SECTION 2

How Do We Measure *Economic Complexity*?

As we have argued, productive knowledge is the key to prosperity. Larger amounts of productive knowledge require increasingly complex webs of human interaction, which we call economic complexity. In this Section we develop measures of the amount of productive knowledge held by different societies. How can we go about doing this, given that there are no direct ways to look at a country and know how much knowledge is embedded in it? Our approach is based on the following trick: we can look at what countries make, and from this, we can begin to infer what a country knows.

We can observe how many different kinds of products a country is able to make. We call this the **diversity** of a country (Figure 2.1). We can also observe the number of countries that are able to make a product. We call this the **ubiquity** of a product (Figure 2.1). We assume that countries are only able to make the products for which they have the requisite knowledge. From this simple claim, it is possible to extract a few implications that can be used to construct a measure of economic complexity.

The 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 **A, C** and **R** to construct the words **CAR** or **ARC.** In this analogy, words are like products and letters are like capabilities, or modules of embedded knowledge. We assume that each player has plenty of copies of the letters that they do have. This means that if a country has a certain module of knowledge, it can use that knowledge in many different settings. Our challenge is to measure the number of different letters the players have by looking at two things: first, the number of words that each player can write; second, the number of players who can write a particular word.

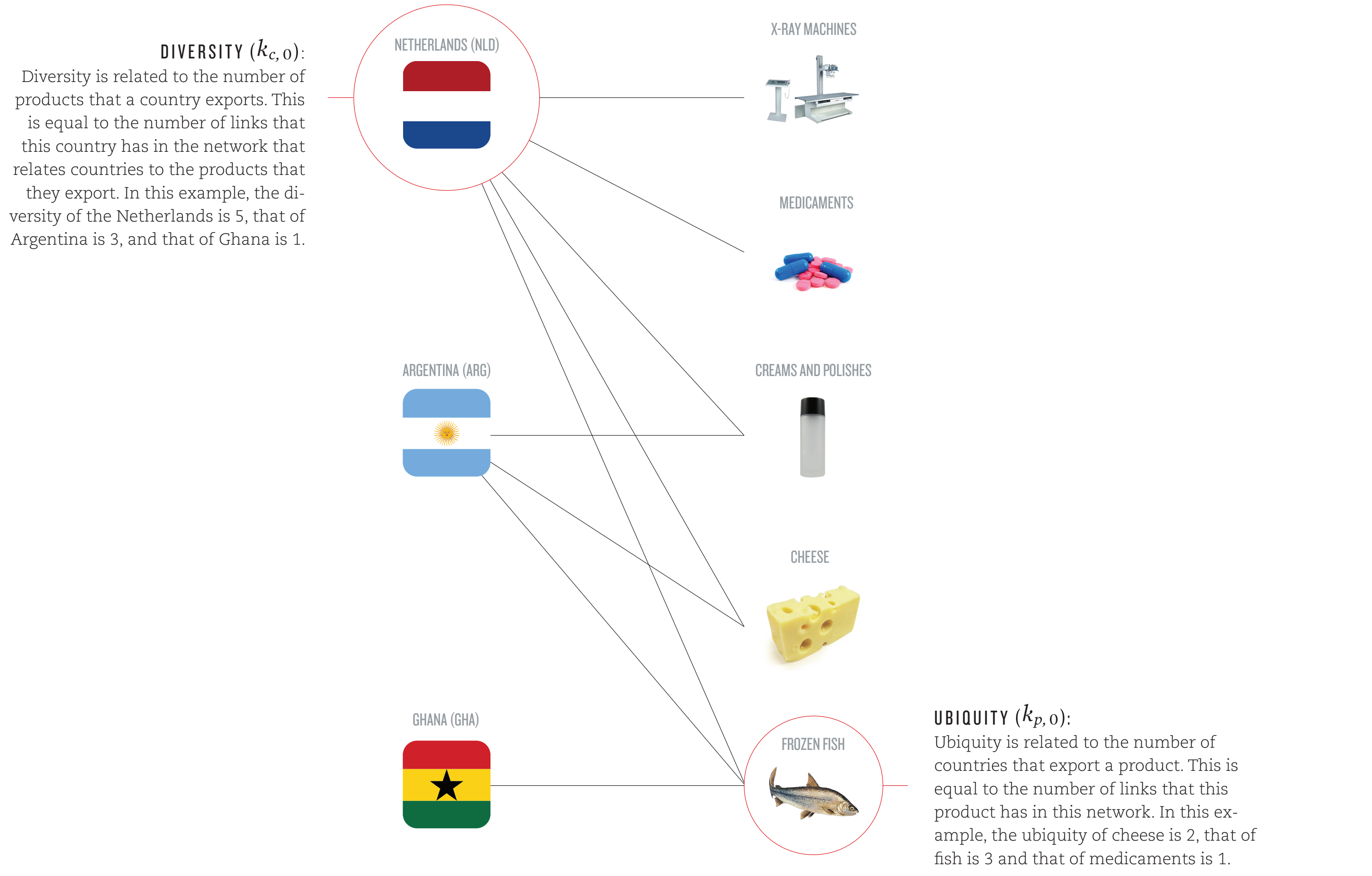
Players who have more letters should be able to make more words. We can expect the diversity of words (products) that players (countries) can make to be strongly related to the number of letters (capabilities) that they have. Hence, diversity is a first measure of how much knowledge a country has.

