team at the MIT Media Lab and to Sandy Sener. We thank the leadership at Harvard Kennedy School and the MIT Media Lab who were early enthusiasts of our work.

The editorial design of this book was produced by DRAFT. We would like to especially acknowledge the contributions of Francisca Barros and Beltrán García.

ISBN-10: 0615546625
ISBN-13: 9780615546629

THE ATLAS OF ECONOMIC COMPLEXITY

MAPPING PATHS TO PROSPERITY

| *Ricardo Hausmann* | *César A. Hidalgo* | *Sebastián Bustos* | *Michele Coscia* |
| *Sarah Chung* | *Juan Jimenez* | *Alexander Simoes* | *Muhammed A. Yıldırım* |

PART 1 WHAT, WHY AND HOW?

SECTION I

What Do We Mean by *Economic Complexity*?

What are things made out of? One way of describing the economic world is to say that things are made with machines, raw materials and labor. Another way is to emphasize that products are made with knowledge. Consider toothpaste. Is toothpaste just some paste in a tube? Or do the paste and the tube allow us to access knowledge about the properties of sodium fluoride on teeth and about how to achieve its synthesis? The true value of a tube of toothpaste, in other words, is that it manifests knowledge about the chemicals that facilitate brushing, and that kill the germs that cause bad breath, cavities and gum disease.

When we think of products in these terms, markets take on a different meaning. Markets allow us to access the vast amounts of knowledge that are scattered among the people of the world. Toothpaste embeds our knowledge about the chemicals that prevent tooth decay, just like cars embody our knowledge of mechanical engineering, metallurgy, electronics and design. Computers package knowledge about information theory, electronics, plastics and graphics, whereas apples embody thousands of years of plant domestication as well as knowledge about logistics, refrigeration, pest control, food safety and the preservation of fresh produce.

Products are vehicles for knowledge, but embedding knowledge in products requires people who possess a working understanding of that knowledge. Most of us can be ignorant about how to synthesize sodium fluoride because we can rely on the few people who know how to create this atomic cocktail, and who together with their colleagues at the toothpaste factory, can deposit it into a product that we can use.

We owe to Adam Smith the idea that the division of labor is the secret of the wealth of nations. In a modern reinterpretation of this idea, the division of labor is what allows us to access a quantity of knowledge that none of us would be able to hold individually. We rely on dentists, plumbers, lawyers, meteorologists and car mechanics to sustain our standard of living, because few of us know how to fill cavities, repair leaks, write contracts, predict the weather or fix our cars. Many of us, however, can get our cavities filled, our cars repaired and our weather predicted. **Markets and organizations allow the knowledge that is held by few to reach many**. In other words, they make us collectively wiser.

The amount of knowledge embedded in a society, however, does not depend mainly on how much knowledge each individual holds. It depends, instead, on the diversity of knowledge across individuals and on their ability to combine this knowledge, and make use of it, through complex webs of interaction. A hunter-gatherer in the Arctic must know a lot of things to survive. Without the knowledge embedded in an Inuit, most of us would die in the Arctic, as has been demonstrated by the number of Westerners who have tried and failed. Yet, the total amount of knowledge embedded in a hunter-gatherer society is not very different from that which is embedded in each one of its members. The secret of modern societies is not that each person holds much more productive knowledge than those in a more traditional society. The secret to modernity is that we collectively use large volumes of knowledge, while each one of us holds only

▶

a few bits of it. Society functions because its members form webs that allow them to specialize and share their knowledge with others.

We can distinguish between two kinds of knowledge: *explicit* and *tacit*. Explicit knowledge can be transferred easily by reading a text or listening to a conversation. Yesterday's sports results, tomorrow's weather forecast or the size of the moon can all be learned quickly by looking them up in a newspaper or on the web. And yet, if all knowledge had this characteristic, the world would be very different. Countries would catch up very quickly to frontier technologies, and the income differences across the world would be much smaller than what we see today. The problem is that crucial parts of knowledge are tacit and therefore hard to embed in people. Learning how to fix dental problems, speak a foreign language, or run a farm requires a costly and time-consuming effort. As a consequence, it does not make sense for all of us to spend our lives learning how to do everything. **Because it is hard to transfer, tacit knowledge is what constrains the process of growth and development**. Ultimately, differences in prosperity are related to the amount of tacit knowledge that societies hold.

Because embedding tacit knowledge is a long and costly process, we specialize. This is why people are trained for specific occupations and why organizations become good at specific functions. To fix cavities you must be able to identify them, remove the decayed material and replace it. To play baseball, you must know how to catch, field and bat, but you do not need to know how to give financial advice or fix cavities. On the other hand, to perform the function of baseball player, knowing how to catch a ball is not enough (you must also be able to field and bat). In other words, in allocating productive knowledge to individuals, it is important that the chunks each person gets be internally coherent so that he or she can perform a certain function. We refer to these modularized chunks of embedded knowledge as **capabilities**. Some of these capabilities have been modularized at the level of individuals, while others have been grouped into organizations and even into networks of organizations.

For example, consider what has happened with undergraduate degrees, which in the US take four years of study. This norm has remained constant for the last four centuries. During the same period, however, knowledge has expanded enormously. The university system did not respond to the increase in knowledge by lengthening the time it takes to get a college degree. Instead, it increased the diversity of degrees. What used to be a degree in philosophy, split into several branches, one being natural philosophy, which later split into physics, chemistry and biology and later into other disciplines such as ecology, earth sciences and psychology. The Bureau of Labor Statistics' Standard Occupation Classification for 2010 lists 840 different occupations, including 78 in healthcare, 16 in engineering, 35 kinds of scientists – in coarse categories such as "economists", "physicists" and

“chemists” – five types of artists, and eight kinds of designers. We can all imagine a much more nuanced classification in our respective fields. For instance, we could distinguish between economists that specialize in labor, trade, finance, development, industrial organization, macro and econometrics, among others. If we did this further disaggregation for all occupations, we would easily go into the tens of thousands. The only way that society can hold all of the knowledge we have is by distributing coherent pieces of it among individuals. It is the way the world adapts to expanding knowledge.

Most products, however, require more knowledge than can be mastered by any individual. Hence, those products require that individuals with different capabilities interact. Assume that a person has the capacity to hold an amount of tacit knowledge equal to one ***personbyte***. How can you make a product that requires 100 different personbytes? Obviously, it cannot be made by a micro-entrepreneur working on her own. It has to be made either by an organization with at least 100 individuals (with a different personbyte each), or by a network of organizations that can aggregate these 100 personbytes of knowledge. How can a society hold a kilo-, mega- or giga-personbyte? Only through a deep division of labor, in which individuals become experts in small pieces of the available knowledge and then aggregate their personbytes into *peoplebytes* through organizations and markets.

For example, to make a shirt you need to design it, procure the fabric, cut it, sew it, pack it, brand it, market it and distribute it. In a firm that manufactures shirts, expertise in each of these knowledge chunks will be held by different people. And shirts require all of them. Moreover, you need to finance the operation, hire the relevant people, coordinate all the activities and negotiate everybody’s buy-in, which in itself require different kinds of knowhow. We can say that putting together this operation requires *know-who* and *know-where*. Know-who can be thought of as knowledge of who has the requisite chunks of knowledge, and know-where as knowledge of where the people and organizations that have this knowledge are located. To make shirts, you can import the fabric and access the knowledge about looms and threading that is embedded in a piece of cloth. Yet some of the knowledge required cannot be accessed through shipped inputs. The people with the relevant knowledge must be near the place where shirts are made.

In fact, just as knowhow is modularized in people in the form of individual capabilities, larger amounts of knowhow are modularized in organizations, and networks of organizations, as organizational or collective capabilities. For example, to operate a garment plant you need power and water. You need to be able to move raw materials in and ship the final product out. Workers need access to urban transportation, day care centers and health facilities. To be able to operate, the plant manager needs all of these services to be locally available. This implies that others must be aggre-

▶

gating the personbytes required to generate power, provide clean water, and run a transportation system. The relevant capabilities to perform all of these functions reside in organizations that are able to package the relevant knowledge into transferable bundles. These are bundles of knowhow that are more efficiently organized separately and transferred as intermediate inputs. We can think of these bundles as organizational capabilities the manufacturer needs.

Ultimately, **the complexity of an economy is related to the multiplicity of useful knowledge embedded in it**. For a complex society to exist, and to sustain itself, people who know about design, marketing, finance, technology, human resource management, operations and trade law must be able to interact and combine their knowledge to make products. These same products cannot be made in societies that are missing parts of this capability set. **Economic complexity, therefore, is expressed in the composition of a country's productive output and reflects the structures that emerge to hold and combine knowledge.**

Knowledge can only be accumulated, transferred and preserved if it is embedded in networks of individuals and organizations that put this knowledge into productive use. Knowledge that is not used, however, is also not transferred, and will disappear once the individuals and organization that have it retire or die.

Said differently, countries do not simply make the products and services they need. They make the ones they can. To do so, they need people and organizations that possess relevant knowledge. Some goods, like medical imaging devices or jet engines, embed large amounts of knowledge and are the results of very large networks of people and organizations. By contrast, wood logs or coffee, embed much less knowledge, and the networks required to support these operations do not need to be as large. Complex economies are those that can weave vast quantities of relevant knowledge together, across large networks of people, to generate a diverse mix of knowledge-intensive products. Simpler economies, in contrast, have a narrow base of productive knowledge and produce fewer and simpler products, which require smaller webs of interaction. Because individuals are limited in what they know, the only way societies can expand their knowledge base is by facilitating the interaction of individuals in increasingly complex webs of organizations and markets. **Increased economic complexity is necessary for a society to be able to hold and use a larger amount of productive knowledge**, and we can measure it from the mix of products that countries are able to make.

SECTION 2

How Do We Measure *Economic Complexity*?

How do we go from what a country makes to what a country knows? If making a product requires a particular type and mix of knowledge, then the countries that make the product reveal having the requisite knowledge (see *Technical Box 2.1*). From this simple observation, it is possible to extract a few implications that can be used to construct a measure of economic complexity. First, countries whose residents and organizations possess more knowledge have what it takes to produce a more diverse set of products. In other words, the amount of embedded knowledge that a country has is expressed in its productive diversity, or the number of distinct products that it makes. Second, products that demand large volumes of knowledge are feasible only in the few places where all the requisite knowledge is available. We define ubiquity as the number of countries that make a product (*Figure 2.1*). Using this terminology, we can observe that complex products –those that contain many personbytes of knowledge–are less ubiquitous. The ubiquity of a product, therefore, reveals information about the volume of knowledge that is required for its production. Hence, the amount of knowledge that a country has is expressed in the diversity and ubiquity of the products that it makes.

A game of scrabble is a useful analogy. In scrabble, players use tiles containing single letters to make words. For instance, a player can use the tiles **R**, **A** and **C** to construct the word **CAR** or **ARC**. In this analogy, each product is represented by a word, and each capability, or module of embedded knowledge, is represented by a letter. We assume that each player has plenty of copies of the letters they have. Our measure of economic complexity corresponds to estimating what fraction of the alphabet a player possesses, knowing only how many words he or she can make, and how many other players can also make those same words.

Players who have more letters will be able to make more words. So we can expect the diversity of words (products) that a player (country) can make to be strongly related to the number of letters (capabilities) that he (it) has. Long words will tend to be rare, since they can only be put together by players with many letters. Hence, the number of players that can make a word tells us something about the variety of letters each word requires: longer words tend to be less ubiquitous, while shorter words tend to be more common. Similarly, ubiquitous products are more likely to require few capabilities, and less ubiquitous products are more likely to require a large variety of capabilities.

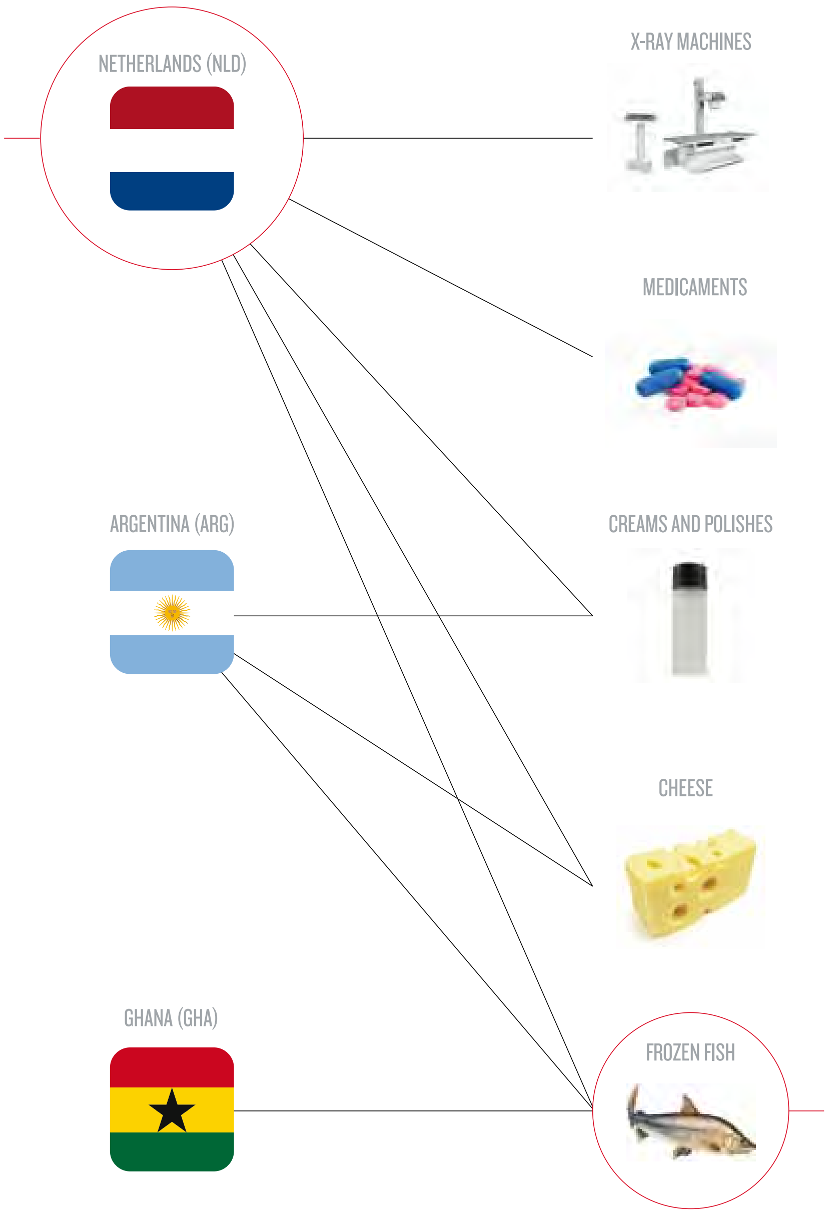
Diversity and ubiquity are, respectively, crude approximations of the variety of capabilities available in a country or required by a product. Both of these mappings are affected by the existence of rare letters, such as Q and X. For instance, players holding rare letters will be able to put together words that few other players can make, not because they have many letters, but because the letters that they have are rare. This is just like rare natural resources, such as uranium or diamonds. Yet, we can see whether low ubiquity originates in scarcity or complexity by looking at the number of other words that the makers of rare words are able to form. If these players can only make a few other words, then it is likely that rarity explains the low ubiquity. However, if the players that can make these rare words are, in general, able to put together many other words, then it is likely that the low ubiquity of the word reflects the fact that it requires a large number of letters and not just a few rare ones.

Diversity can therefore be used to correct the information carried by ubiquity, and ubiquity can be used to correct the information carried by diversity. We can take this process a step further by correcting diversity using a measure of ubiquity that has already been corrected by diversity and vice versa. In fact, we can do this an infinite number of times using mathematics. This process converges after a few iterations and represents our quantitative measures of

FIGURE 2.1:

Graphical explanation of diversity and ubiquity.

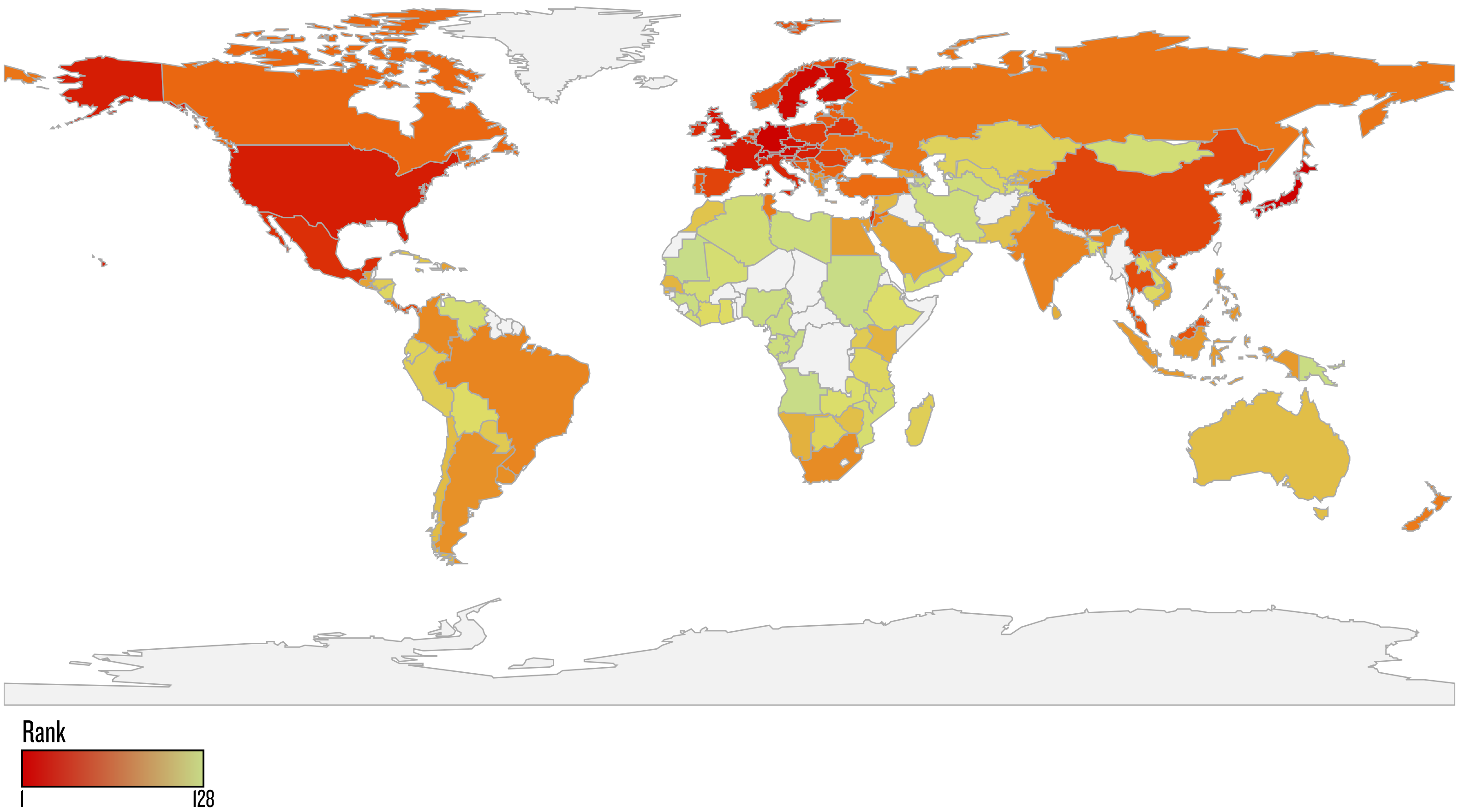
DIVERSITY ($k_{c,0}$):
Diversity is related to the number of products that a country is connected to. This is equal to the number of links that this country has in the network. In this example, using a subset of the 2009 data, the diversity of Netherlands is 5, that of Argentina is 3, and that of Gana is 1.



UBIQUITY ($k_{p,0}$):
Ubiquity is is related to the number of countries that a product is connected to. This is equal to the number of links that this product has in the network. In this example, using a subset of the 2009 data, the ubiquity of Cheese is 2, that of Fish is 3 and that of Medicaments is 1.

FIGURE 2.2:

Map of the World colored according to ECI Ranking.

