Economic Development as Global Learning

Andres Gomez-Lievano [*], **Michele Coscia** [†]

[*]Harvard University, and [†]Harvard University

Submitted to Proceedings of the National Academy of Sciences of the United States of America

The process of accumulating and coordinating increasingly complex productive capabilities, resulting in economic development, has been well documented in many empirical studies and has been the basis for the Theory of Economic Complexity. In order to quantify the size of the body of collective know-how an economy has accumulated, the index of economic complexity was developed some years ago. Some controversies surrounding this index have emerged, and some confusion has surrounded the meaning of the index. Here we provide a clarification of its meaning, why it works, and what it means for how economic development unfolds in the world.

Introduction

One of most important questions in the history of ideas is why are some societies poor and others rich. This is a question for which there is no definite answer yet. Among many of the theories and explanations, one that has recently gained attention is the Theory of Economic Complexity (TEC) (1; 2; 3; 4; 5; 6; 7). TEC emphasizes the flows of know-how as opposed to the flows of capital or labor. But more importantly, TEC emphasizes the complementarity of qualitatively different pieces of know-how, as opposed to spillovers or flows of undifferentiated/unstructured ideas (8). The main take-away from TEC is that economic development is the process of both accumulating, coordinating, and successfully deploying, qualitatively different pieces of productive know-how.

TEC has already proven to be a unifying paradigm. It explains why rich countries are diverse while poor countries are not (9; 10), which itself suggests that economic development is process of diversification, not specialization. TEC also suggests an explanation for the phenomenon of urban scaling, whereby larger cities are disproportionately more productive (11; 12), and it explains and predicts how do regions (cities, regions or countries) diversify over time (13; 14; 15; 2; 16; 17). Central to this paradigm is the question of *how much does a team of individuals know?* In other words, it pushes the research agenda towards the task of quantifying *collective know-how*, and tracking the way it grows, contracts, and is transmitted from society to society and from generation to generation (18; 19).

The notion of collective know-how has to be distinguished from the notion of human capital. Human capital, as it has been used in the literature (20; 21; 22) is an individual-specific property. In contrast, collective know-how is a group-specific property. Societies may or may not have individuals with high human capital, but still have high collective know-how. That is, a collective of individuals that know is different from a collective that knows. In an extended interpretation, collective know-how can be seen as the body of culture, as is defined in the field of Cultural Evolution (23; 24; 25; 26; 27; 28; 29; 18; 19; 30; 31). In this field, the term "cultural accumulation" is precisely the process of expanding a society's body of collective know-how. In other words, cultural accumulation is the process of collective learning.

The field of cultural evolution has shown the difficulty of formalizing the notion of collective know-how. Hence, there many ongoing efforts about measuring the collective know-how of groups of individuals (32; 33; 34; 35; 36; 37). Conceptually, it is even difficult to define what know-how is. In the literature of cultural evolution, the amount of culture (i.e., the size of a society's collective know-how) is proxied by the number of cultural elements (e.g., the number of technologies and tools in the cultural repertoir). Similarly, one good proxy for the collective know-how of a country is the number of products it exports to the world. Hence, it is reasonable to define know-how as "the size of all the different things the society knows how to do", analogously to how we assess students knowledge in standardized tests. However, knowing how to do difficult things counts more than know how to do simple things. Hence, in the reference (10) an attempt was made to derive a measure of collective know-how, and thus the Economic Complexity Index, or ECI, was born.

In this paper we introduce a new way of understanding how collective know-how expands, is maintained or lost, and we take the opportunity to clarify confusions and controversies surrounding the Economic Complexity Index (9; 10), which was expressed mathematically in its current form in the reference (38).

Mathematical Definition of the Economic Complexity Index

The calculation of the ECI starts from the matrix of countries (rows) and the products (columns) they export, $\mathbf{M}$. Let this matrix have size $C \times P$. This is a matrix that has been discretized so that $M_{c,p}$ is 1 if the product p is exported in country c, and 0 otherwise. From this matrix, one creates two stochastic matrices. First, the right-stochastic (i.e., row-stochastic or row-normalized) transition matrix of "countries to products",

$$\mathbf{R} = \mathrm{diag}\left(1/\mathbf{d}\right) \cdot \mathbf{M},$$

and second, the left-stochastic (i.e., column-stochastic or column-normalized) transition matrix of "products to countries",