Let us look now at words. The number of players who can make a word is indicative of how many letters the word has. Longer words will tend to be less common, since they can only be put together by players who have all the requisite letters. Similarly, more complex products will be less common because only the countries that have all the requisite knowledge will be able to make them. Products that require little knowledge should be more ubiquitous and vice versa.

The diversity of a country's exports is a crude approximation of the variety of capabilities available in the country, just as the ubiquity of a product is a crude approximation of the variety of capabilities required by a product. Consider medical imaging devices. These machines are made in few places, and the countries that are able to make them, such as the United States or Germany, also export a large number of other products. From this we can infer that medical imaging devices are complex because few countries make them and those that do tend to be diverse. 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? Not at all! If raw diamonds were complex, then the countries that extract diamonds should also be able to make many other things because they would have the many capabilities required by diamonds. We see though that Sierra Leone and Botswana principally export diamonds. This indicates that, unlike medical imaging devices, something other than large volumes of knowledge makes diamonds rare. Both of these measures are affected by the existence of rare capabilities, which, using the Scrabble analogy, would be represented by letters like Q and X. So, here we have used the diversity of the countries making a product (say, diamonds) to nuance the first impression given by the (low) ubiquity of the product.

FIGURE 2.1:

Graphical representation of diversity and ubiquity.