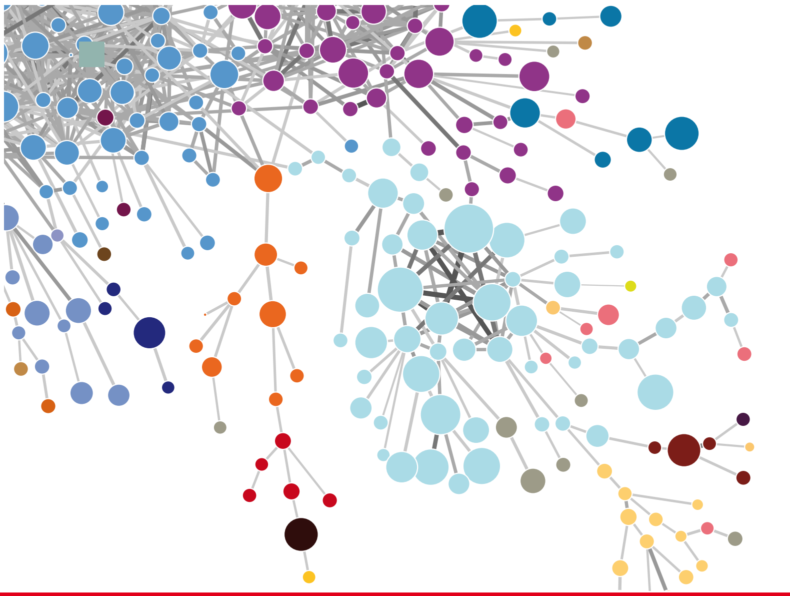


complexity. For countries, we refer to this as the Economic Complexity Index (ECI). The corresponding measure for products gives us the Product Complexity Index. *Technical Box 2.2* presents the mathematical definition of these two quantities and *Ranking 1* lists countries sorted by their ECI. *Figure 2.2* shows a map of the world colored according to a country's ECI ranking.

Consider the case of Singapore and Pakistan. The population of Pakistan is 34 times larger than that of Singapore. At market prices their GDPs are similar since Singapore is 38 times richer than Pakistan in per capita terms. Under the classification we use in this Atlas, they both export a similar number of different products, about 133. How can products tell us about the conspicuous differences in the level of development that exist between these two countries? Pakistan exports products that are on average exported by 28 other countries (placing Pakistan in the 60th percentile of countries in terms of the average ubiquity of their products), while Singapore exports products that are exported on average by 17 other countries (1st percentile). Moreover, the products that Singapore exports are exported by highly diversified countries, while those that Pakistan exports are exported by poorly diversified countries. Our mathematical approach exploits these second, third and higher order differences to create measures that approximate the amount of productive knowledge held in each of these countries. Ultimately, what countries make reveals what they know (see *Information Box 2.1*).

Take medical imaging devices. These machines are made in few places, but the countries that are able to make them, such as the United States or Germany, also export a large number of other products. We can infer that medical imaging devices are complex because few countries make them, and those that do tend to be diverse. By contrast, wood logs are exported by most countries, indicating that many countries have the knowledge required to export them. Now consider the case of raw diamonds. These products are extracted in very few places, making their ubiquity quite low. But is this a reflection of the high knowledge-intensity of raw diamonds? Of course not. If raw diamonds were complex, the countries that would extract diamonds should also be able to make many other things. Since Sierra Leone and Botswana are not very diversified, this indicates that something other than large volumes of knowledge is what makes diamonds rare (see *Information Box 2.2* on Product Complexity).

This Atlas relies on international trade data. We made this choice because it is the only dataset available that has a rich detailed cross-country information linking countries to the products that they produce in a standardized classification. As such, it offers great advantages, but it does have limitations. First, it includes data on exports, not production. Countries may be able to make things that they do not export. The fact that they do not export them, however, suggests that they may not be very good at them. Countries may also export things they do not make. To circumvent this issue we require that countries export a fair share of the products we connect them to (see *Technical Box 2.1*). Second, because the data is collected by customs offices, it includes only goods and not services. This is an important drawback, as services are becoming a rising share of international trade. Unfortunately, the statistical efforts of most countries of the world have not kept up with this reality. Finally, the data does not include information on non-tradable activities. These are an important part of the economic eco-system that allows products and services to be made. Our current research is focused on finding implementable solutions to these limitations, and we hope we will be able to present them in future versions of this Atlas.



INFORMATION BOX 2.1: AN ALTERNATIVE WAY TO UNDERSTAND OUR MEASURES OF ECONOMIC COMPLEXITY

Understanding the measures of economic complexity described in this Atlas can be sometimes challenging. Analogies, however, can help get our minds around what the economic complexity index is able to capture.

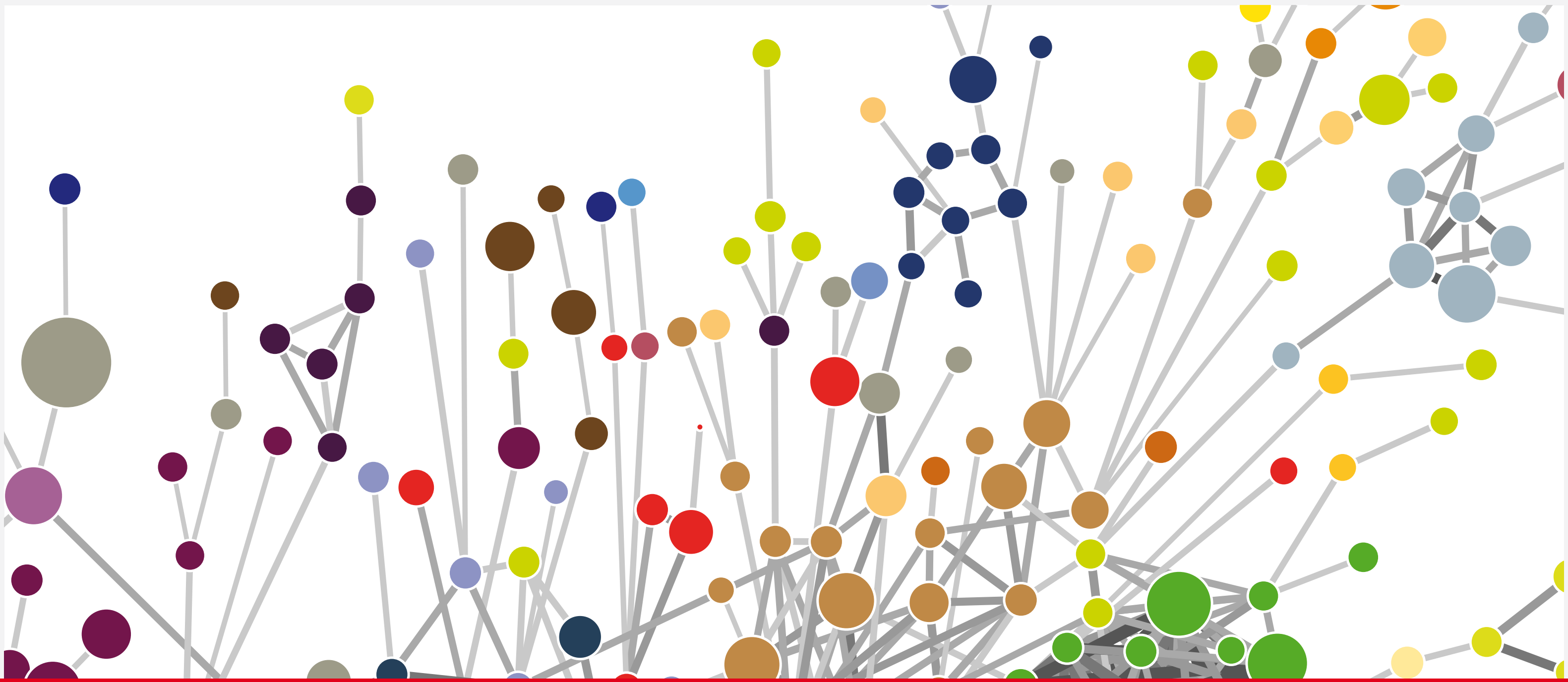
Think of a particular country and consider a random product. Now, ask yourself the following question: If this country cannot make this product, in how many other countries can this product be made? If the answer is many countries, then this country probably does not have a complex economy. On the other hand, if few other countries are able to make a product that this country cannot make, this would suggest that this is a complex economy.

Let us illustrate this with a few examples. According to our measures, Japan and Germany are the two countries with the highest levels of economic complexity. Ask yourself the question: If a good cannot be produced in Japan or Germany, where else can it be made? That list of countries is likely to be a very short one, indicating that Japan and Germany are complex economies. Now take an opposite example: if a product cannot be made in Mauritania or Sudan, where else can it be made? For most products this is likely to be a long list of countries, indicating that Sudan and Mauritania are among the world's least complex economies.

This analogy is useful to understand the difference between economic complexity and the level of income per capita of a country. Two countries that have high levels of economic complexity, but still low levels of per capita income are China and Thailand. Ask yourself the question, if you cannot produce it in China or Thailand, where else can you produce it? That list of countries will tend to be relatively short. The comparison becomes starker if we restrict it to countries with a similar level of per capita income, like Iran, Peru and Venezuela, countries that do not make things that many other can.

At the opposite end of this comparison, there are countries with high levels of per capita income but relatively low levels of economic complexity. Examples of this are Qatar, Kuwait, Oman, Venezuela, Libya and Chile. These countries are not rich because of the productive knowledge they hold but because of their "geological luck", given the large volumes of natural resources based wealth. Ask yourself the question; if you cannot build it in Chile or Venezuela, where else can you build it? The fact that there are many countries where it would be possible to produce many things that are not being made in Chile or Venezuela, including countries with a similar level of income such as Hungary or the Czech Republic, indicates that the level of economic complexity of these countries is low, despite their fairly high level of income.

In fact, as we show in this Atlas, the gap between a country's complexity and its level of per capita income is an important determinant of future growth: countries tend to converge to the level of income that can be supported by the knowhow that is embedded in their economy.



TECHNICAL BOX 2.1: MEASURING ECONOMIC COMPLEXITY:

If we define M_{cp}, as a matrix that is 1 if country c produces product p, and 0 otherwise, we can measure diversity and ubiquity simply by summing over the rows or columns of that matrix. Formally, we define:

$$Diversity = k_{c,0} = \sum_p M_{cp} \quad (1)$$

$$Ubiquity = k_{p,0} = \sum_c M_{cp} \quad (2)$$

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

$$k_{c,N} = \frac{1}{k_{c,0}} \sum_p M_{cp} \cdot k_{p,N-1} \quad (3)$$

$$k_{p,N} = \frac{1}{k_{p,0}} \sum_c M_{cp} \cdot k_{c,N-1} \quad (4)$$

We then insert (4) into (3) to obtain

$$k_{c,N} = \frac{1}{k_{c,0}} \sum_p M_{cp} \frac{1}{k_{p,0}} \sum_{c'} M_{c'p} \cdot k_{c',N-2} \quad (5)$$

$$k_{c,N} = \sum_{c'} M_{c'p} \cdot k_{c',N-2} \sum \frac{M_{cp} M_{c'p}}{k_{c,0} k_{p,0}} \quad (6)$$

and rewrite this as :

$$k_{c,N} = \sum_{c'} \widetilde{M}_{cc'} k_{c',N-2} \quad (7)$$

where

$$\widetilde{M}_{cc'} = \sum_p \frac{M_{cp} M_{c'p}}{k_{c,0} k_{p,0}} \quad (8)$$

We note (7) is satisfied when $k_{c,N} = k_{c,N-2} = 1$. This is the eigenvector of $\widetilde{M}_{cc'}$ which is associated with the largest eigenvalue. Since this eigenvector is a vector of ones, it is not informative. We look, 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 our measure of economic complexity. Hence, we define the Economic Complexity Index (ECI) as:

$$ECI = \frac{\vec{K} - \langle \vec{K} \rangle}{\text{stdev}(\vec{K})} \quad (9)$$

where < > represents an average, stdev stands for the standard deviation and

$$\vec{K} = \text{Eigenvector of } \widetilde{M}_{cc'} \text{ associated with second largest eigenvalue.} \quad (10)$$

Analogously, we define a Product Complexity Index (PCI). Because of the symmetry of the problem, this can be done simply by exchanging the index of countries (c) with that for products (p) in the definitions above. Hence, we define PCI as:

$$PCI = \frac{\vec{Q} - \langle \vec{Q} \rangle}{\text{stdev}(\vec{Q})} \quad (11)$$

where

$$\vec{Q} = \text{Eigenvector of } \widetilde{M}_{pp'} \text{ associated with second largest eigenvalue.} \quad (12)$$

INFORMATION BOX 2.2: THE WORLD'S MOST AND LEAST COMPLEX PRODUCTS

Table 2.2.1 and Table 2.2.2 show respectively the products that rank highest and lowest in the complexity scale. The difference between the world's most and less complex products is stark. The most complex products are sophisticated chemicals and machinery that tend to emerge from organizations where a large number of high skilled individuals participate. The world's least complex products, on the other hand, are raw minerals or simple agricultural products.

The economic complexity of a country is connected intimately to the complexity of the products that it exports. Ultimately, countries can only increase their score in the Economic Complexity Index by becoming competitive in an increasing number of complex industries.

TABLE 2.2.1: TOP 5 PRODUCTS BY COMPLEXITY

Product Code (SITC4)	Product Name	Product Community	Product Complexity Index
7284	Machines & appliances for specialized particular industries	Machinery	2.27
8744	Instrument & appliances for physical or chemical analysis	Chemicals & Health	2.21
7742	Appliances based on the use of X-rays or radiation	Chemicals & Health	2.16
3345	Lubricating petrol oils & other heavy petrol oils	Chemicals & Health	2.10
7367	Other machine tools for working metal or metal carbide	Machinery	2.05

TABLE 2.2.2: BOTTOM 5 PRODUCTS BY COMPLEXITY

Product Code (SITC4)	Product Name	Product Community	Product Complexity Index
3330	Crude oil	Oil	-3.00
2876	Tin ores & concentrates	Mining	-2.63
2631	Cotton, not carded or combed	Cotton, Rice, Soy & Others	-2.63
3345	Cocoa beans	Tropical Agriculture	-2.61
7367	Sesame seeds	Cotton, Rice, Soy & Others	-2.58

TECHNICAL BOX 2.2: WHO MAKES WHAT?

When associating countries to products it is important to take into account the size of the export volume of countries and that of the world trade of products. This is because, even for the same product, we expect the volume of exports of a large country like China, to be larger than the volume of exports of a small country like Uruguay. By the same token, we expect the export volume of products that represent a large fraction of world trade, such as cars or footwear, to represent a larger share of a country's exports than products that account for a small fraction of world trade, like cotton seed oil or potato flour.

To make countries and products comparable we use Balassa's definition of Revealed Comparative Advantage or RCA. Balassa's definition says that a country has Revealed Comparative Advantage in a product if it exports more than its "fair" share, that is, a share that is equal to the share of total world trade that the product represents. For example, in 2008, with exports of $42 billion, soybeans represented 0.35% of world trade. Of this total, Brazil exported nearly $11 billion, and since Brazil's total exports for that year were $140 billion, soybeans accounted for 7.8% of Brazil's exports. This represents around 21 times Brazil's "fair share" of soybean exports (7.8% divided by 0.35%), so we can say that Brazil has revealed comparative advantage in soybeans.

Formally, if X_{cp} represents the exports of country c in product p, we can express the Revealed Comparative Advantage that country c has in product p as:

$$RCA_{cp} = \frac{X_{cp}}{\sum_c X_{cp}} \Big/ \frac{\sum_p X_{cp}}{\sum_{c,p} X_{cp}} \qquad (1)$$

We use this measure to construct a matrix that connects each country to the products that it makes. The entries in the matrix are 1 if country c exports product p with Revealed Comparative Advantage larger than 1, and 0 otherwise. Formally we define this as the M_{cp} matrix, where

$$M_{cp} = \begin{cases} 1 & \text{if } RCA_{cp} \geq 1; \\ 0 & \text{otherwise.} \end{cases} \qquad (2)$$

M_{cp} is the matrix summarizing which country makes what, and is used to construct the product space and our measures of economic complexity for countries and products. In our research we have played around with cutoff values other than 1 to construct the M_{cp} matrix and found that our results are robust to these changes.

Going forward, we smooth changes in export volumes induced by the price fluctuation of commodities by using a modified definition of RCA in which the denominator is averaged over the previous three years.

SECTION 3

Why Is *Economic Complexity* Important?

As we have argued, economic complexity reflects the amount of knowledge that is embedded in the productive structure of an economy. Seen this way, it is no coincidence that there is a strong correlation between our measures of economic complexity and the income per capita that countries are able to generate.

Figure 3.1 illustrates the relationship between the Economic Complexity Index (ECI) and Income per capita for the 128 countries studied in this Atlas. Here, we separate countries according to their intensity in natural resource exports. We color in red those countries for which natural resources, such as minerals, gas and oil, represent at least 10% of GDP. For the 75 countries with a limited relative presence of natural-resource exports (in blue), economic complexity accounts for 75 percent of the variance in income per capita. But as the *Figure 3.1* illustrates, countries with a large presence of natural resources can be relatively rich without being complex. If we control for the income that is generated from extractive activities, which has more to do with geology than knowhow, economic complexity can explain about 73 percent of the variation in income across all 128 countries. *Figure 3.2* shows the tight relationship between economic complexity and income per capita that emerges after we take into account a country's natural resource income.

Economic complexity, therefore, is related to a country's level of prosperity. As such, it is just a correlation of things we care about. The relationship between income and complexity, however, goes deeper than this. **Countries whose economic complexity is greater than what we would expect, given their level of income, tend to grow faster than those that are "too rich" for their current level of economic complexity. In this sense, economic complexity is not just a symptom or an expression of prosperity: it is a driver.**

Technical Box 3.1 presents the regression that we use to relate economic complexity to subsequent economic growth. The equation is simple. We regress the growth in per capita income over 10-year periods on economic complexity, while controlling for initial income and for the increase in real natural resource income experienced during that period. We also include an interaction term between initial income per capita and the ECI. The increase in the explanatory power of the growth equation that can be attributed to the Economic Complexity Index is at least 15 percentage points, or more than a third of the variance explained by the whole equation. Moreover, the size of the estimated effect is large: **an increase of one standard deviation in complexity, which is something that Thailand achieved between 1970 and 1985, is associated with a subsequent acceleration of a country's long-term growth rate of 1.6 percent per year. This is over and above the growth that would have been expected from mineral wealth and global trends.**

The ability of the ECI to predict future economic growth suggests that countries tend to move towards an income level that is compatible with their overall level of embedded knowhow. On average, their income tends to reflect their embedded knowledge. But when it does not, it gets corrected through accelerated or diminished growth. The gap between a country's level of income and complexity is the key vari-

FIGURE 3.1:

Shows the relationship between income per capita and the Economic Complexity Index (ECI) for countries where natural resource exports are larger than 10% of GDP (red) and for those where natural resource exports are lower than 10% of GDP (blue). For the latter group of countries, the Economic Complexity Index accounts for 75% of the variance. Countries in which the levels of natural resource exports is relatively high tend to be significantly richer than what would be expected given the complexity of their economies, yet the ECI still correlates strongly with income for that group.

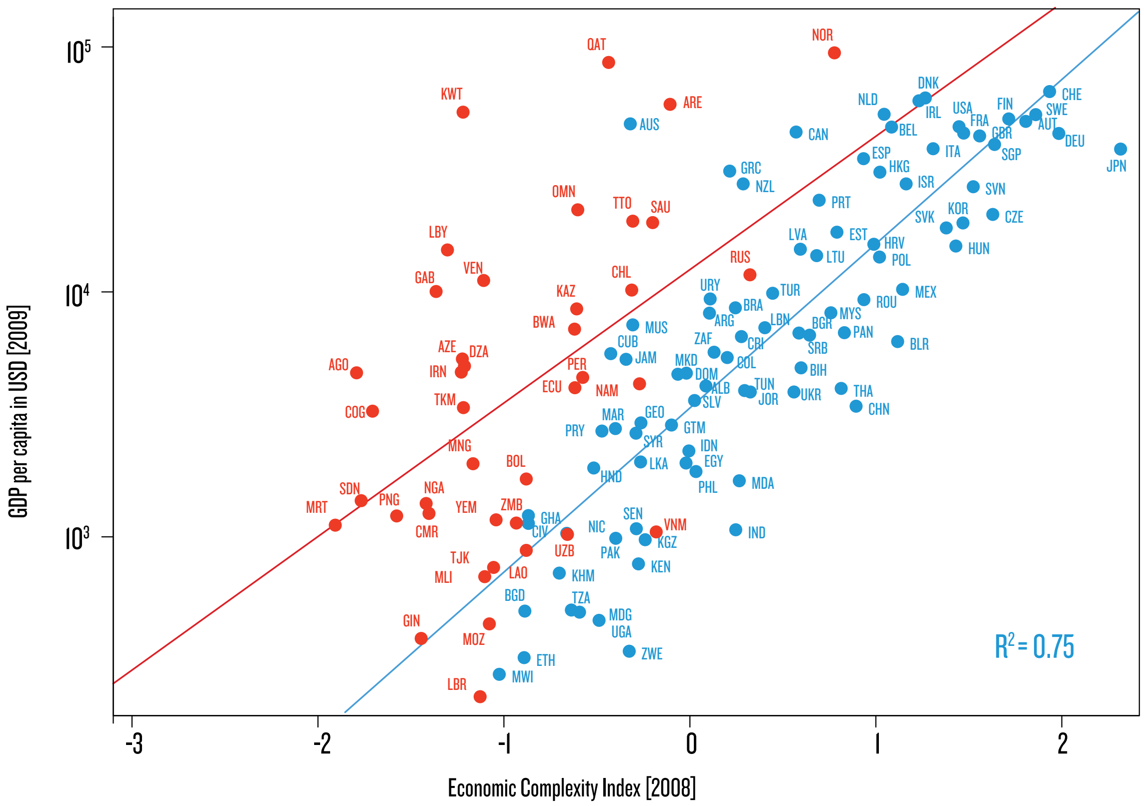


FIGURE 3.2:

Shows the relationship between economic complexity and income per capita obtained after controlling for each country's natural resource exports. After including this control, through the inclusion of the log of natural resource exports per capita, economic complexity and natural resources explain 73% of the variance in per capita income across countries.