$$\mathbf{L} = \mathbf{M} \cdot \mathrm{diag}\left(1/\mathbf{u}\right),$$

where $\mathbf{d} = \mathbf{M} \cdot \mathbf{1}$ is the vector that contains the number of products a country exports (i.e., its *diversity*) and where $\mathbf{u} = \mathbf{M}^T \cdot \mathbf{1}$ is the vector that contains the number of countries from which the product is exported (i.e., its *ubiquity*). We use the notation $\mathrm{diag}\left(\mathbf{x}\right)$ to mean the matrix whose diagonal is the vector $\mathbf{x}$ and the other values are zero, and $\mathbf{1}$ to denote a vector of 1's.

Four important characteristics of left-stochastic matrices are worth mentioning, as they will be useful below:[1]

Reserved for Publication Footnotes

[1] These properties also hold for right-stochastic matrices, by simply swapping the words "right" and "left".

1. A stochastic matrix for discrete markov chain can be represented as a network of nodes connected through directed edges.
2. Multiplying on the right of the matrix is the way of *propagating probabilities* through the network of connected nodes.
3. Multiplying on the left of the matrix is the way of *averaging some node-specific property* conditioned on standing on each of the nodes and only observing the nodes to which probabilities propagate to.

Let us construct the left-stochastic transition probability matrix of "countries to countries" (39),

$$\mathbf{C} = \mathbf{L} \cdot \mathbf{R}^T.$$

For mathematical convenience, we will assume that the stochastic matrix $\mathbf{C}$ is irreducible and aperiodic.[2]

Now, let $\mathbf{l_i}^T$ and $\mathbf{r_i}$ be the ith left-eigenvector and right-eigenvector, respectively, so that the eigenvalues are ordered in decreasing value, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_C$. The list of ECIs for countries is defined as the left sub-dominant eigenvector, $\mathbf{ECI}^T \equiv \mathbf{l_2}^T$:

$$\mathbf{ECI}^T \cdot \mathbf{C} = \lambda_2 \mathbf{ECI}^T. \qquad [1]$$

It is easy to prove that the vector $\mathbf{d}$ of the diversity of countries is orthogonal to the vector of ECIs, $\mathbf{ECI}$, once you realize that $\mathbf{d}$ is actually the dominant right-eigenvector (sometimes referred to as the "perron" eigenvector, or just simply, the stationary distribution of the discrete markov chain defined by $\mathbf{C}$). Thus, multiplying $\mathbf{d}$ on the right of $\mathbf{C}$, and expanding $\mathbf{C}$ into its components,

$$\begin{aligned}
\mathbf{C} \cdot \mathbf{d} &= (\mathbf{L} \cdot \mathbf{R}^T) \cdot \mathbf{d}, \\
&= (\mathbf{M} \cdot \mathrm{diag}\,(1/\mathbf{u})) \cdot (\mathrm{diag}\,(1/\mathbf{d}) \cdot \mathbf{M})^T \cdot \mathbf{d}, \\
&= (\mathbf{M} \cdot \mathrm{diag}\,(1/\mathbf{u})) \cdot (\mathbf{M}^T \mathrm{diag}\,(1/\mathbf{d})) \cdot \mathbf{d}, \\
&= (\mathbf{M} \cdot \mathrm{diag}\,(1/\mathbf{u})) \cdot \mathbf{M}^T \cdot \mathbf{1}, \\
&= (\mathbf{M} \cdot \mathrm{diag}\,(1/\mathbf{u})) \cdot \mathbf{u}, \\
&= \mathbf{M} \cdot \mathbf{1}, \\
&= \mathbf{d}.
\end{aligned} \qquad [2]$$

Thus, $\mathbf{d}$ is a right-eigenvector of $\mathbf{C}$ associated with the eigenvalue $\lambda_1 = 1$, which from the Perron-Frobenius theorem one concludes that $\mathbf{d}$ is the *dominant* right-eigenvector. This means, given classical results from discrete markov chains, that the stationary distribution of the stochastic process defined by $\mathbf{C}$ is $\boldsymbol{\pi} = \mathbf{d}/\sum_c d_c$. Therefore, since left-eigenvectors are orthogonal to right-eigenvectors, $\mathbf{l_i}^T \cdot \mathbf{r_j} = \delta_{i,j}$ (assuming eigenvectors have norm one), we conclude that

$$\mathbf{ECI}^T \cdot \mathbf{d} = 0,$$

which is a result that had been noted before already in reference (40).