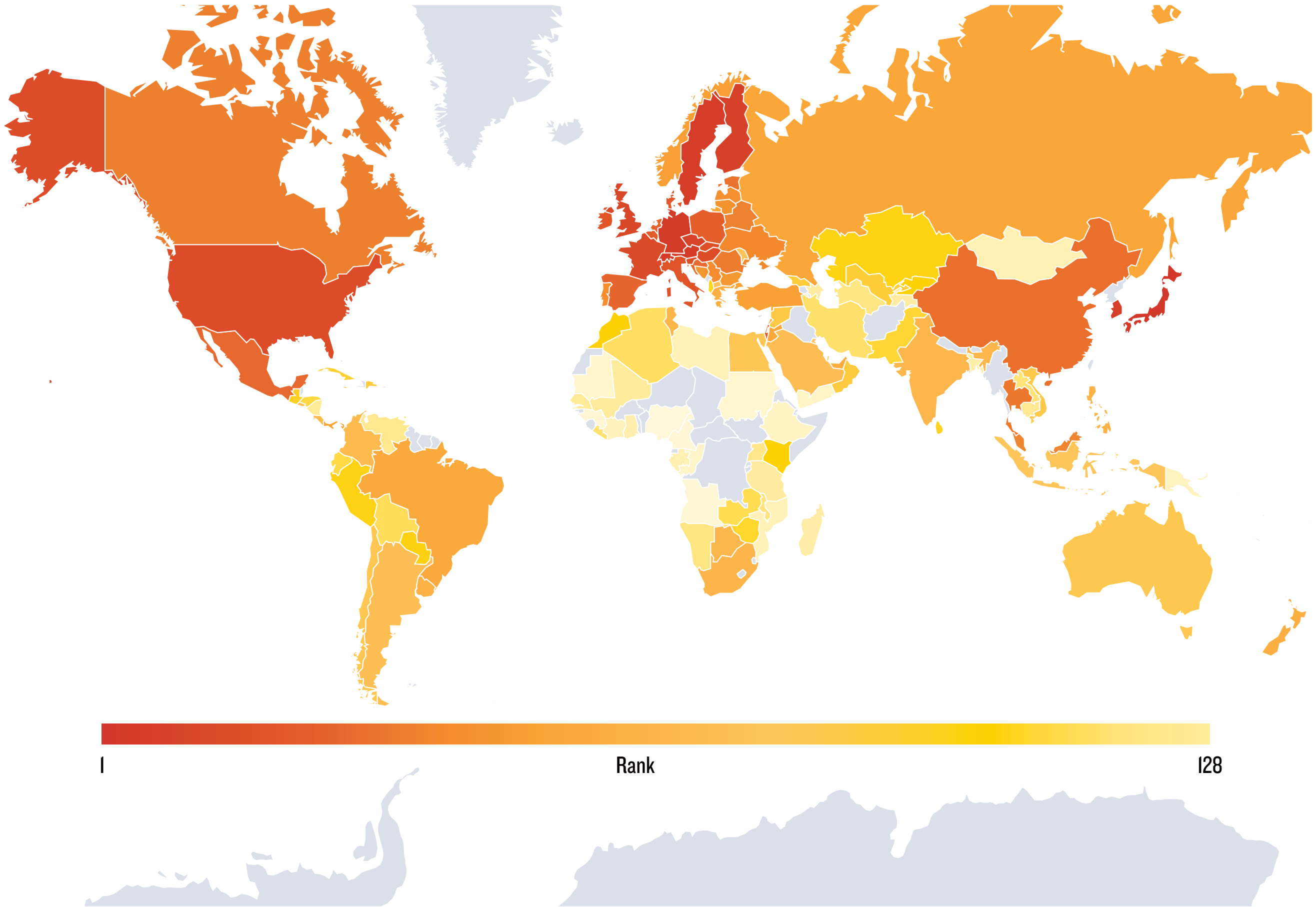
By the same token, we can improve the first impression about the complexity of a country that is given by its diversity, by also looking at the ubiquity of the products that it makes. Consider a country that chooses to concentrate in a few very complex products. It does so, not because it has few letters, but because it prefers to use them in very long words. Hence, the diversity of the country may give the wrong impression about the availability of capabilities. But if we look at the ubiquity of the products that the country makes, we would see that it specializes in low ubiquity products. We can look further into how diversified the countries that make those products are, and we will find that highly diversified countries make them. The information about how many capabilities the country has is contained not only in the number of products that it makes, but also in the ubiquity of those products and in the diversity of the other countries that make them.

Consider the case of Switzerland and Egypt. The population of Egypt is 11 times larger than that of Switzerland. At purchasing power prices their GDPs are similar since Switzerland is about 8 times richer than Egypt in per capita terms. Under the classification we use in this Atlas, they both export a similar number of different products, about 180. How can products tell us about the conspicuous differences in the level of development that exist between these two countries? Egypt exports products that are on average exported by 28 other countries (placing Egypt in the 60th percentile of countries in terms of the average ubiquity of its products), while Switzerland exports products that are exported on average by only 19 other countries, putting it in the 5th percentile. Moreover, the products that Switzerland exports are exported by highly diversified countries, while those that Egypt 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. Because of these differences, Switzerland is ranked way above Egypt in productive knowledge (Switzerland is ranked 3rd, and Egypt is ranked 67th out of 128 countries in year 2010). Ultimately, what countries make reveals what they know.