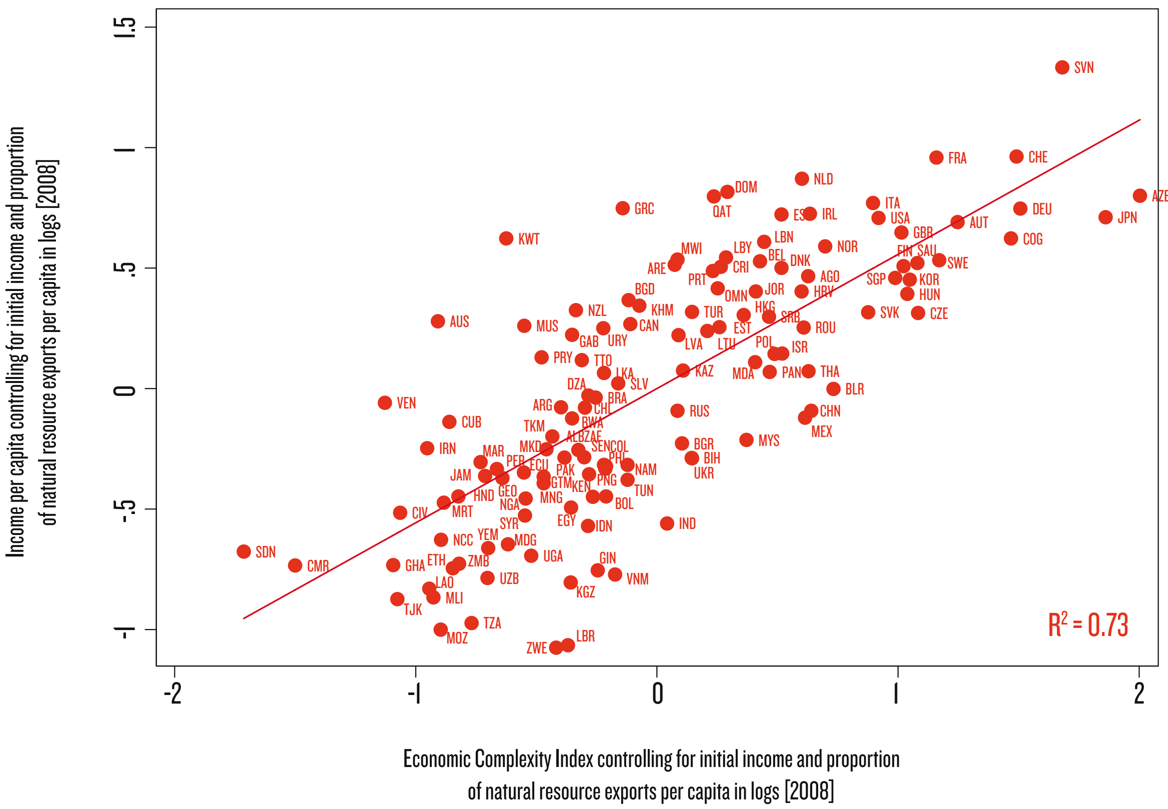
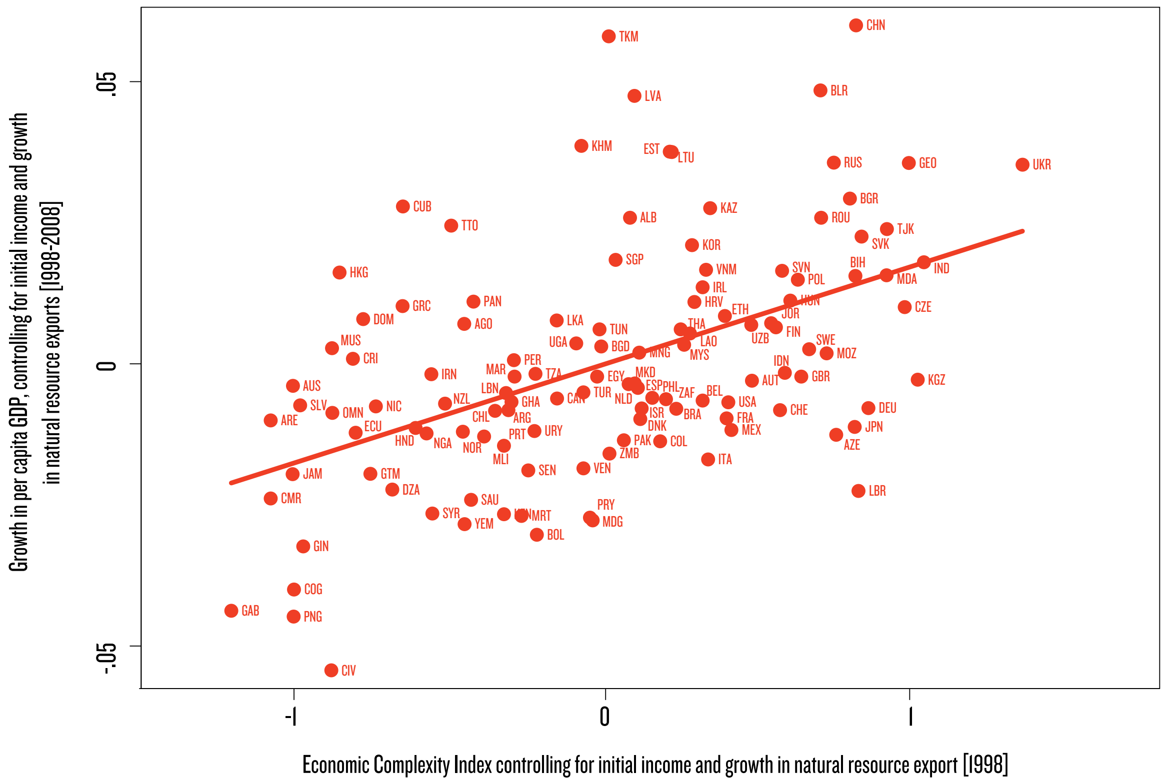


FIGURE 3.3:

Shows the relationship between the annualized GDP per capita growth for the period between 1998 and 2008 and the Economic Complexity Index for 1998, after taking into account the initial level of income and the increase in natural resource exports during that period (in constant dollars as a share of initial GDP).



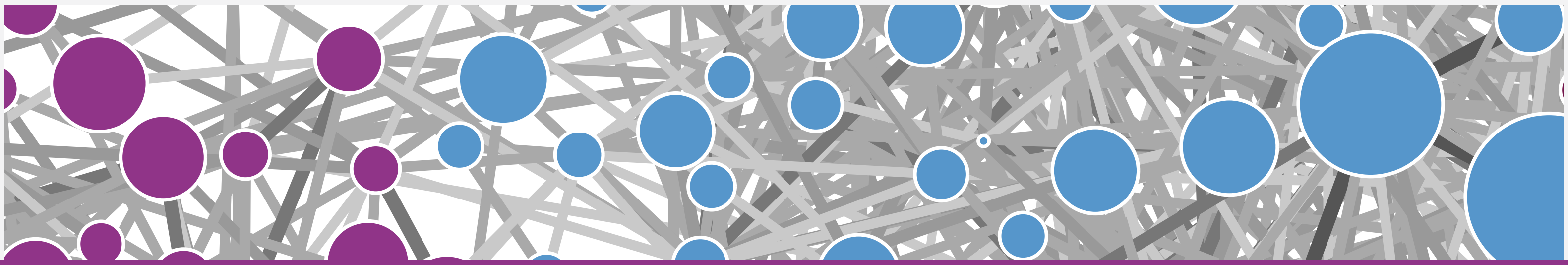
able that we use here to estimate the growth potential of countries (*Figure* 3.3).

It is important to note what the Economic Complexity Index is not about: it is not about export-oriented growth, openness, export diversification or country size. Although we calculate the ECI using export data, the channel through which it contributes to future growth is not limited to its impact on the growth of exports. Clearly, countries whose exports grow faster, all other things being equal, will necessarily experience higher GDP growth. This is simply because exports are a component of GDP. However, as *Technical Box* 3.2 shows, the contribution of the ECI to future economic growth remains strong after accounting for the growth in real exports.

The ECI is also not about openness to trade: the impact of the ECI on growth is essentially unaffected if we account for differences in the ratio of exports to GDP. And the ECI is not a measure of export diversification. Controlling for standard measures of export concentration, such as the Herfindahl-Hirschman Index, does not affect our results. In fact, neither openness nor export concentration are statistically significant determinants of growth after controlling for the ECI (see *Technical Box* 3.2).

Finally, the ECI is not about a country's size. The ability of the ECI to predict growth is unaffected when we take into account a country's size, as measured by its population, while the population itself is not statistically significant (see *Technical Box* 3.2).

In short, economic complexity matters because it helps explain differences in the level of income of countries, and more important, because it predicts future economic growth. Economic complexity might not be simple to accomplish, but the countries that do achieve it, tend to reap important rewards.



TECHNICAL BOX 3.1: THE GROWTH REGRESSION

To analyze the impact of the Economic Complexity Index (ECI) on future economic growth we estimate two regressions where the dependent variable is the annualized growth rate of GDP per capita for the periods 1978-1988, 1988-1998 and 1998-2008. In the first of these equations we do not include ECI and use only two control variables: the logarithm of the initial level of GDP per capita in each period and the increase in natural resource exports 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 the economic literature as convergence. The second control variable captures the effect on growth of increases in income that come from natural resource wealth, which complexity does not explain. In addition, we include a dummy variable 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 28.5 percent of the variance in countries' growth rates. This is shown in the first column of Table 3.1.1.

In addition to the above mentioned variables, the second regression includes the effect of economic complexity on growth. We do this by adding two additional terms: the ECI at the beginning of the decade and an interaction term between the ECI and the initial level of GDP per capita. The interaction attempts to capture the idea that the contribution of economic complexity to future economic growth depends on the level of per capita income. The second column of Table 3.1.1 shows that economic complexity is strongly associated with future economic growth. The negative coefficient on the interaction term indicates that the impact of complexity on growth declines with a country's level of income. For example, according to the estimation in Column 2, and using data for 1998, an increase in the ECI of one standard deviation would accelerate growth by 2.3 percent per year in a country at the 10th percentile of income, by 1.6 percent in a country at the median income, and by 0.7 percent for countries in the 90th percentile. The variables contained in Column 2 jointly account for 43.4 percent of the variance in growth rates. The difference between these two regressions indicates that the ECI increases the regression's R^2 in 15 percentage points. This represents over a third of the explained fraction of the 43.4 percent of the variance that the equation explains as a whole.

The estimates of the second column of Table 3.1.1 are used to forecast the growth in GDP per capita and rank countries according to their growth potential (See Table 3.1.1). To predict average annualized growth between 2008 and 2020 we make two assumptions. First, we assume a worldwide common growth term for the decade, which we take to be the same as that observed in the 1998-2008 period. Changing this assumption would affect the growth rate of all countries by a similar amount but would not change the rankings. Second, we assume that there will be no change in the real value of natural resource exports as a share of initial GDP. This implies that the we assume that natural resource exports in real terms in the next decade will remain at the record-high levels achieved in 2008. This assumption may underestimate the effect on countries whose volumes of natural resource extraction will increase significantly and over-estimate the growth in countries that will see their natural-resource export volumes declines. A higher (lower) constant dollar price of natural resource exports would improve (reduce) the projected growth performance of countries by an amount proportional to their natural resource intensity.

TABLE 3.1.1

	Annualized growth in GDP pc (by decade)	
	(1978-1988, 1988-1998, 1998-2008)	
VARIABLES	(1)	(2)
Initial Income per capita, log	-0.00017	-0.00638***
	(0.001)	(0.001)
Increase in natural resource exports	0.03960***	0.03682***
- in constant dollars (as a share of initial GDP)	(0.008)	(0.010)
Initial Economic Complexity Index (ECI)		0.04430***
		(0.009)
[ECI] X [Income per capita, log]		-0.00371***
		(0.001)
Constant	0.03036***	0.08251***
	(0.008)	(0.011)
Observations	291	291
R^2	0.285	0.434
Year FE	Yes	Yes

TECHNICAL BOX 3.2: ECONOMIC COMPLEXITY: THE VOLUME AND CONCENTRATION OF EXPORTS AND COUNTRY SIZE

This box explores the robustness of the impact of the Economic Complexity Index on growth. While the ECI is constructed using export data, its relationship with future growth is not driven by export volumes or concentration. To show this, we start with our basic growth equation (Table 3.2.1, column 1). Column 2 adds to this equation the increase in the real value of the exports of goods and services in the decade in question as a fraction of initial GDP. Exports are a component of GDP, and therefore, we expect them to contribute to growth. Nevertheless, after including the increase in exports, the effect of ECI on growth remains strong and significant. Column 3 introduces export as a share of GDP. We use this as a measure of openness. Column 4 includes the Herfindahl-Hirschman index as a measure of export concentration. Column 5 includes the log of initial population as a measure of size. This is equivalent to introducing total GDP, given that we are already controlling for GDP per capita. The contribution to growth of the variables introduced in columns 3, 4 and 5 are estimated to be very close to zero, are not statistically significant and do not affect the ability of the ECI to predict future economic growth.

TABLE 3.2.1

	Annualized growth in GDP pc (by decade)				
	(1978-1988, 1988-1998, 1998-2008)				
VARIABLES	(1)	(2)	(3)	(4)	(5)
Initial Economic Complexity Index (ECI)	0.04430***	0.03005***	0.04240***	0.04143***	0.04389***
	(0.009)	(0.007)	(0.008)	(0.010)	(0.009)
[ECI] X [Income per capita, log]	-0.00371***	-0.00244***	-0.00345***	-0.00354***	-0.00381***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Increase exports (goods and services)		0.04549***			
- in constant dollars (as a share of initial GDP)		(0.007)			
Exports to GDP			0.00009		
			(0.000)		
Export Concentration				-0.00890	
				(0.008)	
Population, log					0.00168
					(0.001)
Initial Income per capita, log	-0.00638***	-0.00562***	-0.00729***	-0.00611***	-0.00558***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Increase in natural resource exports	0.03682***	0.00169	0.03441***	0.03699***	0.03758***
- in constant dollars (as a share of initial GDP)	(0.010)	(0.005)	(0.008)	(0.010)	(0.010)
Constant	0.08251***	0.06741***	0.08616***	0.08145***	0.04878**
	(0.011)	(0.011)	(0.011)	(0.011)	(0.022)
Observations	291	260	284	291	291
R^2	0.434	0.584	0.449	0.436	0.440
Year FE	Yes	Yes	Yes	Yes	Yes

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

SECTION 4

How Is *Complexity* Different from Other Approaches?

We are certainly not the first ones to look for correlates or causal factors of income and growth. There are plenty of others who have come before us. One strand of the literature has looked at the salience of institutions in determining growth, whereas others have looked at human capital or broader measures of competitiveness. Clearly, more complex economies have better institutions, more educated workers and more competitive environments, so these approaches are not completely at odds with each other or with ours. In fact, institutions, education, competitiveness and economic complexity emphasize different aspects of the same intricate reality. It is not clear, however, that these different approaches have the same ability to capture factors that are verifiably important for growth and development. In this section, we compare each of these measures with the Economic Complexity Index and gauge their marginal contribution to income and economic growth.

MEASURES OF GOVERNANCE AND INSTITUTIONAL QUALITY

Some of the most respected measures of institutional quality are the six Worldwide Governance Indicators (WGIs), which the World Bank has published biennially since 1996. These indicators are used, for example, as eligibility criteria by the Millennium Challenge Corporation (MCC) when selecting the countries they chose to support. These criteria are based on the direct connection between governance and growth and poverty reduction.

To the extent that governance is important to allow individuals and organizations to cooperate, share knowledge and make more complex products, it should be reflected in the kind of industries that a country can support. Therefore, the Economic Complexity Index indirectly captures information about the quality of governance in the country. Which indicator captures information that is more relevant for growth is an empirical question.

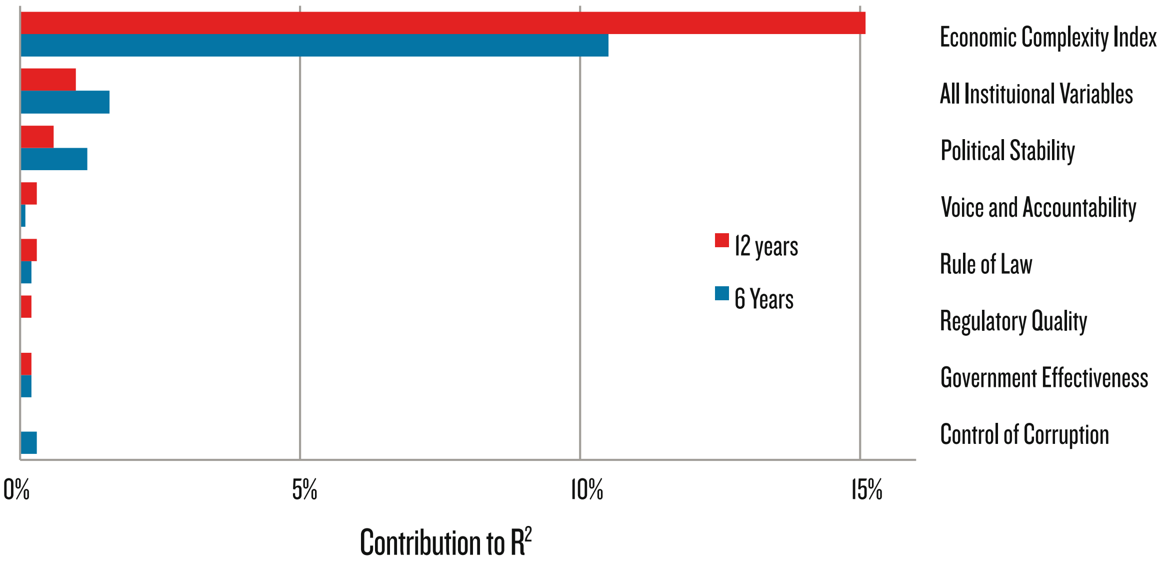
Here we compare the contribution to future economic growth implied by the WGIs and the ECI using a technique described in *Technical Box 4.1*. Since the WGIs are available only since 1996, we perform this exercise using the 1996-2008 period as a whole and as two consecutive 6-year periods. We also compare with each individual WGI and with the six of them together.

Figure 4.1 shows that the ECI accounts for 15.1 percent of the variance in economic growth during the 1996-2008 period, while the six WGIs combined account only for 1.0 percent. For the estimation using the two six year periods, we find that ECI accounts for 10.5% of the variance in growth, whereas the six WGIs combined account for 1.6%.

We conclude that as far as future economic growth is concerned, **the Economic Complexity Index captures significantly more growth-relevant information than the 6 World Governance Indicators, either individually or combined.** This does not mean that governance is not important for the economy. It suggests that the aspects of governance important for growth are weakly reflected in the WGIs and appear to be more strongly reflected in the economic activities that thrive in each country. These may be more effectively captured by the Economic Complexity Index.

FIGURE 4.1:

Contribution to the variance of economic growth from the Economic Complexity Index (ECI) and from the measures of governance and institutional quality.



MEASURES OF HUMAN CAPITAL

Another strand of the growth and development literature has looked at the impact of human capital on economic growth. The idea that human capital is important for income and growth is not unrelated to our focus on the productive knowledge that exists in a society. The human capital literature, however, has placed its attention on measures of formal education. Instead, our approach emphasizes the tacit productive knowledge that is embedded in a country's economic activities.