All these results apply to the product space matrix as well, $\mathbf{P} = \mathbf{R}^T \cdot \mathbf{L}$. Namely, the sub-dominant left-eigenvector is the list of product complexity indices, PCIs, and the dominant right-eigenvector is proportional to the list of ubiquities.

One of the ways the economic complexity index has been defined is by postulating that products have a complexity, and that the complexity of countries is the average complexity of the products it exports. Conversely, one defines the complexity of the products as the average complexity of the countries where it is exported. It is claimed that this uniquely defines these two vectors. This is true, although any of the left-eigenvectors of the matrices $\mathbf{C}$ and $\mathbf{P}$ have this precise property. To show it, recall that $\mathbf{l_i}^T$ and $\mathbf{r_i}$ are the ith left-eigenvector and right-eigenvector of $\mathbf{C}$, respectively, and assume denote now $\mathbf{q_i}^T$ and $\mathbf{s_i}$ be the ith left-eigenvector and right-eigenvector, respectively, of $\mathbf{P}$. First, note that

$$\begin{aligned}
\mathbf{C} \cdot \mathbf{L} &= \left(\mathbf{L}\mathbf{R}^T\right) \cdot \mathbf{L}, \\
&= \mathbf{L} \cdot \left(\mathbf{R}^T \mathbf{L}\right), \\
&= \mathbf{L} \cdot \mathbf{P}.
\end{aligned} \qquad [3]$$

And second, start from the fact that $\lambda_i \mathbf{l_i}^T = \mathbf{l_i}^T \cdot \mathbf{C}$, and multiply it on the right by the matrix $\mathbf{L}$, and then use Eq. (3):

$$\begin{aligned}
\left(\lambda_i \mathbf{l_i}^T\right) \cdot \mathbf{L} &= \left(\mathbf{l_i}^T \cdot \mathbf{C}\right) \cdot \mathbf{L}, \\
\lambda_i \left(\mathbf{l_i}^T \cdot \mathbf{L}\right) &= \mathbf{l_i}^T \cdot (\mathbf{C} \cdot \mathbf{L}), \\
&= \mathbf{l_i}^T \cdot (\mathbf{L} \cdot \mathbf{P}), \\
&= \left(\mathbf{l_i}^T \cdot \mathbf{L}\right) \cdot \mathbf{P},
\end{aligned} \qquad [4]$$

which can be re-written as

$$\lambda_i \mathbf{q_i}^T = \mathbf{q_i}^T \cdot \mathbf{P}. \qquad [5]$$

It is easy to see from Equations (4) and (5) that if $\mathbf{l_i}^T$ is a left-eigenvector of $\mathbf{C}$, then the vector $\mathbf{l_i}^T \cdot \mathbf{L}$ is the ith left-eigenvector of $\mathbf{P}$, such that $\mathbf{q_i}^T = \mathbf{l_i}^T \cdot \mathbf{L}$. Since multiplying on the left of a left-stochastic matrix takes averages, we see that the left-eigenvectors of the product space are the averages of the left-eigenvectors of the country space. In particular, for the complexity indices. Furthermore, we obtain that the eigenvalues of both matrices are the same.[3] This result, obviously, applies to the sub-dominant eigenvectors of both matrices, which are the complexity indices. But the point is that if we use the definition of complexities based on the averages, one has to choose among $\min\{C, P\}$ choices.

Now, the values of ECI have been shown to be positively associated with income levels and income growth of countries (10). However, a clear and direct interpretation of the physical meaning of ECI, as the sub-dominant left-eigenvector, and its association with a measure of collective know-how, and its link to economic growth, has been lacking. The reason for this confusion is born out, first, from its flawed interpretation as a linear measure to rank countries, and second, as that interpretation that states that it is that quantity which is the average property of products.

In the next section, we clarify these issues, and we show why it is that ECI is, in fact, a good indicator of the size of the collective know-how of countries. The argument is, however, both obvious and non-trivial. Obvious precisely because of its interpretation as the sub-dominant left-eigenvector, but non-trivial, from the ultimate significance about the emphasis it gives to the underlying process of economic development as a process of collective learning.