This example illustrates that we can improve the estimate of the productive knowledge of a country that we infer from its diversity by looking at the ubiquity of the products that it makes. We can refine it further by looking at the diversity of the countries that make those products and at the ubiquity of the products that those countries make. Similarly, we can improve the estimate of the productive

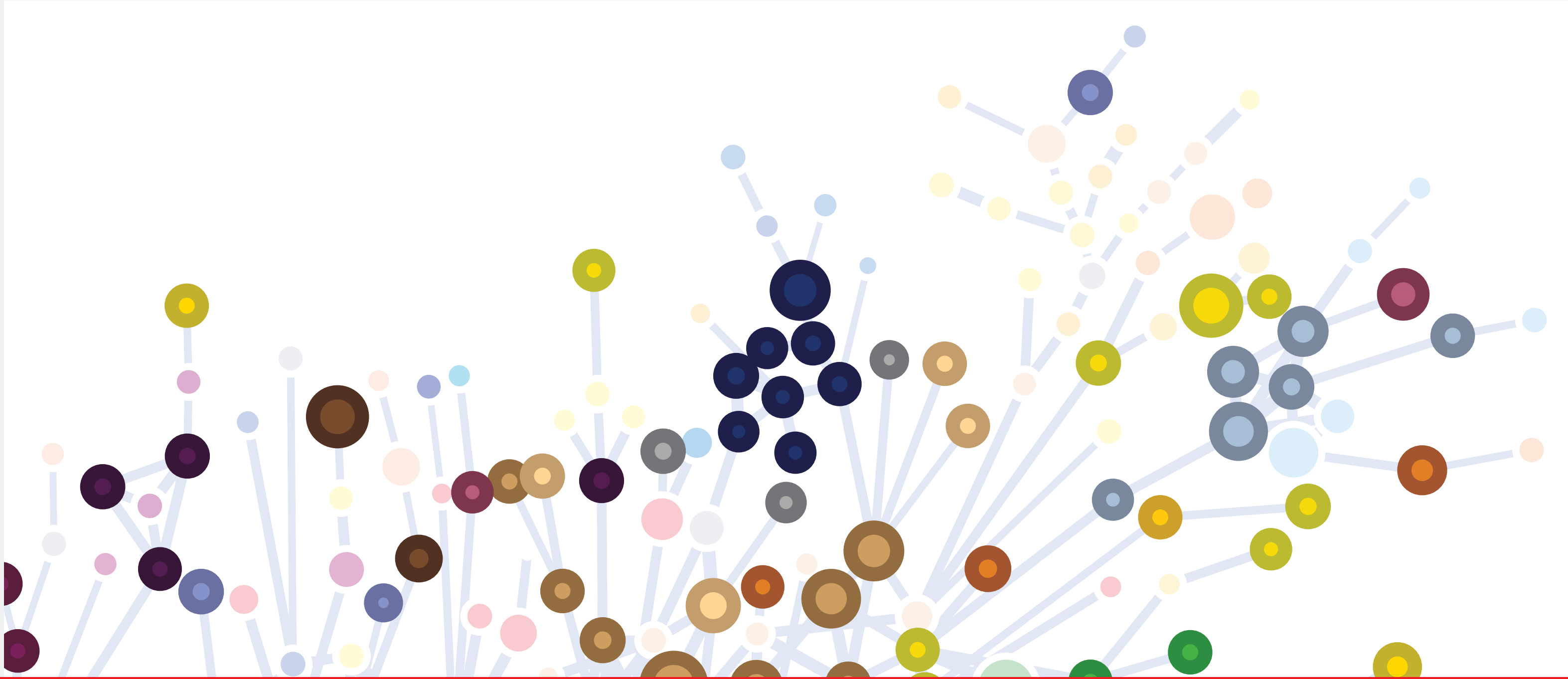
FIGURE 2.2:

Map of the World colored according to ECI Ranking.



knowledge a product requires that we infer from its ubiquity by looking at the diversity of the countries that make it, as we did with diamonds and Botswana. We can refine it further by looking at the ubiquity of the other products that diamond exporters make and at the diversity of the countries that make those other products. We can do this an infinite number of times using mathematics. This process converges after a few iterations and represents our quantitative measures of 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 in Part 2 lists countries sorted by their ECI. Figure 2.2 shows a map of the world colored according to a country's ECI ranking. Information Box 2.1 lists the most and least complex products.