The standard variables used as a proxy for human capital are the number of years of formal schooling attained by those currently of working age, or the school enrollment of the young (Barro and Lee, 2010). Since these indicators do not take into account the quality of the education received by pupils, they have been subject to criticism resulting in new measures of educational quality. These measures use test scores from standardized international exams, such as the OECD Programme for International Student Assessment (PISA) or the Trend in International Mathematics and Science Study (TIMSS). Hanuschek and Woessmann (2008) collected data for all the countries that participated in either program and used this information to generate a measure of the cognitive ability of students for a cross-section of countries around the year 2000.

The information on productive knowhow captured by the Economic Complexity Index and by measures of human capital, are not just two sides of the same coin. Analytically, human capital indicators try to measure how much of the same knowledge individuals have, whether knowledge is measured as years of study of the national curriculum or as the skills mastered by students according to standardized international tests. In contrast, the Economic Complexity Index tries to capture the total amount of productive knowledge that is embedded in a society as a whole and is related to the *diversity* of knowledge that a society holds. Clearly, for a complex economy to exist, its members must be able to read, write and manipulate symbols, such as numbers or mathematical functions. This is what is taught in schools. Yet, the converse is not true: the skills acquired in school may be a poor proxy for the productive knowledge of society.

For example, if a country were to achieve the goal of having everybody finish a good secondary education and if this was the extent of its productive knowledge, nobody would know how to make a pair of shoes, a metal knife, a roll of paper or a patterned piece of cotton fabric. There is a reason why job offers request years of experience and not just years of schooling. This means that what a society makes affects what kinds of knowledge new workers can acquire on the job. The human capital approach emphasizes the opposite logic: what workers formally study is what affects what a society can produce.

Figure 4.2 shows the relationship between our measure of economic complexity and years of schooling for the year 2000. It is clear that there is positive relationship between the two (R^2=50%). Countries like India and Uganda, or Mongolia and Mexico, have very similar levels of average formal education. Yet, they differ dramatically in economic complexity. India is much more complex than Uganda, and Mexico is much more complex than Mongolia.

Figure 4.3 shows that the relationship between cognitive ability and economic complexity is also positive.Here we

FIGURE 4.2:

Relationship between Years of Schooling and the Economic Complexity Index (ECI) for the year 2000.

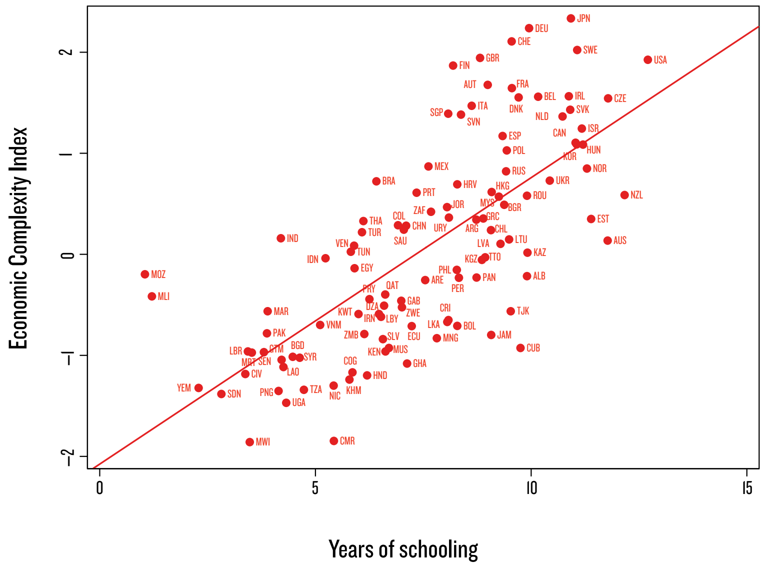
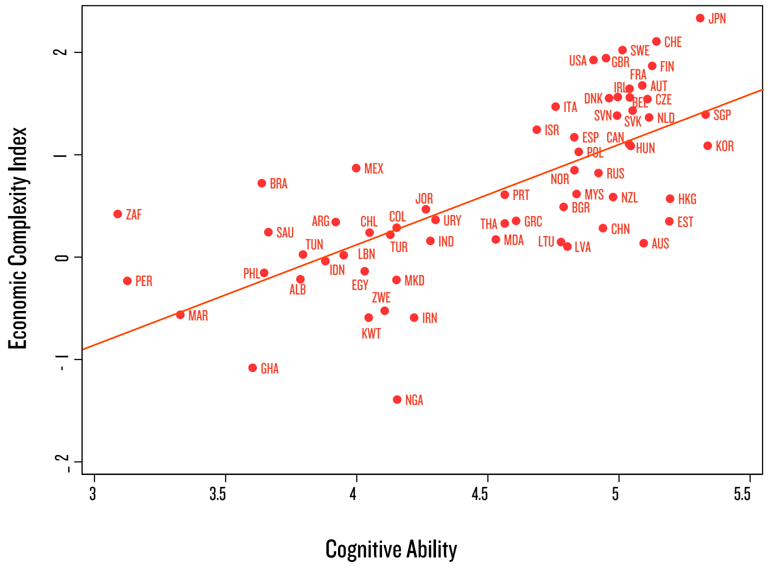


FIGURE 4.3:

Relationship between Cognitive Ability and the Economic Complexity Index (ECI) for the year 2000.



find that Brazil and Ghana are two countries with similar levels of cognitive ability, but very different levels of economic complexity. Brazil is two standard deviations more complex than Ghana. The same is true for Colombia and Nigeria. Their measured cognitive abilities are the same, but Colombia is nearly 1.5 standard deviations more complex than Nigeria.

For illustration purposes, consider the case of Ghana and Thailand. Both countries had similar levels of schooling in 1970, but Ghana expanded education more vigorously than Thailand in the subsequent 40 years (*Figure* 4.4). But Ghana's economic complexity and income stagnated as it remained an exporter of cocoa, aluminum, fish and forest products. By contrast, between 1970 and 1985 Thailand underwent a massive increase in economic complexity, equivalent to a change of one standard deviation in the Economic Complexity Index (*Figure* 4.5). This caused a sustained economic boom in Thailand after 1985. As a consequence, the level of income per capita between Ghana and Thailand has since diverged dramatically (*Figure* 4.6).

Next, we measure these indicators' ability to predict future economic growth, using the same technique that we employed to compare ECI to the WGIs (see *Technical Box* 4.2). We begin by looking at the relationship between education, complexity and a country's level of income per capita. While data on years of schooling and school enrollment is available for several years, the data on educational quality exists only for a cross-section of countries around the year 2000. We use the data for this year to estimate equations where the dependent variable is the level of income per capita and the independent variables are the years of schooling of the labor force, the Hanushek and Woessmann measure of cognitive ability, and the ECI. We do not use school enrollment as this variable affects future human capital but not the human capital invested in creating today's income. The results, presented in *Figure* 4.7, indicate that the Economic Complexity Index explains 17.2 percent of the variance while years of schooling and cognitive ability account for only 3.6 percent of the variance when combined.

We also look at the ability of human capital and complexity to explain future growth. To do this we follow a similar methodology as before (see *Technical Box* 4.2). In this case, we include data on school enrollment at the secondary and tertiary levels as these would affect the years of schooling of the labor force going forward. We do not include cognitive ability as this variable exists only for a single year.

Figure 4.8 shows that economic complexity accounts for 12.1 percent of the variance in economic growth rates for the three decades between 1978 and 2008. All education variables, on the other hand, account only for 2.6 percent when combined.

These results show that the Economic Complexity Index contains information that is more directly related to a country's level of income and its future rate of growth than the standard variables used to measure human capital.

MEASURES OF COMPETITIVENESS

Finally, we look at measures of competitiveness. The most respected source of these measures is the World Economic Forum's Global Competitiveness Index (GCI). The GCI has been published since 1979. Over the course of more than 30 years, the coverage of the GCI has been expanded and

FIGURE 4.4:

Years of schooling of Thailand and Ghana as a function of time.

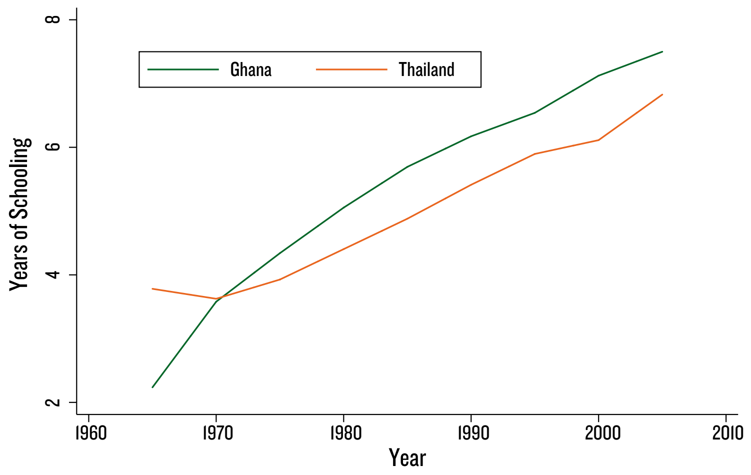


FIGURE 4.5:

Economic Complexity Index (ECI) of Thailand and Ghana as a function of time.

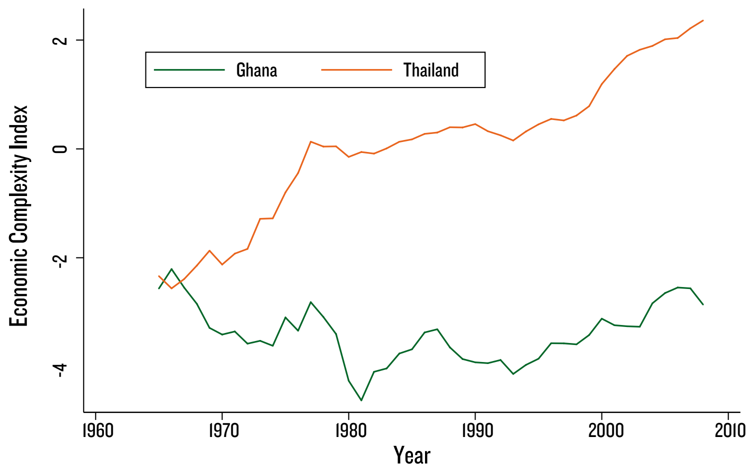
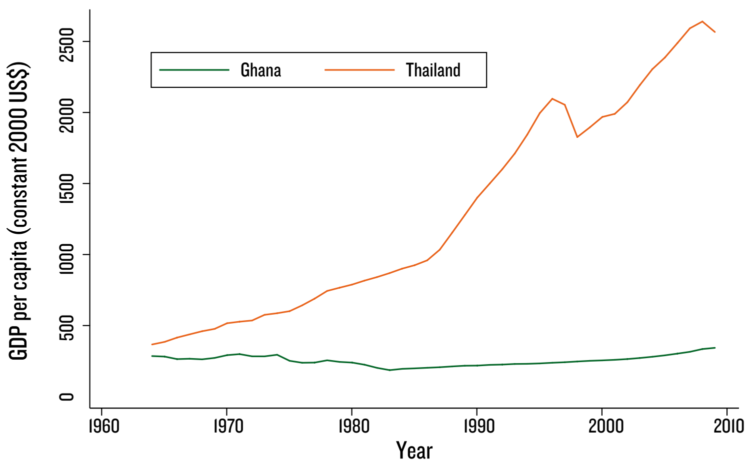


FIGURE 4.6:

Evolution of the GDP per capita of Thailand and Ghana as a function of time.



improved methodologically, going through two major revisions in 2001 and 2006. By 1995, the GCR ranked less than 50 countries, but over the years this number has increased, now reaching over 130 countries. The claim of the Global Competitiveness Report is that the index captures the fundamental variables that drive growth over the medium term:

> *"We define competitiveness as the set of institutions, policies, and factors that determine the level of productivity of a country. Because the rates of return are the fundamental drivers of the growth rates of the economy, a more competitive economy is one that is likely to grow faster in the medium to long run."*
> (GLOBAL COMPETITIVENESS REPORT 2010. CHAPTER 1.1, PAGE 4)

The GCI develops over 150 measures of elements that it considers important for competitiveness and then averages them. The ECI looks, instead, at the actual kinds of industry that a country can support. Both should capture information that is relevant to an economy's ability to grow. Which one does so more effectively is an empirical question that we address next.

Since we only have data for the GCI rankings, and not the underlying value of the index, we do the analysis using the rankings of the Economic Complexity Index instead of its value. This allows for a fairer comparison. **We do comparisons using 5 and 10 year panels starting in 1979 and find that the GCI rankings contribute significantly less to the variance of economic growth than the ECI** (see *Technical Box* 4.3 and *Figure* 4.9).

We conclude that the Economic Complexity Index can account for a significant fraction of the cross-country variation in income per capita and economic growth, and that the ECI is a much stronger predictor of growth than other commonly used indicators that measure human capital, governance or competitiveness.

FIGURE 4.7:

Contribution to the variance of income from the Economic Complexity Index (ECI) and measures of Human Capital.

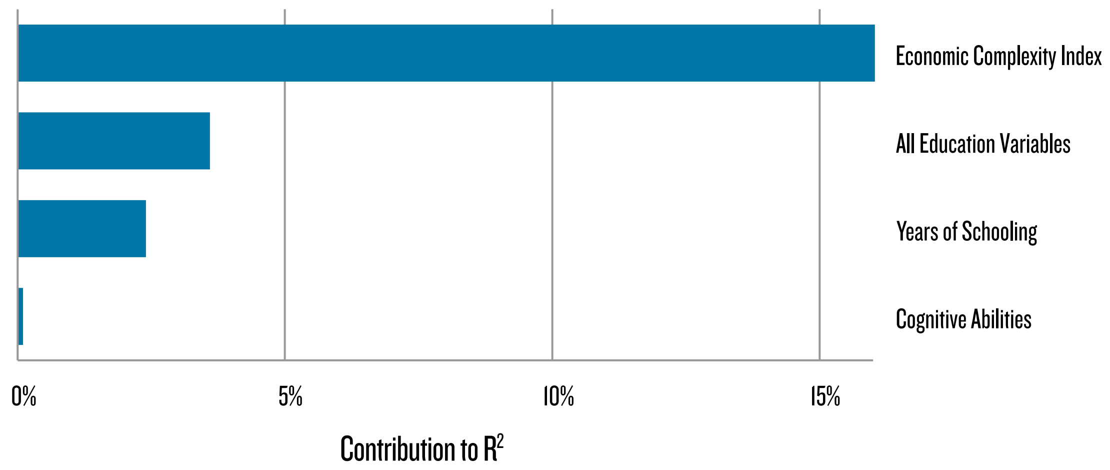


FIGURE 4.8:

Contribution to the variance of economic growth from the Economic Complexity Index (ECI) and measures of Human Capital.

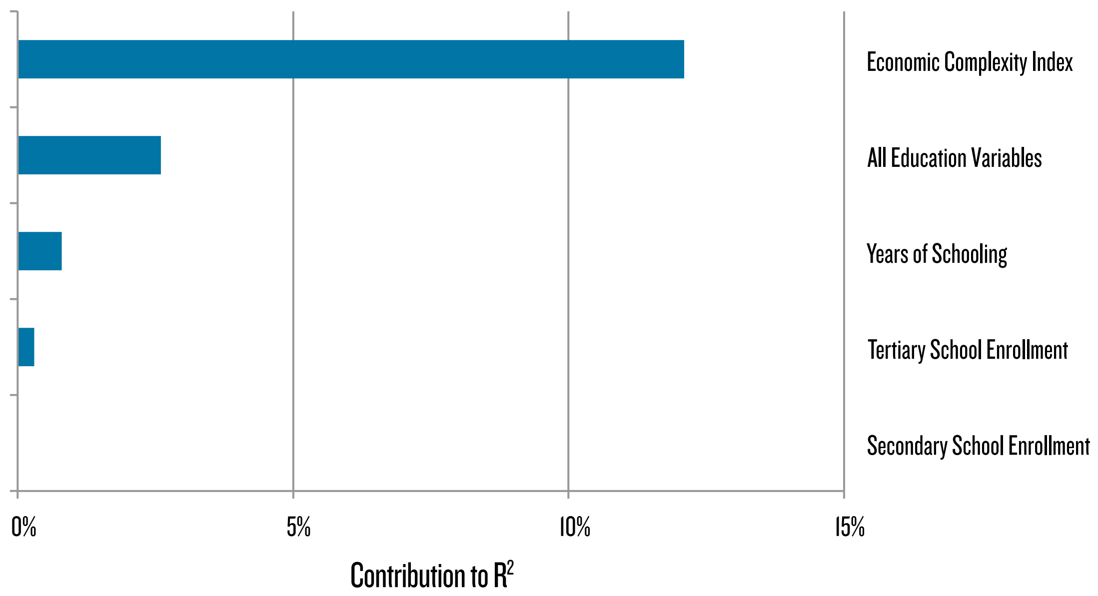
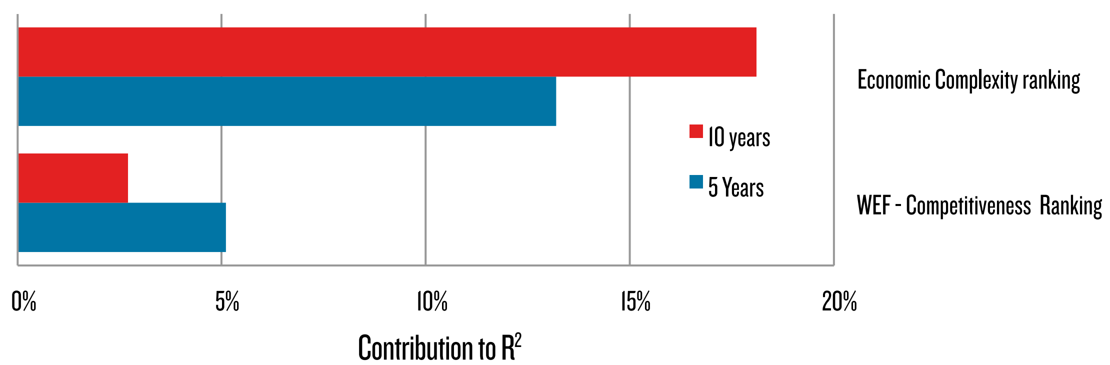
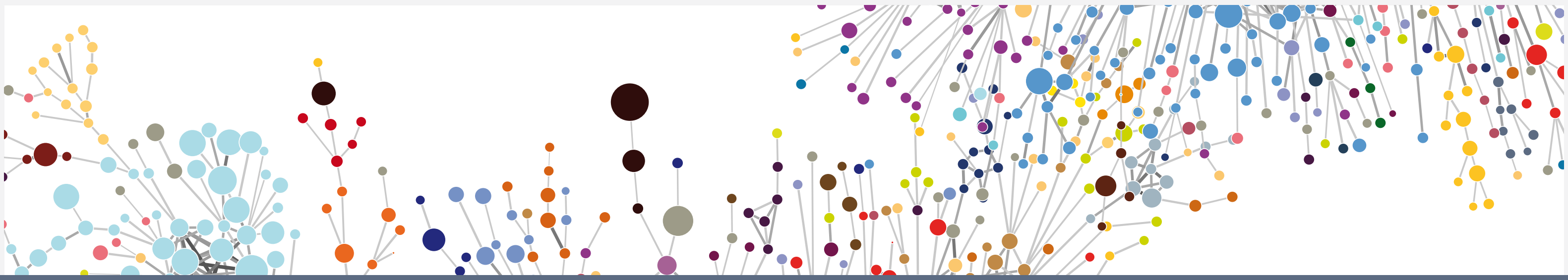


FIGURE 4.9:

Contribution to the variance of economic growth from the Economic Complexity Index (ECI) and measures of competitiveness.





TECHNICAL BOX 4.1: GOVERNANCE AND COMPLEXITY

We compare the contribution to economic growth of the Worldwide Governance Indicators (WGIs) and economic complexity by estimating a growth regression where all of the WGIs and the Economic Complexity Index are used as explanatory variables. As controls we include the logarithm of per capita income, the increase in natural resource exports during the period and the initial share of GDP represented by natural resource exports. The contribution of each variable is estimated by taking the difference between the R^2 obtained for the regression using all variables and that obtained for the regression where the variable was removed.