[2] When $\mathbf{C}$ is constructed using real data, it is irreducible since all countries produce at least one product that some other country also produces, and it is aperiodic since, by construction, it has self-loops.

[3] Since $\mathbf{M}$ is typically not square, and there are more products than countries, $P > C$, this results also indicates that the matrix $\mathbf{P}$ must have some degenerate eigenvalues, which in turn explains why in the calculation of the PCIs one observes many products with identical values.

Left is for communities, right for capabilities [IS THERE A GAME OF WORDS HERE, THAT MAY HAVE BEEN USED IN THE CONTEXT OF (left, communist) AND (right, capitalist)?]

The ECI is an index that separates the set of countries in two communities, a core and a periphery. Mathematically, we show that it maximizes a community discovery task. Numerically, we create a synthetic world of several communities, and show that **ECI** is the answer to the question *which of two communities does each country belong to?* Finally, we will show that the role of the ECI becomes evident when we use the flattened export vs. product-importer matrix.

The takeaway from this section is more general, though. We show both left-eigenvectors and right-eigenvectors contain the information about the community structure of the network. However, left-eigenvectors are better about the statistical identification of communities, while the right-eigenvectors are better about measuring capabilities. These two realms are related since the extent to which a country is embedded in a community is itself a measure of the number of capabilities it has. The ultimate reason is Anna Karenina's Principle. Richly diversified countries tend to be all alike, while poorly diversified countries are poorly diversified in their own way.

Numerical demonstration Recalling that $\mathbf{C}$ is a left-stochastic matrix, there are several known results about its spectral properties and their relation to community structure (41; 42).

First, if the nodes of the network are organized in well-defined k clusters, then there are $k-1$ relatively large, non-trivial, eigenvalues, in addition to the dominant eigenvalue with value equal to 1 (43). Thus, a heuristic that can be used to infer the number of communities in a network is to count the number of eigenvalues larger than, say, 0.1, in the spectrum of the stochastic matrix. Second, the eigenvectors (both left- and right-eigenvectors) associated with those $k-1$ non-trivial eigenvalues reveal the structure of the clusters. Hence, if the clusters are well-defined, carrying out a k-means clustering on the matrix $\mathbf{E}_{C\times k-1}$ where the columns are the $k-1$ left-eigenvectors of $\mathbf{C}$ would identify the k clusters.

In our case, the left-eigenvectors of $\mathbf{C}$ can be used to discover the communities of countries, where the communities are the based on how similar are the export baskets of countries. But as we will show, both left and right-eigenvectors can carry out this function.

As a comparison, we create three $\mathbf{M}$ matrices. The first, we created by putting $M_{c,p}=1$ with a probability of 0.6 if c and p belong to the same community, and with probability 0.1 if they belong to different ones. The second way is following (44; 45), who hypothesize the interaction between two matrices, $\mathcal{C}$ and $\mathcal{P}$ determines $\mathbf{M}$. These underlying matrices are also binary matrices, which can be thought of as the matrix of countries and the capabilities they have on the one hand, and the matrix of products and the capabilities they require to be produced on the other, such that $\mathbf{M}=\mathcal{C}\odot\mathcal{P}$, where the operator $\odot$ is a production function operator, which we choose as the Leontief (i.e., a country produces a product if it has all the capabilities required to produce it). Finally, the third way to construct $\mathbf{M}$ is from real data. We choose the year 2015, 224 countries and 773 products (SITC4 codes).

Figure 1 shows the results from the matrix filled uniformly, with five communities. Figure 2 shows the results from the matrix created based on an underlying structure of capabilities, with also five communities. Figure 3 show the results from real data.