This Atlas relies on international trade data. We made this choice because it is the only dataset available that has a rich and detailed cross-country information linking countries to the products that they produce using a standardized classification. This data offers great advantages, but does have limitations. First, it includes data on exports, not production. Countries may be able to make things that they do not export, although the fact that they are unable to sell those products abroad may be indicative of low productivity or quality, and hence knowledge deficiencies. Countries may also export things they do not make but only re-export. To circumvent this issue we require that countries export a *"fair share"* of the products we associate with them (see Technical Box 2.1). A second limitation is that this dataset includes only goods and not services, because the latter do not go through customs offices, which are the source of the statistical records. This is an important drawback, as services are a rising share of international trade. Unfortunately, the statistical efforts of most countries have not kept up with this reality and it is difficult to capture international flows of services in a reliable way. We explored a very coarse dataset of services and found it did not add to the precision with which we can measure economic complexity (see Technical Box 3.3). Finally, the data does not include information on non-tradable activities, such as construction, electricity distribution and restaurants. These activities are not exported because producers and consumers need to meet in the same place. They are an important part of the economic eco-system, but at present there are no global datasets that capture this information. Our current research is focused on finding implementable solutions to these limitations.



INFORMATION BOX 2.1: THE WORLD'S MOST AND LEAST COMPLEX PRODUCTS

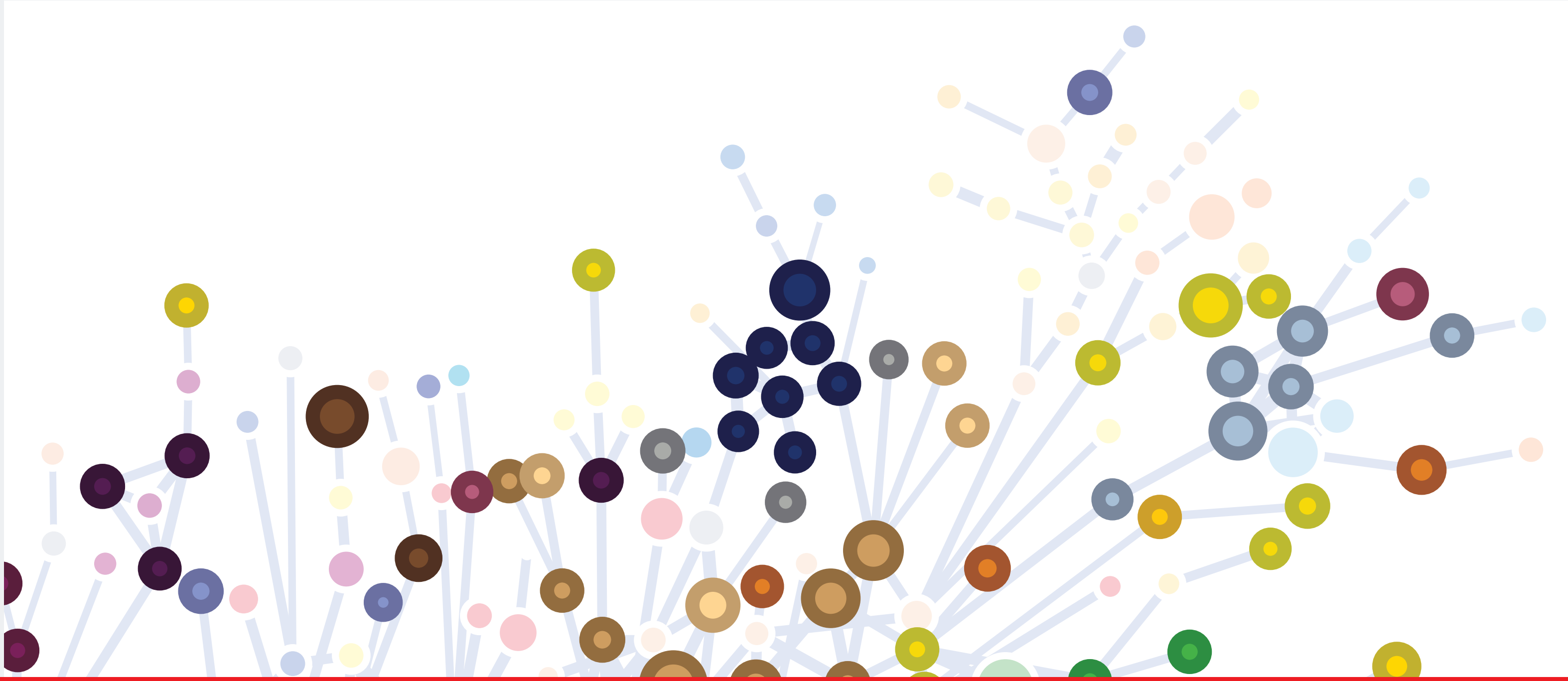
Table 2.1.1 and Table 2.1.2 show respectively the products that rank highest and lowest in the complexity scale. The difference between the world's most and less complex products is stark. The most complex products are sophisticated chemicals and machinery that tend to emerge from organizations where 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 2.1.1: TOP 5 PRODUCTS BY COMPLEXITY