TABLE 4.1.1

	Annualized growth in GDP pc (6 years)								
	1996-2002, 2002-2008								
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Income per capita, logs	-0.00253***	-0.00149***	-0.00251***	-0.00251***	-0.00257***	-0.00253***	-0.00258***	-0.00254***	-0.00210***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Increase in natural resource exports	0.00420**	0.00246	0.00418**	0.00417**	0.00426**	0.00423**	0.00431**	0.00411**	0.00393**
- in constant dollars (as a share of initial GDP)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Initial NR exports to GDP	0.01181***	0.00879***	0.01178***	0.01177***	0.01169***	0.01184***	0.01214***	0.01174***	0.01121***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Initial Economic Complexity Index (ECI)	0.00547***		0.00547***	0.00536***	0.00598***	0.00551***	0.00531***	0.00537***	0.00500***
	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
[ECI] x GDP per capita, log	-0.00035**		-0.00035**	-0.00034**	-0.00042**	-0.00036**	-0.00032**	-0.00034**	-0.00026*
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Initial Control of Corruption	-0.00006	-0.00123		0.00003	-0.00005	0.00000	-0.00006	-0.00006	
	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
Initial Government Effectiveness	0.00029	0.00110	-0.00006		0.00050	0.00036	0.00013	0.00025	
	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)	
Initial Political Stability	0.00102**	0.00119**	0.00091**	0.00104**		0.00105**	0.00106**	0.00107**	
	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	
Initial Rule of Law	0.00022	0.00105	0.00046	0.00031	0.00082		0.00003	0.00029	
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	
Initial Regulatory Quality	-0.00056	-0.00125	-0.00050	-0.00052	-0.00068	-0.00053		-0.00039	
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	
Initial Voice and Accountability	0.00032	0.00053	0.00034	0.00031	0.00050	0.00033	0.00014		
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
Constant	0.02428***	0.01621***	0.02412***	0.02410***	0.02452***	0.02421***	0.02453***	0.02429***	0.01868***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Observations	243	243	243	243	243	243	243	243	243
Adjusted R^2	0.420	0.315	0.417	0.422	0.408	0.422	0.420	0.421	0.404
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 difference		10.5%	0.3%	0.2%	1.2%	0.2%	0.0%	0.1%	1.6%

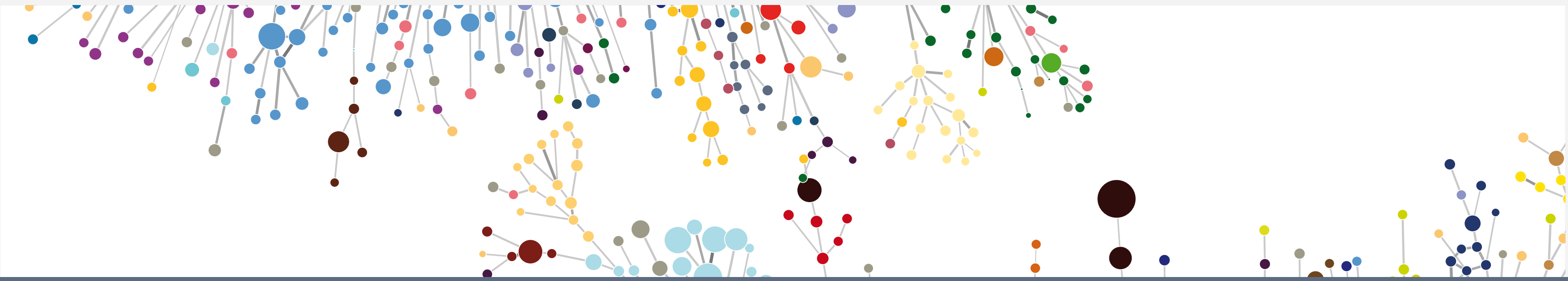
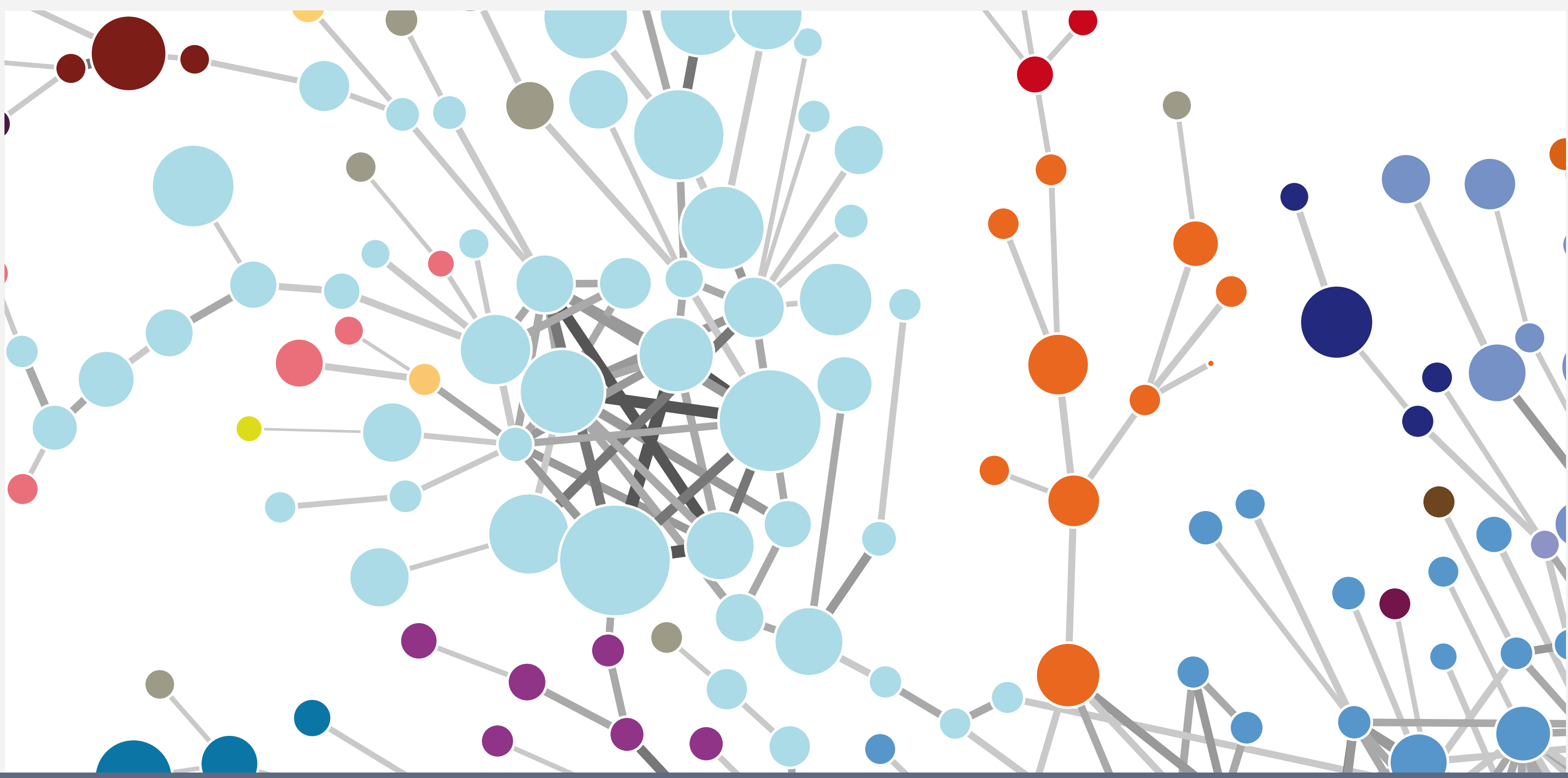


Table 4.1.1 shows the results of this procedure using two consecutive six year periods. Table 4.1.2 shows the same procedure using one twelve year period (1996-2008). Figure 4.1 of the main text, illustrates the differences in R^2 between the regression using all variables and those where individual variables were removed.

TABLE 4.1.2

	Annualized growth in GDP pc (12 years)								
	1996-2008								
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Income per capita, logs	-0.00202***	-0.00127***	-0.00199***	-0.00213***	-0.00203***	-0.00196***	-0.00207***	-0.00201***	-0.00202***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Increase in natural resource exports	0.00228***	0.00255***	0.00226***	0.00233***	0.00228***	0.00227***	0.00233***	0.00232***	0.00235***
- in constant dollars (as a share of initial GDP)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Initial NR exports to GDP	0.00361***	0.00314**	0.00357***	0.00367***	0.00342***	0.00367***	0.00383***	0.00361***	0.00369***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Initial Economic Complexity Index [ECI]	0.00380**		0.00371***	0.00436***	0.00405**	0.00415***	0.00364**	0.00392**	0.00467***
	(0.002)		(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
[ECI] x [GDP per capita, log]	-0.00017		-0.00015	-0.00026	-0.00020	-0.00020	-0.00014	-0.00019	-0.00026**
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Initial Control of Corruption	0.00004	-0.00079		-0.00027	0.00001	0.00029	0.00005	-0.00003	
	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
Initial Government Effectiveness	-0.00097	-0.00040	-0.00117		-0.00080	-0.00069	-0.00107	-0.00087	
	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)	
Initial Political Stability	0.00065*	0.00079*	0.00054	0.00057*		0.00078**	0.00069*	0.00059*	
	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	
Initial Rule of Law	0.00103	0.00237***	0.00121*	0.00073	0.00135*		0.00087	0.00094	
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	
Initial Regulatory Quality	-0.00040	-0.00125*	-0.00036	-0.00051	-0.00051	-0.00024		-0.00053	
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	
Initial Voice and Accountability	-0.00030	-0.00014	-0.00028	-0.00019	-0.00016	-0.00021	-0.00042		
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
Constant	0.01960***	0.01384***	0.01936***	0.02052***	0.01963***	0.01918***	0.01993***	0.01961***	0.01954***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Observations	118	118	118	118	118	118	118	118	118
Adjusted R^2	0.516	0.365	0.516	0.514	0.510	0.513	0.518	0.519	0.506
Difference in R^2		15.1%	0.0%	0.2%	0.6%	0.3%	0.2%	0.3%	1.0%



TECHNICAL BOX 4.2: EDUCATION, COGNITIVE ABILITY AND ECONOMIC COMPLEXITY

We compare the contribution to income of education, cognitive ability and economic complexity by regressing income against years of schooling, cognitive ability and the Economic Complexity Index. The contribution to income of each variable is estimated by taking the difference between the R^2 obtained for the regression using all variables and that obtained for a regression where the variable in question was removed.

Table 4.2.1 shows the results of this procedure for the year 2000, when cognitive ability data is available. Figure 4.7 in the main text summarizes the results.

We compare the contribution to growth of education and economic complexity by regressing growth against years of schooling secondary school enrollment, tertiary school enrollment, and the Economic Complexity Index. As additional controls we include the change in natural resource exports during the period, the logarithm of per capita income and year fixed effects. The contribution of each variable to growth is estimated by taking the difference between the R^2 obtained for a regression using all variables and one obtained for a regression where the variable in question was removed.

Table 4.2.2 shows the results of this procedure for ten year panels starting in 1978, 1988 and 1998. Figure 4.8 in the main text summarize the results.

TABLE 4.2.1

	Income per capita, log – Year 2000				
VARIABLES	(1)	(2)	(3)	(4)	(5)
Economic Complexity Index	0.998***		1.079***	1.042***	1.264***
	(0.204)		(0.214)	(0.160)	(0.138)
Years of schooling	0.134*	0.213**		0.14750**	
	(0.077)	(0.085)		(0.068)	
Cognitive ability	0.118	0.875***	0.344		
	(0.263)	(0.290)	(0.249)		
Constant	6.294***	2.861***	6.38972***	6.688***	7.826***
	(0.917)	(0.995)	(0.977)	(0.562)	(0.183)
Observations	59	59	59	59	59
R^2	0.620	0.448	0.596	0.619	0.584

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

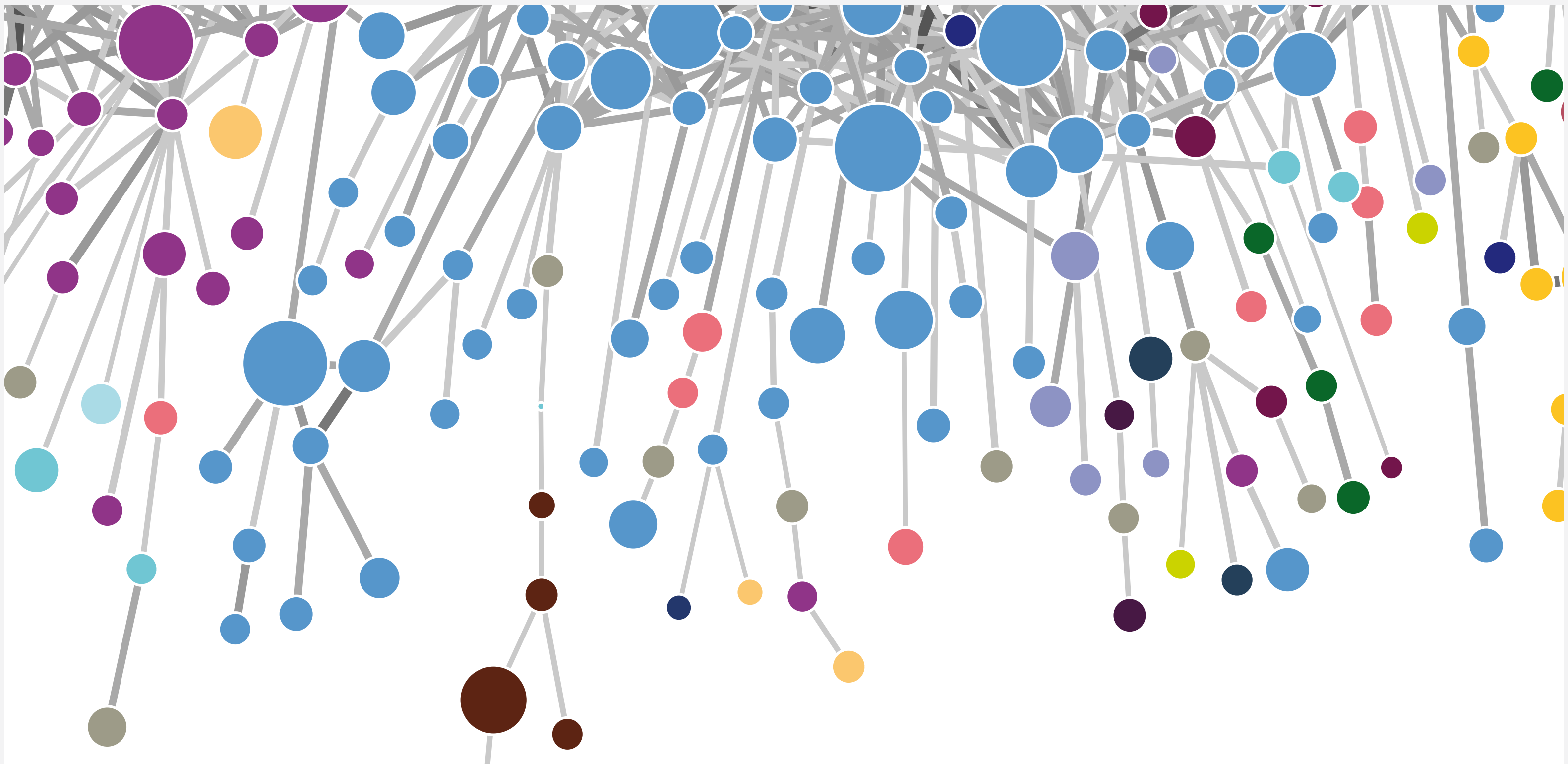
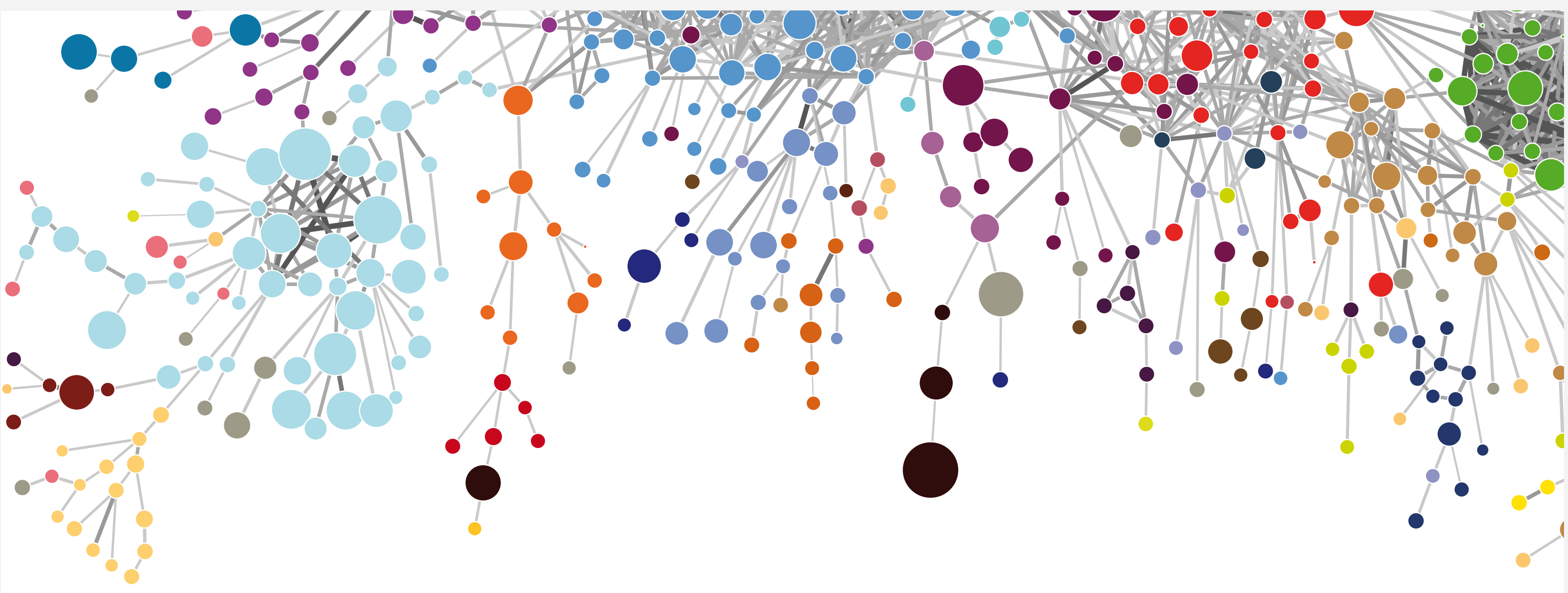


TABLE 4.2.2

	Annualized growth in GDP pc (by decade)					
	(1978-1988, 1988-1998, 1998-2008)					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Initial Income per capita, log	-0.009***	-0.006***	-0.008***	-0.009***	-0.009***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Increase in natural resource exports	0.055***	0.049**	0.055***	0.055***	0.055***	0.061***
- in constant dollars (as a share of initial GDP)	(0.018)	(0.021)	(0.018)	(0.018)	(0.018)	(0.017)
Initial Economic Complexity Index	0.044***		0.046***	0.044***	0.045***	0.052***
	(0.011)		(0.011)	(0.011)	(0.011)	(0.011)
Initial Economic Complexity Index X Income per capita, log	-0.004***		-0.004***	-0.004***	-0.004***	-0.004***
	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)
Years of schooling (standardized)	0.008*	0.013***		0.008***	0.005	
	(0.004)	(0.004)		(0.003)	(0.003)	
Secondary school enrollment (standardized)	0.000	0.001	0.005**		0.002	
	(0.003)	(0.003)	(0.002)		(0.003)	
Tertiary school enrollment (standardized)	-0.003	-0.003	-0.000	-0.003		
	(0.002)	(0.002)	(0.001)	(0.002)		
Constant	0.097***	0.071***	0.091***	0.097***	0.099***	0.083***
	(0.012)	(0.012)	(0.011)	(0.012)	(0.012)	(0.010)
Observations	263	263	263	263	263	263
R^2	0.409	0.288	0.401	0.409	0.406	0.383
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

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



TECHNICAL BOX 4.3: GLOBAL COMPETITIVENESS INDEX AND GROWTH

Here we compare the contribution of the Global Competitiveness Index (GCI) and the ECI to economic growth. We use the ranking of countries in the GCI and the ranking of countries in ECI to predict growth using 5 and 10 year panels. As controls, we use the increase in natural resource exports during the period as well as the logarithm of the initial GDP per capita and year fixed effects. We estimate the contribution of the GCI and ECI to growth by taking the difference between the R^2 obtained for the equation in which they were both included and that in which one or the other is missing.

Table 4.3.1 shows that eliminating the rank of the ECI from the regression results in a much larger loss of explanatory power than removing the rank of GCI. This is true for both 5 and 10 year panels. Figure 4.9 on the main text illustrates these results.