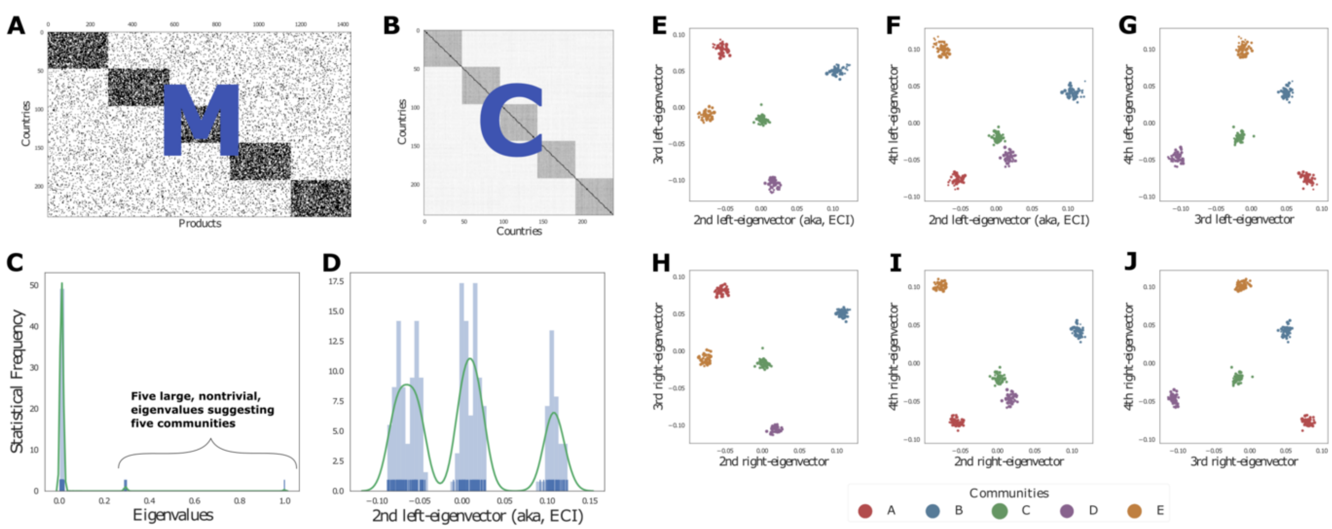

Figure 1: Example of a matrix connecting countries with products with a uniform probability. The within-community probability was set at 0.6 and the between-community probability at 0.1.

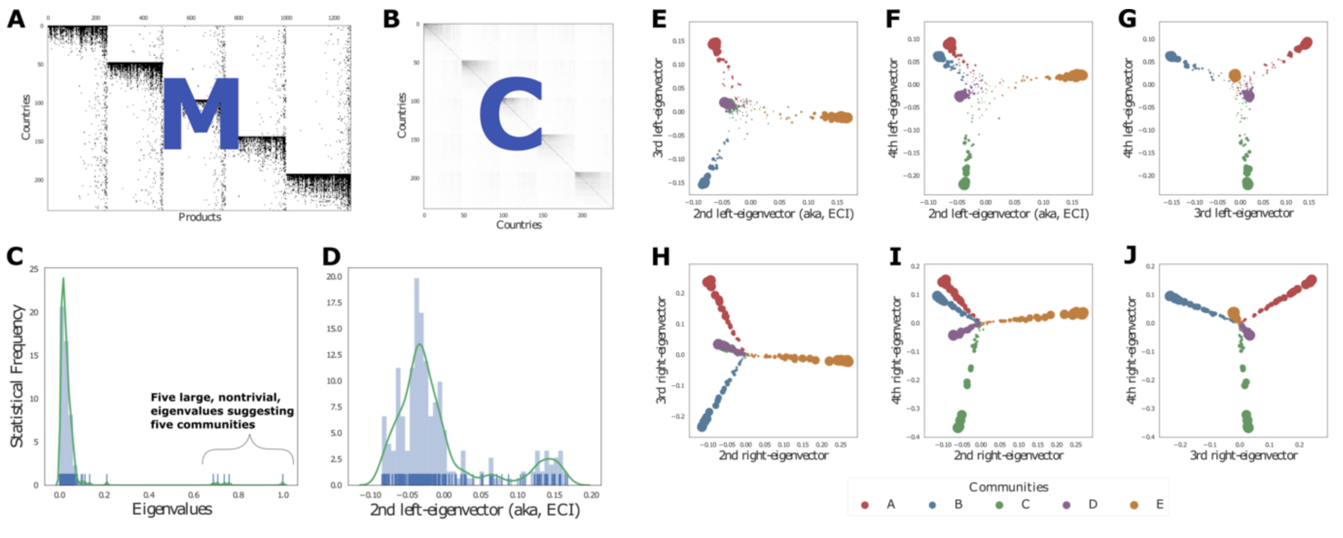

Figure 2: Example of a matrix connecting countries with products as it results from the interaction between the matrix of countries and capabilities, with the matrix of products and the capabilities required. Within each community we model a nested pattern in which some countries have many capabilities and others only a few. We also include the possibility in which some products can be produced countries regardless of the community to which they belong.