Product Code (SITC4)	Product Name	Product Community	Product Complexity Index
7367	Other machine tools for working metal or metal carbide	Machinery	2.08
8744	Instrument & appliances for physical or chemical analysis	Chemicals & Health	2.02
7742	Appliances based on the use of X-rays or radiation	Chemicals & Health	1.96
8821	Chemical products and flashlight materials for use in photography	Chemicals & Health	1.91
7373	Welding, brazing, cutting, etc. machines and appliances, parts, N.E.S.	Machinery	1.86

TABLE 2.1.2: BOTTOM 5 PRODUCTS BY COMPLEXITY

Product Code (SITC4)	Product Name	Product Community	Product Complexity Index
2631	Raw cotton, excluding linters, not carded or combed	Cotton, rice, soy beans and others	-2.51
2876	Tin ores and concentrates	Mining	-2.57
2320	Natural rubber latex; natural rubber and gums	Tropical tree-crops and flowers	-2.63
2225	Sesame seeds	Cotton, rice, soy beans and others	-2.99
0721	Cocoa beans, raw, roasted	Tropical tree-crops and flowers	-3.10



TECHNICAL BOX 2.1: MEASURING ECONOMIC COMPLEXITY:

Consider M_{cp}, as a matrix in which rows represent different countries and columns represents different products. An element of the matrix is equal to 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:

$$Diversity = k_{c,0} = \sum_{p} M_{cp} \qquad (1)$$

$$Ubiquity = k_{p,0} = \sum_{c} M_{cp} \qquad (2)$$

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:

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

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

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

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

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

and rewrite this equation as:

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

where

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

We note that (7) is satisfied when $k_{c,N} = k_{c,N-2} = 1$. This corresponds to 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:

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

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

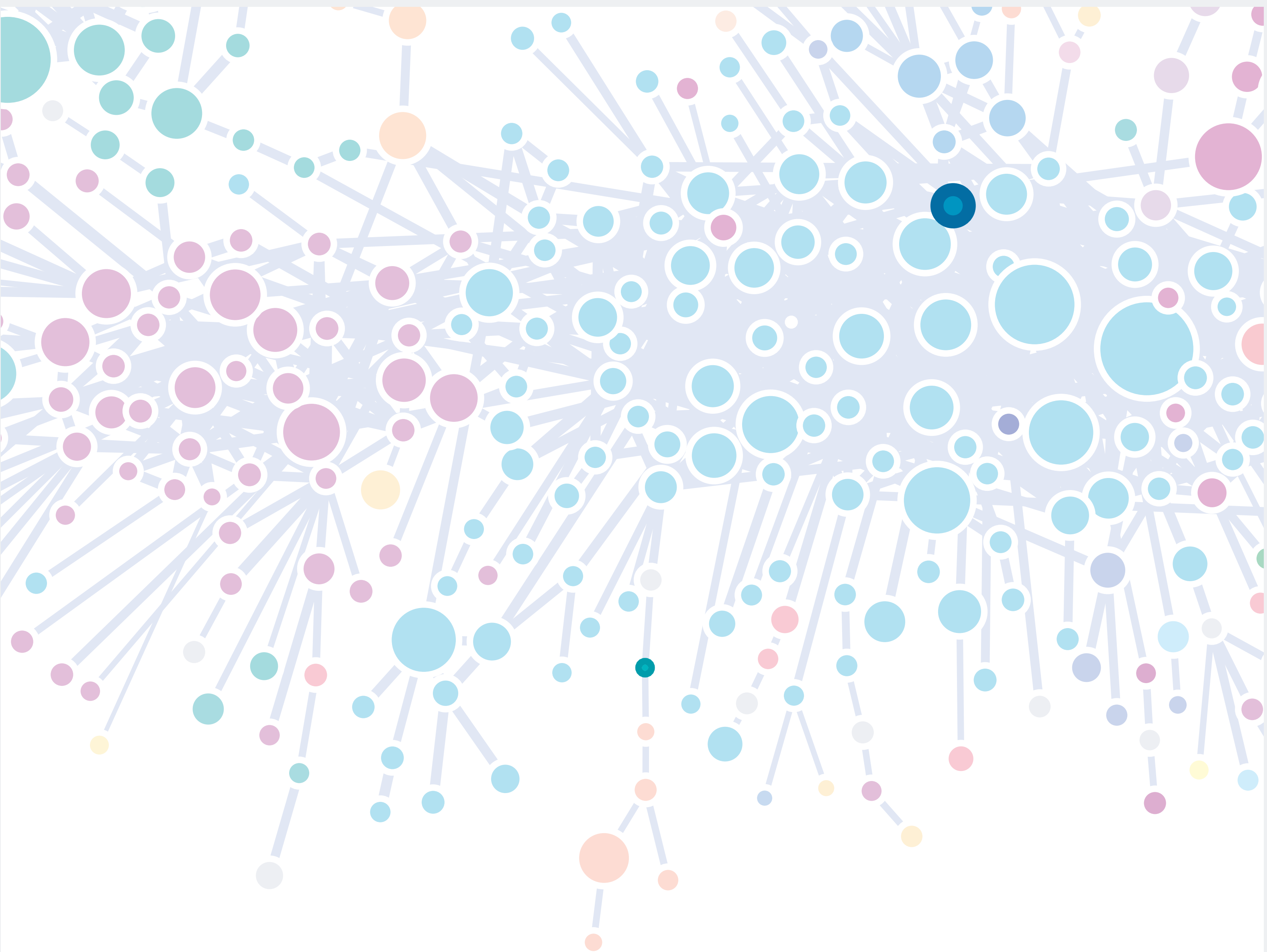
$$\vec{K} = \textit{Eigenvector of } \widetilde{M}_{cc'} \textit{ associated with second largest eigenvalue.} \qquad (10)$$

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:

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

where

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



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 the world trade in each product. 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 2010, 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 22 times Brazil's *"fair share"* of soybean exports (7.8% divided by 0.35%), so we can say that Brazil has a high revealed comparative advantage in soybeans.

Formally, if X_{cp} represents the exports of product p by country c, 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

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

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 moderate changes in export values induced by fluctuations in commodity prices by using a modified definition of RCA in which the denominator is averaged over the previous three years.