TABLE 4.3.1

	Annualized growth in GDP pc					
	5 year panels			10 year panels		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Ranking of GCI	-0.00007*	-0.00004		-0.00023	-0.00005	
	(0.000)	(0.000)		(0.000)	(0.000)	
Ranking of ECI	-0.00009***		-0.00007***	-0.00064***		-0.00060***
	(0.000)		(0.000)	(0.000)		(0.000)
Initial income per capita, logs	-0.00231***	-0.00092***	-0.00180***	-0.01384***	-0.00439*	-0.01200***
	(0.001)	(0.000)	(0.000)	(0.004)	(0.002)	(0.003)
Increase in natural resource exports	0.00709	0.00077	0.00578	0.09135*	0.04545	0.08814*
- in constant dollars (as a share of initial GDP)	(0.008)	(0.009)	(0.008)			
				(0.048)	(0.049)	(0.045)
Constant	0.02744***	0.01181***	0.02059***	0.17473***	0.06694***	0.15084***
	(0.006)	(0.003)	(0.004)	(0.044)	(0.025)	(0.030)
Observations	104	104	104	83	83	83
Adjusted R^2	0.255	0.123	0.204	0.233	0.074	0.228
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

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

SECTION 5

How Does *Economic Complexity* Evolve?

So economic complexity seems to matter: it affects a country's level of income per capita and drives its future growth. It is also distinct from what is captured by measures of human capital, governance and competitiveness. But how does complexity evolve? How do societies increase the amount of productive knowledge embedded in them? What limits the speed of this process? And why does it happen in some places but not in others?

Again, in our interpretation, the complexity of a country's economy reflects the amount of productive knowledge it contains. This knowledge is costly to acquire and transfer, and is modularized into chunks we call capabilities. Capabilities are difficult to accumulate because doing so creates a complicated chicken and egg problem. On the one hand, countries cannot create products that require capabilities they do not have. On the other hand, there are scant incentives to accumulate capabilities in places where the industries that demand them do not exist. This is particularly true when the missing capabilities required by a potential new industry are many. In this case, supplying any single missing capability will not be enough to launch the new industry, given the absence of the other required capabilities. In a world in which capabilities are complementary, new capabilities risk not being demanded simply because other capabilities are not present.

Consider the following example. A country that does not export fresh produce probably does not have either a cold-storage logistic chain, an expedited green lane at the customs service, or a food safety certification system. All are needed by potential produce exporters. Investors planning to provide the capability to refrigerate and transport fresh produce would not get any traction because they would still need to certify their produce and get through customs fairly quickly. Yet, the systems that provide these services are not in place.

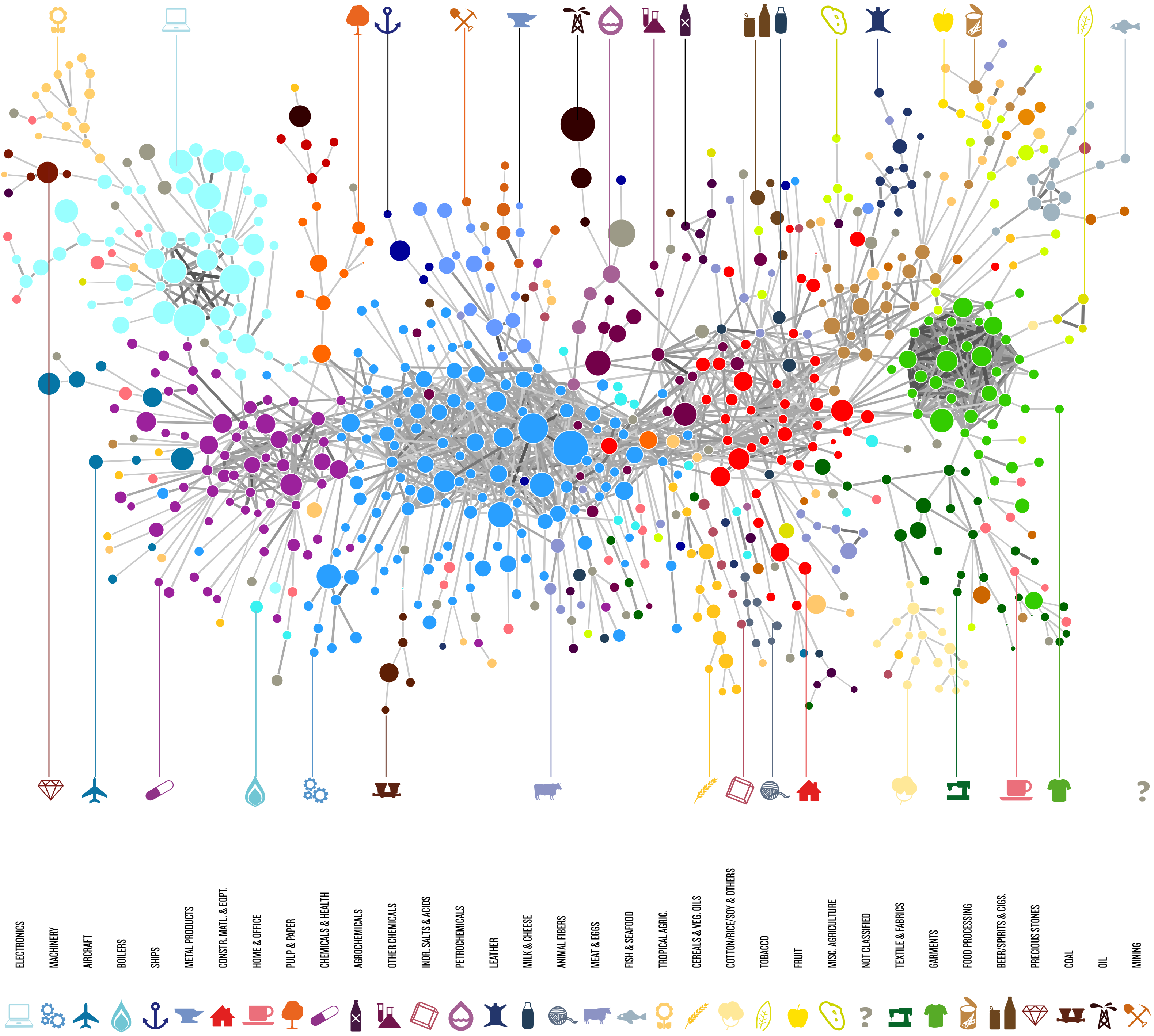
It follows then that new capabilities will be more easily accumulated if they can be combined with others that already exist. This reduces the need to coordinate the accumulation of several new capabilities simultaneously. In our example, if only the green lane at the customs service was missing, it would be easier to develop the fresh produce export industry.

For this reason, countries are more likely to move into products that make use of the capabilities that are already available. These capabilities are available, however, because they are being used to make other products. An implication of this is that a country will diversify by moving from the products they already produce to others that require a similar set of embedded knowledge. Arguably, it is easier to move from shirts to blouses than it is to move from shirts to engines. This is because, in terms of embedded knowledge, shirts are more similar to blouses than to engines. A testable implication of this logic is that countries will move into products that are similar, in terms of the capabilities they require, to the ones they already make.

Measuring the *similarity* in the capability requirements of different products is not simple. Identifying the precise technical and institutional requirements of each product would involve collecting a mindboggling volume of information. Instead, here we measure similarity using a simple trick. If shirts require knowledge that is similar to that required by blouses, but different from that required by engines, then the probability that a country exporting shirts will also export blouses will be higher than the probability that it will

FIGURE 5.1:

The product space.



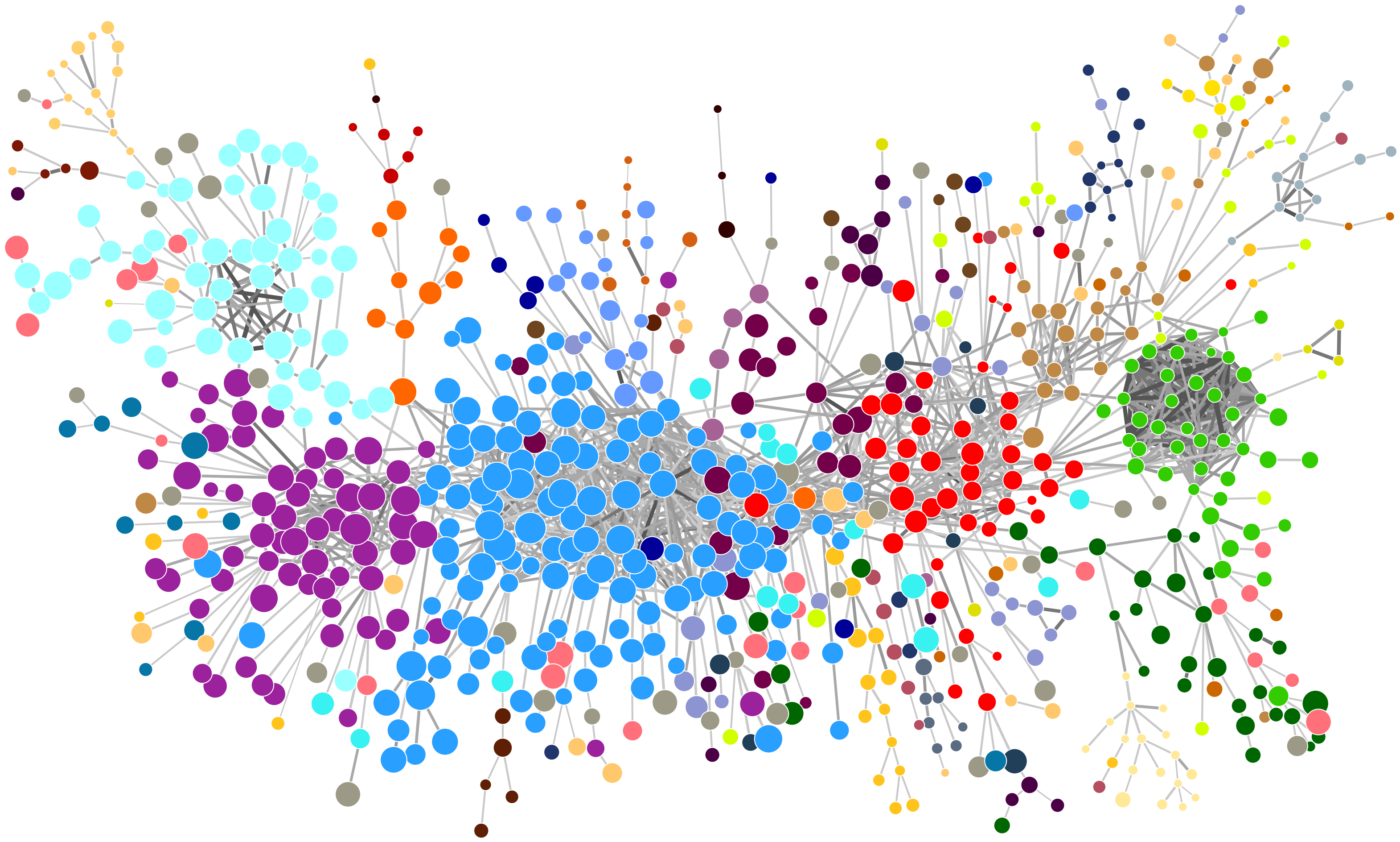
also export engines. So the probability that a pair of products is co-exported carries information about how similar these products are. We use this idea to measure the *proximity* between all pairs of products in our dataset (see *Technical Box* 5.1 on Measuring Proximity). The collection of all proximities is a network connecting pairs of products that are significantly likely to be co-exported by many countries. We refer to this network as the *product space* and use it to study the productive structure of countries.

We care about the structure of the product space because it affects the ability of countries to move into new products. Products that are tightly connected share most of the requisite capabilities. If this is the case, then countries that already have what it takes to make one product will find it relatively easy to move to the next ones. A highly connected product space, therefore, makes the problem of growing the complexity of an economy easier. Conversely, a sparsely connected product space makes it harder.

Once again, a metaphor may help to clarify these ideas. Imagine that the product space is a forest, where every product is a tree. Trees that require similar capabilities are near each other in the forest. Distant trees require very different capabilities. If countries are a collection of firms that make different products, we can think of firms as monkeys that live

FIGURE 5.2:

The product space revisited. The same as *Figure 5.1* but with node sizes proportional to the Product Complexity Index (PCI).



on trees, meaning that they exploit certain products. Countries differ in the number and location of their monkeys in this common forest. The development process, which implies increasing product diversity and complexity, is akin to monkeys colonizing the forest, occupying more trees, and moving especially into the more complex or fruitier ones.

When monkeys jump to nearby trees it minimizes the chicken and egg problem of having to accumulate several missing capabilities at once. Furthermore, if trees are densely packed together it will be relatively easy for monkeys to move from tree to tree and populate the forest. But if trees are far apart, monkeys may be stuck in their current activities. If the product space is heterogeneous, there may be some patches of highly related products, where adding capabilities and expanding into new products would be easier, and other patches of more loosely connected products that make the process of capability accumulation and diversification harder.

What is the shape of the product space we live in? Is it a world in which the forest is dense or sparse? *Figure 5.1* shows a visualization of the product space constructed using international trade data for the years 2006-2008. Here, nodes represent products and their size is proportional to total world trade in that good. Links connect products with a high probability of being co-exported (see *Technical Box 5.2*).

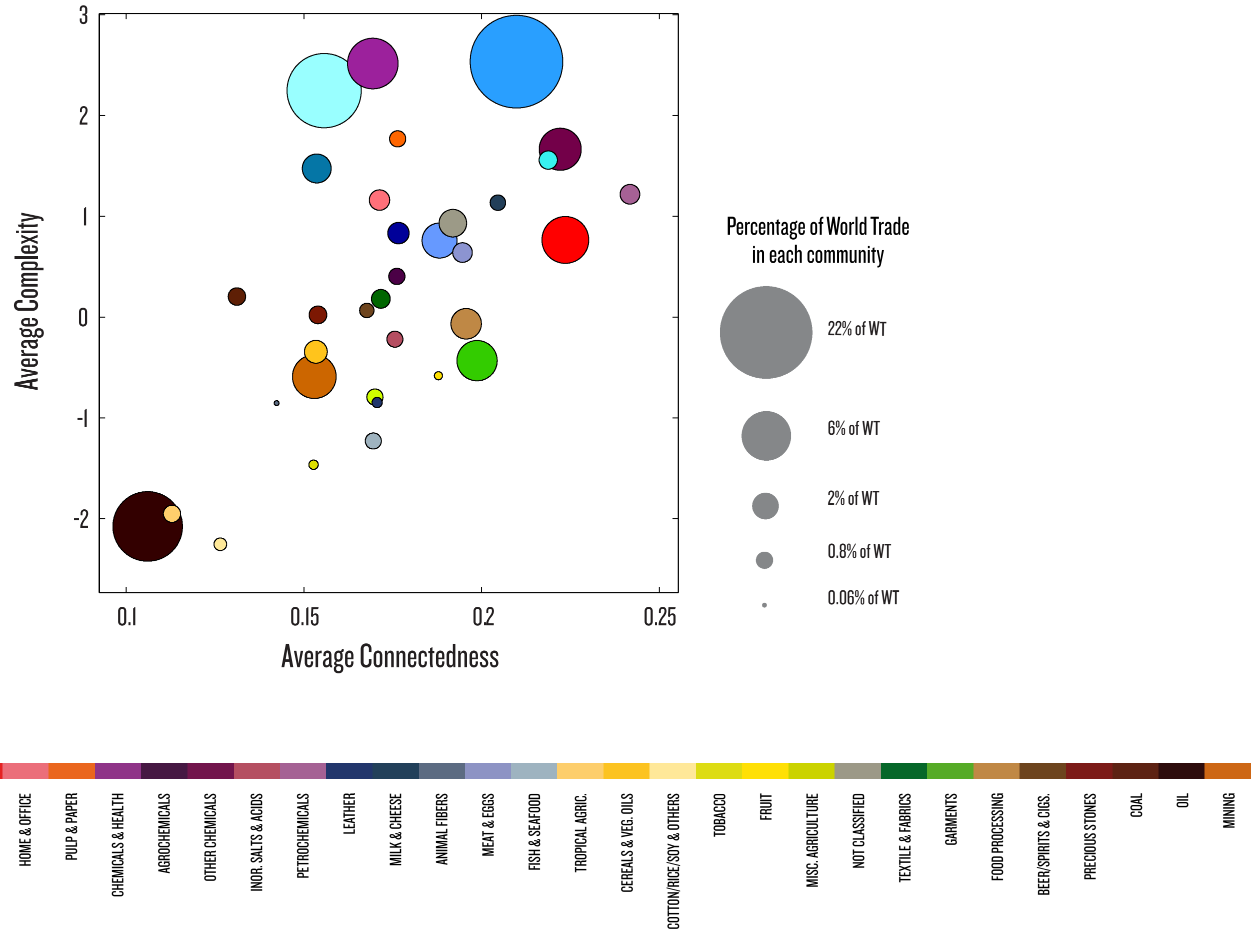
The visualization reveals that the product space is highly heterogeneous. Some sections of it are composed of densely connected groups of products whereas others tend to be more peripheral and sparse.

The product space shows that many goods group naturally into highly connected *communities*. This suggests that products in these communities use a similar set of capabilities. We can identify communities because the products that belong to them are more closely connected to each other than to products outside of the community.Here, we use network science algorithms to discover the communities of products that are hidden in the data (see *Technical Box 5.3* for a discussion of the method). We use these communities to make the discussion of products more tractable. The nearly 800 products in the SITC4 classification were grouped into 34 communities, which we identify by color in our visualization of the product space (*Figure 5.1*). The names, complexity, market size and other characteristics of the communities appear in *Table 5.1*.

Figure 5.2 shows a visualization of the product space that is similar to *Figure 5.1*, but where the size of the nodes is proportional to the complexity of products they represent, as estimated by the Product Complexity Index (PCI). It shows that communities tend to have similar levels of complexity. Products in the Machinery, Electronics and Chemical com-

FIGURE 5.3:

Community characteristics. Average complexity of the products in each community as a function of its connectedness. Bubble size is proportional to the community's participation in world trade.



munities tend to be much more complex than those in the garments cluster or in peripheral communities such as Oil and Tropical Agriculture.

Figure 5.3 shows some of the network characteristics of these communities. *Connectedness* is a measure of how centrally located a community is in the product space. It is the *average proximity* of a community's products to all other products, where *proximity* is the measure of distance between two products used to construct the product space. The figure shows the average connectedness of the products in each community and their average complexity estimated by PCI. The figure reveals a positive relationship between how centrally located the communities are in the product space and how complex their products are. Poorly connected communities such as petroleum, cotton, rice and soybeans tend to be low in complexity. Machinery, by contrast, is very complex and highly connected. Sectors such as garments, textiles and food processing are, on the other hand, in an intermediate position, being connected but not very sophisticated. Electronics and health-related chemicals, however, are also very complex but not as connected as machinery. This suggests they use specific capabilities relevant within their communities but not outside of them.

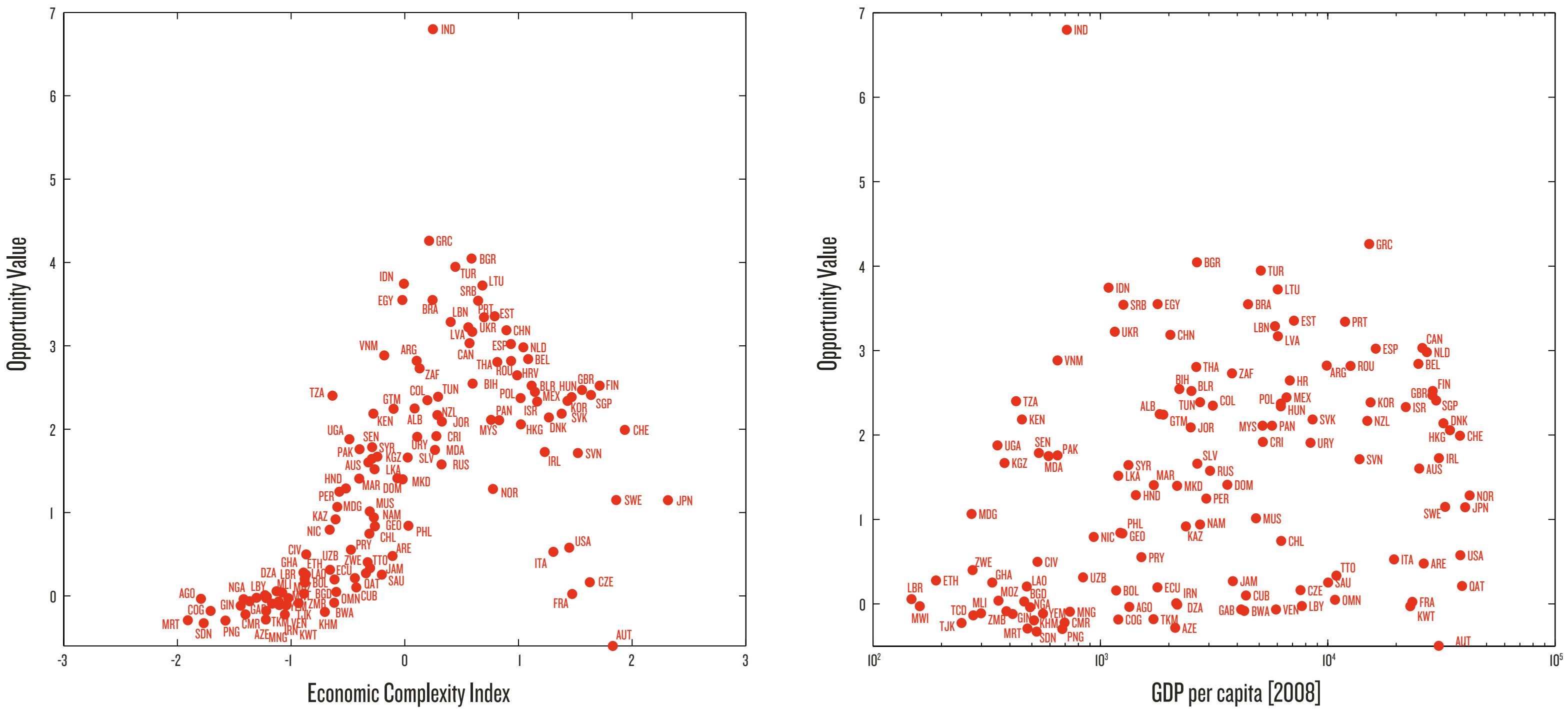
Previous research has shown that the probability that a country will make a new product is strongly related to how close that product is to other products the country already makes. So the location of a country in the product space captures information regarding both the productive knowledge that it possesses and the capacity to expand that knowledge by moving into other nearby products. The ability of countries to diversify and to move into more complex products is crucially dependent on their initial location in the product space.