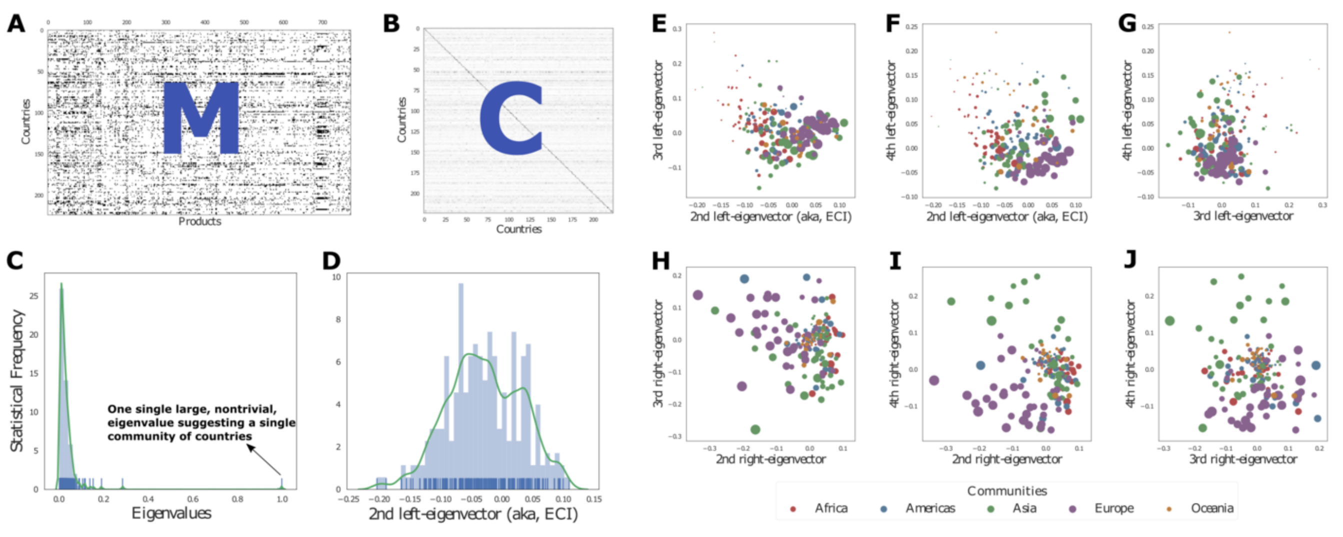

Figure 3: The same exercise as in Figures 1 and 2, but with real data from 2015.

As can be seen from Fig. 1 and Fig. 2, the left and right eigenvectors of the matrix of the country space is helpful in identifying the communities. The

Empirical demonstration (MICHELE'S SECTION) Given the literal interpretation that ECI has received as a physical measure of know-how, and its success predicting economic growth, it would seem hard to demonstrate that ECI is a clustering index. However, we can use a trick in order to disentangle the economic component of the ECI from its community-related properties. The goal is to try to use ECI to reveal *geographical communities*, which is accomplished when we add a *geographical dimension* to the matrix $\mathbf{M}$: the geographical region of destination where the exporter is exporting a product, i.e, the importer. Hence, we will now have a tensor, $M_{c,i,p}$, where c is the index of the exporter country, i the importer, and p the product traded.

Mathematical demonstration Let us begin by showing that **ECI** is a solution to the following optimization problem:

$$\max \sum_{c'} \frac{1}{d_{c'}} \sum_{c,p} \left(\frac{\not\Vdash_{(c \wedge c' \text{ export } p)}}{u_p} \right) \not\Vdash_{(c \wedge c' \text{ same community})} \quad [6]$$

The free parameters of the maximization are the assignments of countries to one of *two communities.* Let us list the properties of this quantity:

- It is higher if countries c and c' export products of low ubiquity (i.e., products that are rare, or difficult to produce).
- It is higher if country c' exports few products.
- It is higher if countries c and c' export the same products.
- It is higher if countries c and c' belong to the same community.

Since the only moving part in the quantity is to which of the communities countries belong to, the algorithm will put countries that export the same products together. Furthermore, the algorithm will weigh much more when countries export rare, and presumably difficult to make, products. Finally, each country contributes to the quantity, and since you don't want diverse countries to dominate the sum, you divide by each country's diversity. [THE ARGUMENT THAT i'D LIKE TO DEVELOP FURTHER IS THAT WHAT THE PREVIOUS PARAGRAPH IS DESCRIBING NOT ONLY THE DEFINITION OF THE ECI; iT IS THE STATE VARIABLE THAT AN ECONOMY IS TRYING TO MAXIMIZE IN THE REAL WORLD. THAT IS, COUNTRIES ARE NOT TRYING TO MAXIMIZE gdp, THEY'RE REALLY (BUT UNCONSCIOUSLY AND IMPLICITLY) TRYING TO MAXIMIZE eci. hENCE, AN ECONOMY IS TRYING TO BELONG TO A COMMUNITY. THIS SHOULD SHELD LIGHT TO DIFFERENT ASPECTS OF THE DIVERSIFICATION PROCESS, AND THUS, IT SHOULD REVEAL THE "DENSITY REGRESSIONS".]

Let us define the vector $\mathbf{s}$, such that $s_c = -1$ if the country c belongs to community 1, and $s_c = 1$ if c belongs to community 2. Recalling the definition of the elements $M_{c,p}$ of the matrix $\mathbf{M}$, the quantity to maximize in Eq. (6) can be written as

$$\begin{aligned} Q &= \sum_{c',c,p} \left(\frac{\not\Vdash_{(c \wedge c' \text{ export } p)}}{u_p} \frac{1}{d_{c'}} \right) \not\Vdash_{(c \wedge c' \text{ same community})}, \\ &= \sum_{c',c,p} \left(\frac{M_{c,p} M_{c',p}}{u_p d_{c'}} \right) \left(\frac{s_c s_{c'} + 1}{2} \right), \\ &\propto \sum_{c',c,p} s_c \left(\frac{M_{c,p} M_{c',p}}{u_p d_{c'}} \right) s_{c'}, \\ &= \mathbf{s}^T \left(\mathbf{M} \cdot \text{diag}\left(1/\mathbf{u}\right) \cdot \left(\mathbf{M}^T \text{diag}\left(1/\mathbf{d}\right)\right) \right) \mathbf{s}, \\ &= \mathbf{s}^T \mathbf{C} \mathbf{s}. \end{aligned} \quad [7]$$

Therefore, the problem is how to choose $\mathbf{s}$ so as to maximize the expression in Equation (7). The strategy for solving this problem is the so-called "spectral approach" (42; 46), expressing the vector $\mathbf{s}^T$ as a linear combination of the eigenvectors of $\mathbf{C}$. Directly replacing this into Eq. (7) will not be useful because the eigenvectors of a non-symmetric matrix are not orthogonal among themselves (only to the other side eigenvectors counterparts). So let us express $\mathbf{s}$ using both the left and right-eigenvectors,

$$\begin{aligned} \mathbf{s}^T &= \sum_i a_i \mathbf{l_i}^T, \\ \mathbf{s} &= \sum_i b_i \mathbf{r_i}. \end{aligned} \quad [8]$$

Replacing these expressions in Eq. (7), we get

$$Q = \sum_i \lambda_i a_i b_i. \quad [9]$$

As has been shown in a number of different instances (47; 48; 46), this quantity can be maximized by choosing the $\mathbf{s}$ to be parallel to eigenvectors with the largest eigenvalues, subject to the constraint that $\mathbf{s}^T \cdot \mathbf{s} = C$. In other words, since the maximization problem was originally set as to find the two main communities, one must choose the subdominant eigenvalue, and therefore the corresponding eigenvectors.

Improvements on the Theory of Economic Complexity

The Theory of Economic Complexity introduced two main ideas to understand economic development. On the one hand, it proposed that economic development is growth at the extensive margin, or, in less technical terms, economic development is a process of diversification. On the other hand, it introduced the notion that economic development is a historical process, in the sense that the extensive margin is constrained by path dependence and therefore its growth depends on previous contingent events.

These conceptual innovations were accompanied by two analytical tools: the technological space that gives structure to the historical process of development, and the quantification of how much a society knows as a collective through an index of economic complexity. Here we show how our understanding of the physical and mathematical meaning of both innovations allow us to improve both notions. We demonstrate this by showing a superior capacity to predict growth.