The product space gives us a glimpse of the embedded knowledge of countries by highlighting the productive capabilities they posses and the opportunities these imply.We can evaluate a country's overall position in the product space by calculating how far it is to alternative products and how complex these products are. We call this measure *opportunity value* and it can be thought of as the value of the option to move into more and more complex products (see *Technical Box* 5.4).

Figure 5.4 compares opportunity value with the Economic Complexity Index and income. It shows that countries with low levels of complexity tend to have few opportunities available. This is because the products they do create tend to be peripheral in the product space. Complex economies tend to have few remaining opportunities because they already occupy a large fraction of the better part of the product space. Countries with an intermediate level of complexity, on the other hand, differ largely in their opportunity value. Some

FIGURE 5.4:

Opportunity Value as a function of the Economic Complexity Index and GDP per capita.



countries like Saudi Arabia, Jamaica and Chile, are located in parts of the product space that imply few opportunities. Others, like India, Greece, Turkey, Brazil and Indonesia are located in parts of the product space where opportunities are plentiful. *Figure* 5.4 also shows that countries with similar incomes face dramatically different opportunities.

Finally, we illustrate how countries move through the product space by looking at Ghana, Poland, Thailand and Turkey (*Figure* 5.5). Here, square nodes with thick black outlines are used to indicate the products that each of these countries was exporting at each point in time. In all cases we see that new industries –new black squares– tend to lie close to the industries already present in these countries. The productive transformation undergone by Poland, Thailand and Turkey, however, look striking compared to that of Ghana. Thailand and Turkey, in particular, moved from mostly agricultural societies to manufacturing powerhouses during the 1975-2009 period. Poland, also "exploded" towards the center of the product space during the last two decades, becoming a manufacturer of most products in both the home and office and the processed foods community and significantly increasing its participation in the production of machinery. These transformations imply an increase in embedded knowledge that is reflected in our Economic Complexity Index. Ultimately, it is these transformations that underpinned the impressive growth performance of these countries.

We started this section asking several questions: How does complexity evolve? And how do societies increase the total amount of productive knowledge embedded in them? Here we have shown that countries expand their productive knowledge by moving into nearby goods. This increases the likelihood that the effort to accumulate any additional capability will be successful, as the complementary capabilities needed to make a new product are more likely to be present in the production of the nearby goods.

What limits the speed of this process? Since capabilities are useful only when combined with others, the accumulation of capabilities is slowed down by the chicken and egg problem. New products may require capabilities that do not exist precisely because the other products that use them are not present. Moreover, since capabilities are chunks of tacit knowledge, accumulating them is difficult even when there is demand for them, because the country does not have any exemplars to copy.

Most important, why does this process of development occur in some places, but not in others? There are many answers to this question. Our approach adds an alternative answer by showing that a country's position in the product space determines its opportunities to expand its productive knowledge and increase its level of economic complexity. But the product space is highly heterogeneous, placing countries in radically different settings. Ultimately, development is the expression of the total amount of productive knowledge that is embedded in a society. But the process by which this knowledge is accumulated has a structure that we are only now starting to understand.

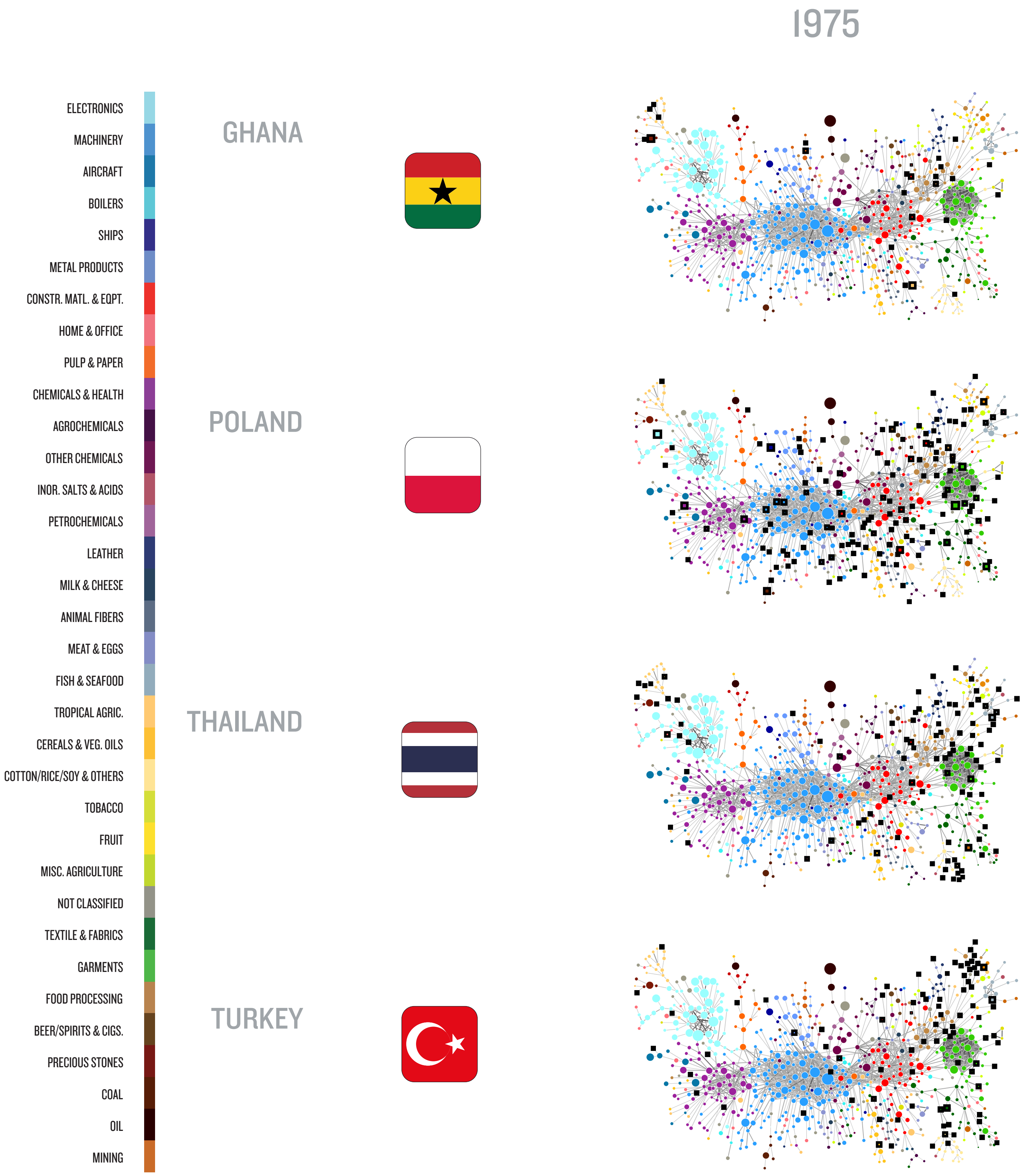
TABLE 5.1:

CHARACTERISTICS OF PRODUCT COMMUNITIES

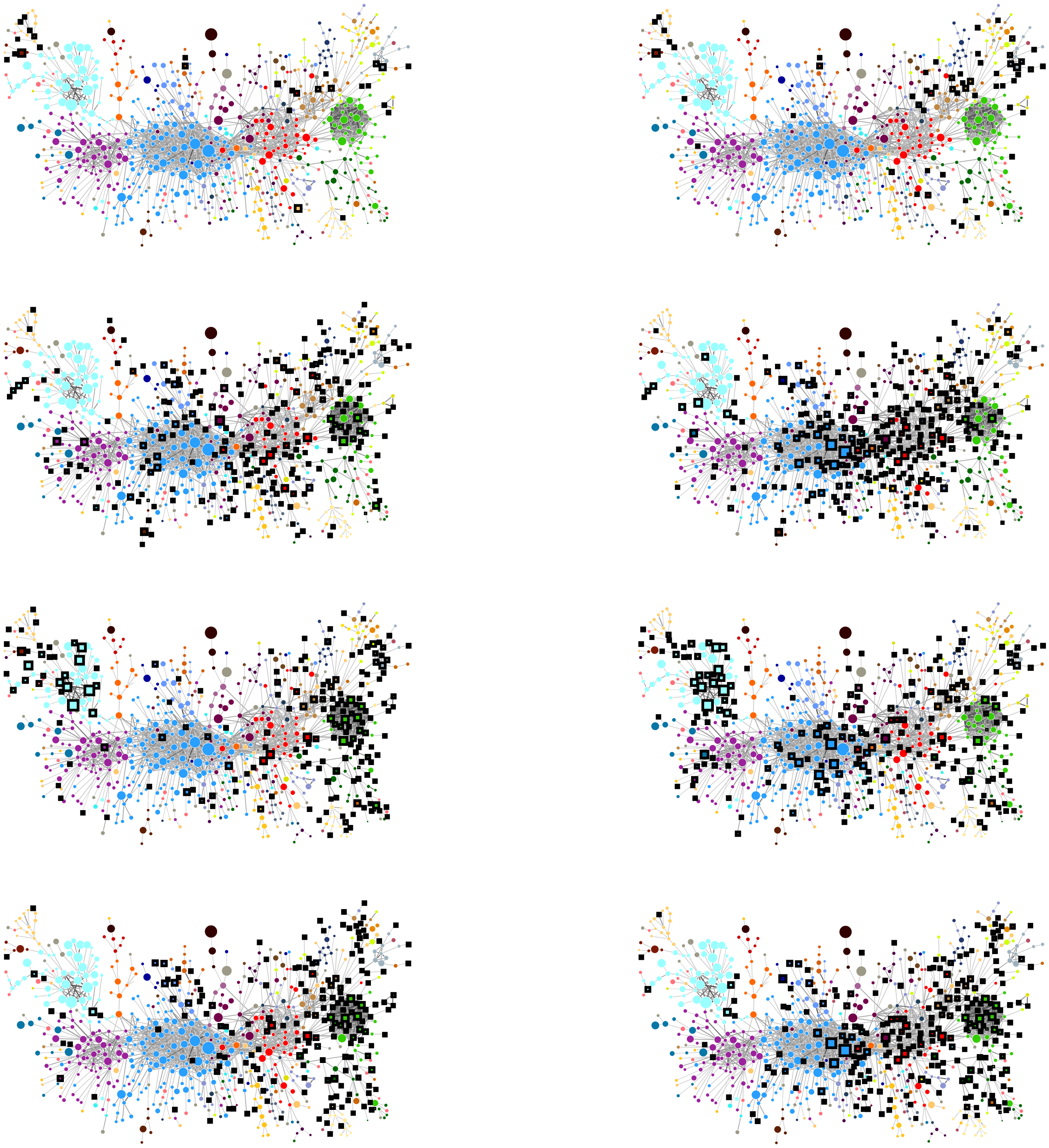
Community Name	Average PCI	Number of Products	World Trade	World Share	Top 3 Countries by Export Volume	Top 3 Countries by Number of Products (RCA>1)
Machinery	2.54	125	4.4T	20.29%	DEU, USA, JPN	DEU, ITA, AUT
Electronics	2.25	52	3.6T	16.71%	CHN, HKG, USA	CHN, HKG, MYS
Oil	-2.08	4	2.3T	10.49%	SAU, RUS, NOR	EGY, KAZ, DZA
Chemicals & Health	2.52	64	1.6T	7.47%	USA, DEU, BEL	USA, BEL, DEU
Other Chemicals	1.67	24	1.2T	5.49%	DEU, USA, FRA	DEU, ITA, ESP
Construction Materials & Equipment	0.77	44	1.1T	5.23%	CHN, DEU, ITA	CZE, POL, SVN
Mining	-0.59	48	1.1T	5.01%	AUS, USA, CHL	CAN, AUS, KAZ
Garments	-0.43	42	1.1T	4.63%	CHN, HKG, ITA	CHN, VNM, TUN
Food Processing	-0.07	26	603B	2.74%	DEU, ITA, USA	SRB, ESP, BEL
Metal Products	0.76	17	496B	2.26%	JPN, DEU, KOR	ZAF, UKR, SVK
Aircraft	1.48	10	440B	2.00%	FRA, DEU, GBR	CAN, GBR, FRA
Not Classified	0.93	36	426B	1.94%	USA, CHN, DEU	CHN, FRA, GBR
Cereals & Vegetable Oils	-0.34	21	295B	1.34%	USA, BRA, ARG	PRY, MDA, ARG
Home & Office	1.16	23	250B	1.14%	CHN, CHE, USA	CHN, PAN, PRT
Meat & Eggs	0.64	23	242B	1.10%	USA, BRA, DEU	FRA, BEL, POL
Ships	0.83	8	232B	1.05%	KOR, CHN, JPN	ROU, POL, HRV
Petrochemicals	1.22	5	220B	1.00%	DEU, USA, BEL	PRT, BEL, FRA
Boilers	1.56	14	193B	0.88%	CHN, DEU, JPN	CHN, TUR, KOR
Fish & Seafood	-1.23	11	191B	0.87%	CHN, NOR, THA	CHL, NAM, SYC
Textile & Fabrics	0.18	32	189B	0.86%	CHN, ITA, HKG	CHN, TUR, IND
Tropical Agriculture	-1.95	16	190B	0.86%	IDN, NLD, MYS	IDN, CIV, CRI
Coal	0.21	6	183B	0.83%	AUS, IDN, RUS	CZE, COL, RUS
Misc Agriculture	-0.79	22	170B	0.78%	BRA, DEU, FRA	ESP, TZA, NIC
Precious Stones	0.02	4	170B	0.77%	IND, ISR, BEL	GBR, LBN, LKA
Pulp & Paper	1.77	11	148B	0.67%	USA, CAN, SWE	SWE, FIN, CAN
Agrochemicals	0.40	13	141B	0.64%	DEU, USA, CAN	BEL, JOR, DEU
Milk & Cheese	1.14	7	134B	0.61%	DEU, FRA, NLD	NLD, BLR, LTU
Beer, Spirits & Cigarettes	0.07	6	124B	0.57%	GBR, NLD, DEU	JAM, BEL, NLD
Inorganic Salts & Acids	-0.22	10	117B	0.53%	USA, CHN, DEU	ISR, JOR, USA
Cotton, Rice, Soy & Others	-2.25	18	96B	0.44%	USA, IND, THA	TZA, MOZ, GRC
Tobacco	-1.46	6	64B	0.29%	DEU, NLD, BRA	PHL, GRC, SEN
Leather	-0.85	14	53B	0.24%	ITA, USA, HKG	ALB, SOM, ESP
Fruit	-0.58	4	45B	0.21%	ESP, USA, CHL	NLD, LBN, LTU
Animal Fibers	-0.85	7	12B	0.06%	AUS, CHN, ITA	URY, NZL, ZAF

FIGURE 5.5:

The evolution of Ghana, Poland, Thailand and Turkey in the product space: 1975, 1990 and 2009. Black boxes indicate the products in which these countries had RCA>1.



1990 2009



TECHNICAL BOX 5.1: MEASURING PROXIMITY

To make products you need chunks of embedded knowledge which we call capabilities. The capabilities needed to produce one good may or may not be useful in the production of other goods. Since we do not observe capabilities directly, we create a measure that infers the similarity between the capabilities required by a pair of goods by looking at the probability that they are co-exported. To quantify this similarity we assume that if two goods share most of the requisite capabilities, the countries that export one will also export the other. By the same token, goods that do not share many capabilities are less likely to be co-exported.

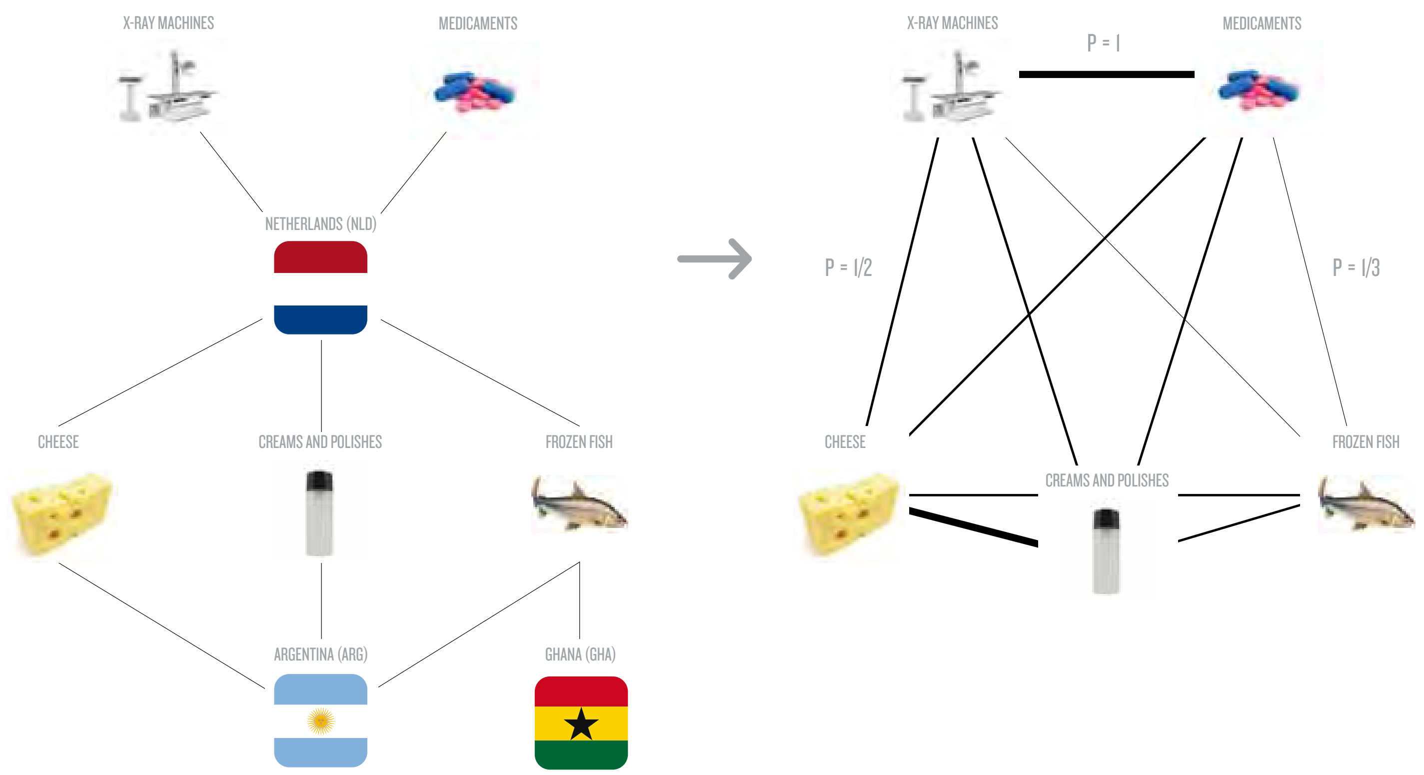
Our measure is based on the conditional probability that a country that exports product p will also export product p' (Figure 5.1.1). Since conditional probabilities are not symmetric we take the minimum of the probability of exporting product p, given p' and the reverse, to make the measure symmetric and more stringent. For instance, in the year 2008, 17 countries exported wine, 24 exported grapes and 11 exported both, all with RCA>1. Then, the proximity between wines and grapes is 11/24=0.46. Note that we divide by 24 instead of 17 to minimize false positives. Formally, for a pair of goods p and p' we define proximity as:

$$\phi_{pp'} = \frac{\sum_c M_{cp} M_{cp'}}{max(k_{p,0}, k_{p',0})}$$

Where M_{cp} =1 if country c exports product p with RCA>1 and 0 otherwise. $k_{p,0}$ is the ubiquity of product p.

FIGURE 5.1.1:

An illustrative example for the proximity measure.



TECHNICAL BOX 5.2: VISUALIZING THE PRODUCT SPACE

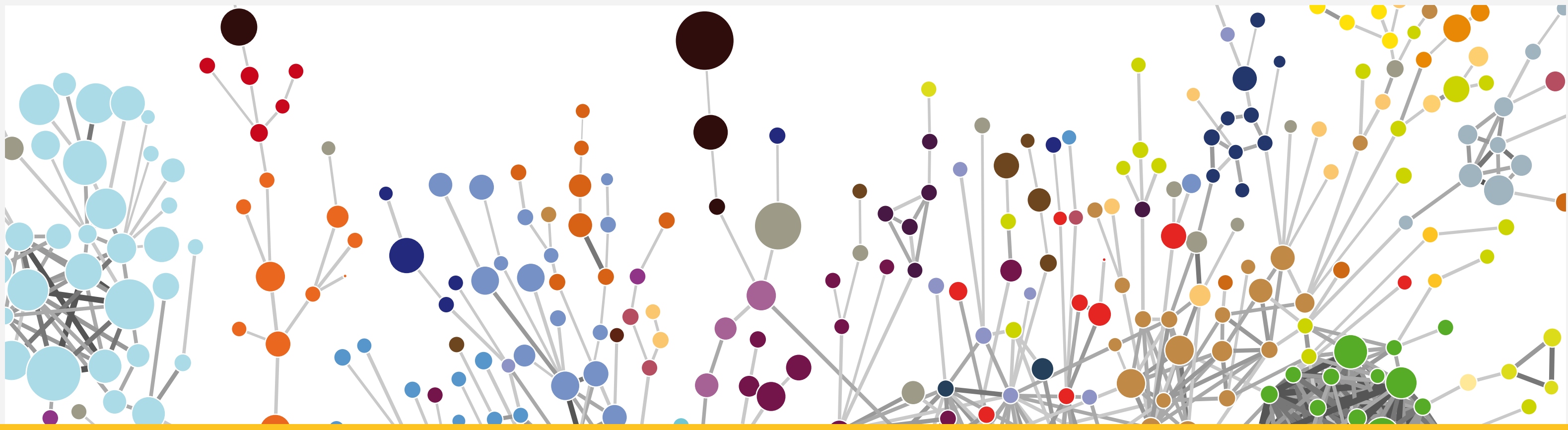
To visualize the product space we use some simple design criteria. First, we want the visualization of the product space to be a connected network. By this, we mean avoiding islands of isolated products. The second criteria is that we want the network visualization to be relatively sparse. Trying to visualize too many links can create unnecessary visual complexity where the most relevant connections will be occluded. This criteria is achieved by creating a visualization in which the average number of links per node is not larger than 5 and results in a representation that can summarize the structure of the product space using the strongest 1% of the links.

To make sure the visualization of the product space is connected, we calculate the maximum spanning tree (MST) of the proximity matrix. MST is the set of links that connects all the nodes in the network using a minimum number of connections and the maximum possible sum of proximities. We calculated the MST using Kruskal's algorithm. Basically the algorithm sorts the values of the proximity matrix in descending order and then includes links in the MST if and only if they connect an isolated product. By definition, the MST includes all products, but the number of links is the minimum possible.

The second step is to add the strongest connections that were not selected for the MST. In this visualization we included the first 1,006 connections satisfying our criterion. By definition a spanning tree for 774 nodes contains 773 edges. With the additional 1,006 connections we end up with 1,779 edges and an average degree of nearly 4.6.

After selecting the links using the above mentioned criteria we build a visualization using a Force-Directed layout algorithm. In this algorithm nodes repel each other, just like electric charges, while edges act as spring trying to bring connected nodes together. This helps to create a visualization in which densely connected sets of nodes are put together while nodes that are not connected are pushed apart.

Finally, we manually clean up the layout to minimize edge crossings and provide the most clearly representation possible.



TECHNICAL BOX 5.3: IDENTIFYING PRODUCT COMMUNITIES

In network science, groups of highly interconnected nodes are known as communities. In the Product Space, communities represent groups of products that are likely to require many of the same capabilities.

We assign products to communities using the algorithm introduced by Rosvall and Bergstrom (2008). This algorithm finds communities using a two step process. First, it explores the network using a collection of random walkers. The intuition behind this first step is that nodes belonging to the same community are more likely to lie close by in the sequence of nodes visited by a random walker. For instance, take *photographic film, photographic chemicals* and *silicones*. These are three products that are interconnected and belong to a densely connected region of the product space. Hence, the random walker is much more likely to go through the sequences {*silicones, photographic chemicals, photographic film*} or {*photographic film, silicones, photographic chemicals*} than {*photographic film, grapes, blouses*}. The emergence of these sequences indicates that photographic film, photographic chemicals and silicones, probably belong to the same community. After several iterations of random walks have been recorded, the algorithm tries to compresses these sequences by looking for ways to rename nodes and minimize the amount of space required to store information about these sequences. For instance, if *silicones, photographic films,* and *photographic chemicals* are grouped into a community called photographic materials this would allow compressing the sequence by replacing each time it appears by a reference to that community. The algorithm looks for a compression that preserves as much information as possible. This avoids the trivial solution in which all products are assigned to the same community.

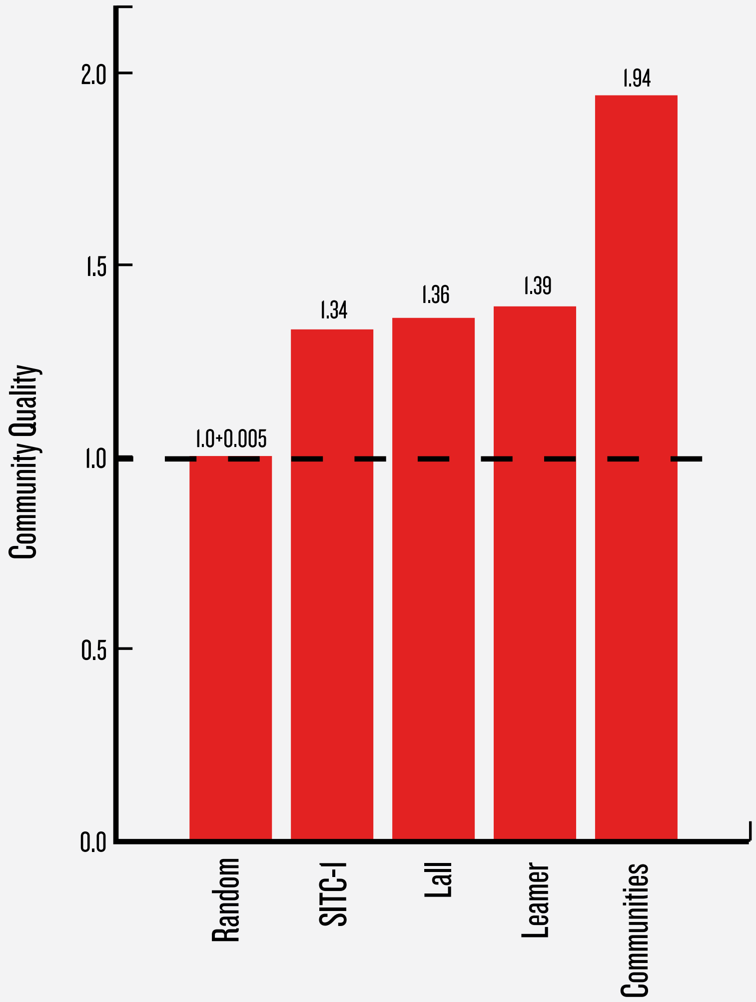
The communities determined through this algorithm were manually named and merged into 34 communities (see Table 5.1 for details).

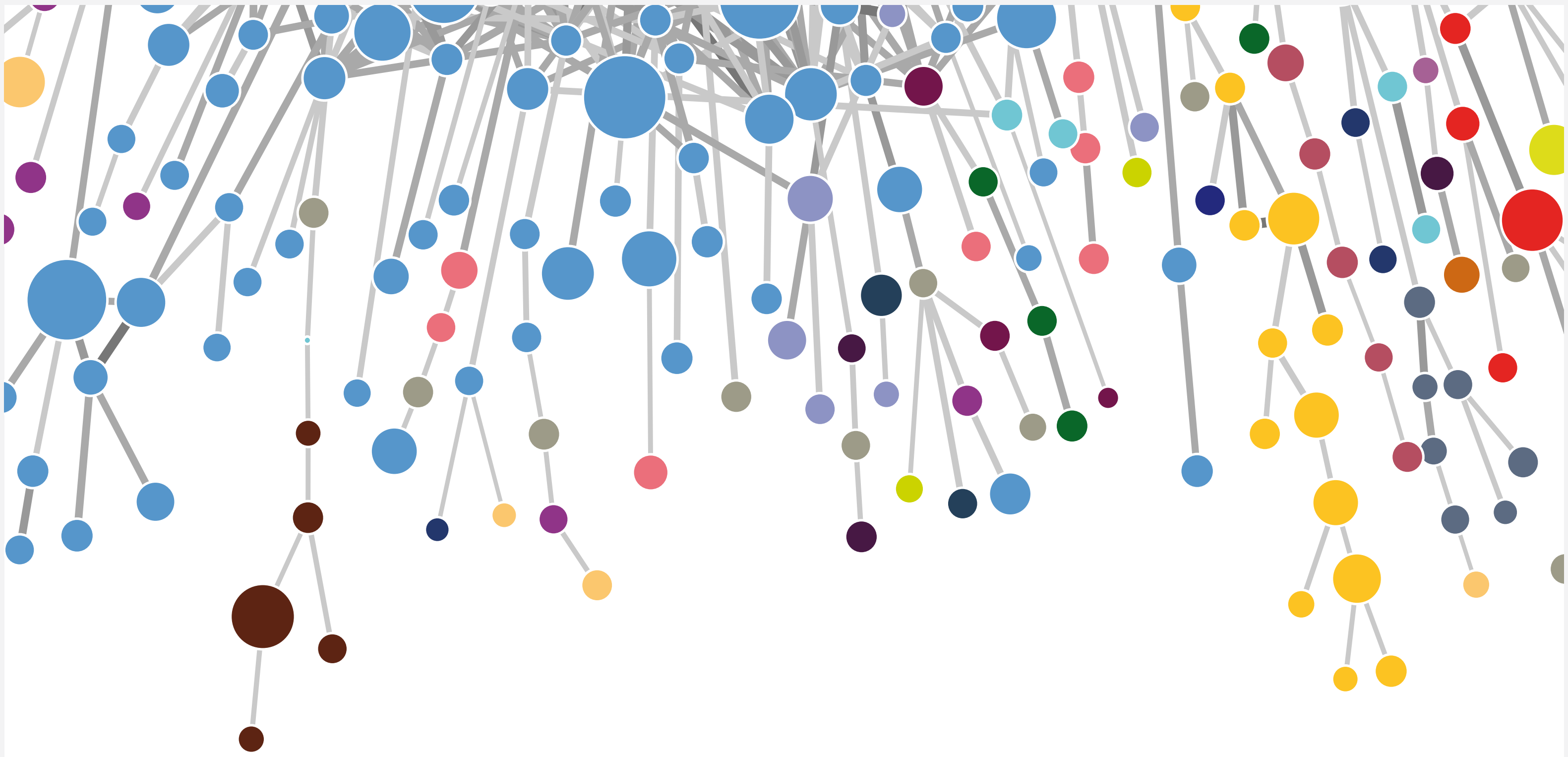
We compare the ability of these communities to summarize the structure of the product space by introducing a measure of *community quality*. This is the ratio between the average proximity of the links within a community, and those connecting products from that community to products in other communities.

To get a sense of the community quality we compare our assignment of products into communities with a baseline null model and three popular categorizations. The baseline null model is given by an ensemble of communities of the same size, where nodes have been assigned to each community at random. In this case, the average strength of the links within communities is equal to the average strength of links between communities, and the community quality is 1. The three categorizations we use as comparators are: the first digit of the Standard International Trade Classification, the categories introduced by Leamer (1984) -based on factor intensities- and the technology categories introduced by Lall (2000). All three classifications produce values of the community quality between 1.3 and 1.4, indicating that links within communities tend to be, on average, 30% to 40% stronger than those between communities. The communities we propose here have a community quality value of 1.94, indicating that the links between nodes in the same community are, on average, 94% stronger than those connecting nodes between communities (Figure 5.3.1). The difference in community quality of our proposed community system and that of the three alternative categorizations is highly statistically significant with a p-value$<1\times10^{-30}$.

FIGURE 5.3.1:

Community quality for five different ways of grouping products.





TECHNICAL BOX 5.4: UNDERSTANDING THE POSITION OF COUNTRIES IN THE PRODUCT SPACE: DISTANCE, OPPORTUNITY VALUE AND OPPORTUNITY GAIN

Empirically, we find that 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. Proximity measures the similarity between a pair of products, so we need another measure to quantify the distance between the products that a country makes and each of the products that it does not. We call this measure distance and define it as the sum of the proximities connecting a new good p to all the products that country c is not currently exporting. We normalize distance 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 0. 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). Formally,

$$d_{cp} = \frac{\sum_{p'}(1 - M_{cp'})\phi_{pp'}}{\sum_{p'}\phi_{pp'}}$$

Distance gives us an idea of how far each product is given a country's current mix of exports. Yet, it would be useful to have a holistic measure of the opportunities implied by a country's position in the product space. Countries that make products that 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. We can think of this as the value of the option to move into other products.

Hence, to quantify the "opportunity value" of a country's unexploited prospects we can add the level of complexity of the products that it is not currently weighted by how close these products are to the country's current export suite. We can write this mathematically as:

$$opportunity\ value_c = \sum_{p'}(1 - d_{cp'})(1 - M_{cp'})PCI_{p'}$$

Where PCI is the Product Complexity Index of product p'. The term $1 - M_{cp'}$ makes sure that we count only the products that the country is not currently producing. Higher opportunity value implies being in the vicinity of more products and/or of products that are more complex.

We can use opportunity value to calculate the potential benefit to a country if it were to move to a particular new product. We call this 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 the doors to more and more complex products. Formally, we can write the opportunity gain as:

$$opportunity\ gain_c = \sum_{p'}\left(\frac{\phi_{pp'}}{\sum_p \phi_{pp'}}(1-M_{cp'})PCI_{p'} - (1-d_{cp}PCI_p)\right)$$

How Can This Atlas Be Used?

SECTION 6

In recent years there has been an explosion in the use of indexes to rank countries, corporations, cities and other entities in terms of a myriad of indicators. We have the Human Development Index, put together by the United Nations Development Program (UNDP), the Freedom House's Freedom Index, the World Bank's Doing Business, Investment Climate Assessment and Governance indexes, the World Economic Forum's Competitiveness, Gender Gap and Trade Facilitation indexes, the OECD's Programme for International Student Assessment (PISA), the credit ratings of governments and corporations put together by Standard & Poor's, Moody's and Fitch, and the Quality of Development Assistance published by the Center for Global Development.

The popularity and usefulness of these efforts reflect the fact that they provide feedback about the relative performance of an entity on dimensions that societies find important and that the publishers want to highlight. The indexes may help the rated entities to evolve towards better outcomes by benchmarking current performance, vis-a-vis the best achievers, and by encouraging learning through imitation and experimentation around best practices.

Here it may be useful to distinguish between the two roles of any indicator. The first one is as a measure and the second one is as a guide to action. As a measure, the critical question is whether the index is able to adequately represent the information that one wants to capture, whether it is freedom, human development, solvency, business environment, governance, educational quality or, in our case, productive knowledge. As a guide to action the question is what interventions will lead an entity to improve its performance on the index.

With respect to the latter question, indexes differ in how clearly they map to a plan of action. At one extreme, the World Bank's Doing Business Index is based on the statutory requirements to perform certain functions such as registering and closing a business or obtaining a license. Here the course of action is clear: change the statutes in a direction that would improve the re-calculated index. However, in other cases, such as improving rule of law, educational quality, gender equality, the control of corruption or competitiveness, the mapping is not so clear. It is hard to know precisely which changes in current practice would be reflected in better performance. This is a general problem in many areas where the relationship between action and performance is complex. Should schools reduce class size, improve teacher training, increase nutritional assistance or implement standardized exams in order to improve educational quality? Should societies get more policemen, revamp the judicial system, encourage press freedom, revise their gun laws or confront organized crime in order to improve rule of law? The best humans have been able to do is to emulate some of the features of the best performers in the hopes that the elements that were selected for imitation may prove crucial to the achievement of better outcomes. If they do, this will hopefully be reflected in future measures of the index. If they do not, other actions will be tried until the action taken is consistently followed by a change in the desired direction.

In our case, the Economic Complexity Index (ECI) is based on the number and the complexity of the products that a country exports with comparative advantage. Empirically, countries that do well in this index, given their income level, tend to achieve higher levels of economic growth. The ability to successfully export new products is a reflection of the fact that the country has acquired new productive knowledge that will then open up further opportunities for progress. What a country needs to do to achieve this will be highly specific to the context of the country and the product. Better seeds could cause an agricultural revolution; improved infrastructure could open up new possibilities for light manufactures; clarifying property rights and human subject regulations may allow for participation in pharmaceutical research; changing the responsiveness of training institutions to the needs of new sectors may unleash their growth; etc.

The Economic Complexity Index is not easy to manipulate through a narrow set of decisions. Ultimately, countries improve on the index by being able to increase the number of different activities they can successfully engage in and by moving towards activities that are more complex. The policy message for most countries is clear: create an environment where a greater diversity of productive activities can thrive and, in particular, activities that are relatively more complex. Countries are more likely to succeed in this agenda if they focus on products that are close to their current set of productive capabilities, as this would facilitate the identification and provision of the missing capabilities. The ECI therefore, is accompanied by maps that help chart the opportunities and rewards available for each country. These are maps that are specific to each country and do not represent one-size-fits-all development advice.

These maps, however, could also be used by firms searching for a new location or looking to diversify into a new product. These maps carry information about the productive capabilities that are present in a given country and the degree to which these capabilities are relevant to support a particular new industry.

A map does not tell people where to go, but it does help them determine their destination and chart their journey towards it. A map empowers by describing opportunities that would not be obvious in the absence of it. If the secret to development is the accumulation of productive knowledge, at a societal rather than individual level, then the process necessarily requires the involvement of many explorers, not just a few planners. This is why the maps we provide in this Atlas are intended for everyone to use.

Which Countries Are Included in This Atlas?

SECTION 7

Countries are highly heterogeneous. When it comes to the size of their population, territory, income and economy, countries differ by orders of magnitude. When it comes to land, Russia is 1,000 times larger than Kuwait. When it comes to population, China is more than 600 times more populous than Slovenia. When it comes to Gross Domestic Product, the United States is more than 1,300 times larger than Namibia. All of these are countries that made it into this Atlas, illustrating the large cross section of the world captured in this book.

Products also differ enormously in terms of their world market size. Depending on the year, crude oil represents five to ten percent of world trade while goat skins represent less than one part in one hundred thousand of world trade. To make countries and products comparable we control for the size of the country and of the product by calculating their Revealed Comparative Advantage (see *Technical Box* 2.2). This means that large and small countries and products with big and small markets count the same as far as our method is concerned. Moreover, the data of each country affects the calculations of all others so including data that is noisy or unreliable greatly affects the integrity of our calculations.

Countries that are too small in terms of their export base, such as Tuvalu or Vanuatu, or with data that is highly unreliable or not adequately classified, do not provide us with a sufficiently broad sample to infer their structure. This is why we restrict the analysis to 128 countries, which account for 99% of world trade, 97% of the world's total GDP and 95% of the world's population.

To generate this list we used a variety of criteria. First, we limit ourselves only to the set of countries for which there is product-level trade data available in the UN COMTRADE and income data available for 2008. Second, we only use data on countries with a population above 1,200,000. Third, we only consider countries that exported at least 1 billion dollars per year, on average, between 2006 and 2008. Finally we remove from this sample Iraq, Macau and Chad, three countries with severe data quality issues (*Figure 7.1*).

FIGURE 7.1:

Schematic of the procedure used to determine the countries that were included in the Atlas.