Hypothesis 1: A better algorithm to calculate the embeddedness of a country in a community may predict the appearance of specific products better than before.

Hypothesis 2: A better measure (remember to make a difference between an index, and a measurement/inference) of economic complexity may predict GDP growth better than before.

The significance of belonging to a community

The discovery presented in Section that ECI is a community discovery index is very important piece to understand the process of economic development. In concise terms, it suggests that countries accumulate capabilities by "looking" like one another. They accomplish this by processes whereby local firms imitate foreign firms, by people migrating in and out of countries. This process drives the fact that there are no separate communities of countries, but only one. But what explains the fact that is nested?

Nestedness comes from the fact that the underlying matrix of capabilities that countries have, and the matrix of capabilities required by products, are *also* nested.

Discussion

The results of this paper can be concisely summarized as follows: economic development is first and foremost a process of global imitation with local exploration. Imitation gives rise to a single connected network of countries, and both imitation and local exploration gives rise to the nested structure in which richer countries are more diverse and build on what less diverse countries know-how to do. The main significance of this is that humanity as a whole is exploring a single branch of the technological tree.

Methodologically, our results imply that identifying the community in the network to which an economy belongs provides a wealth of information about the current and future productive capabilities of the economy.

Conclusion

The field of economics has improved of our understanding of markets. In markets, the two organizing concepts are price and quantity, and the main organizing mechanism is the equalization of supply and demand. The Theory of Economic Complexity complements our understanding of economic systems by emphasizing the role of collective know-how. The adaptability, welfare, and robustness of an economy all are determined by the size of the collective know-how of a society. While the traditional market-centered approach to economics has developed sophisticated tools to track the flows of value and money, TEC is in the process of developing better tools and formalisms to track the expansion, maintenance, and transference, of collective know-how. In this paper we have build on previous ideas to reinforce the claim that networks are behind the processes of economic development: networks highlight collective-level properties like communities. And the main questions are thus: how do economic communities are formed? Why do nodes in the same community grow together? Why do we see convergence in income when nodes belong to the same community (Coscia and Hausmann!)(49; 50)?

Hello

this is cursive and ghere is a citation (51)

References

1. Hausmann R, Klinger B (2006) Structural Transformation and Patterns of Comparative Advantage in the Product Space. *SSRN Electronic Journal.* doi:10.2139/ssrn.939646.

2. Hidalgo CA, Klinger B, Barabasi A-L, Hausmann R (2007) The Product Space Conditions the Development of Nations. *Science* 317(5837):482–487.

3. Arthur WB Complexity and the Economy. doi:10.4337/9781781952665.00007.

4. Arthur WB, Polak W (2006) The evolution of technology within a simple computer model. *Complexity* 11(5):23–31.

5. Hausmann R, Hidalgo C (2010) Country Diversification Product Ubiquity, and Economic Divergence. *SSRN Electronic Journal.* doi:10.2139/ssrn.1724722.

6. Bahar D, Hausmann R, Hidalgo C (2012) International Knowledge Diffusion and the Comparative Advantage of Nations. *SSRN Electronic Journal.* doi:10.2139/ssrn.2087607.

7. Frenken K, Boschma R Notes on a Complexity Theory of Economic Development. doi:10.4337/9780857930378.00022.

8. Lucas R (2008) *Ideas and Growth* (National Bureau of Economic Research) doi:10.3386/w14133.

9. Hausmann R, Hidalgo CA (2011) The network structure of economic output. *Journal of Economic Growth* 16(4):309–342.

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

11. Gomez-Lievano A, Patterson-Lomba O, Hausmann R (2016) Explaining the prevalence scaling and variance of urban phenomena. *Nature Human Behaviour* 1(1):0012.

12. Neffke F (2017) Coworker Complementarity. *SSRN Electronic Journal.* doi:10.2139/ssrn.2929339.

13. Neffke F, Henning M (2012) Skill relatedness and firm diversification. *Strategic Management Journal* 34(3):297–316.

14. Neffke FMH, Henning M, Boschma R (2011) The impact of aging and technological relatedness on agglomeration externalities: a survival analysis. *Journal of Economic Geography* 12(2):485–517.

15. Neffke F, Henning M, Boschma R (2011) How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Economic Geography* 87(3):237–265.

16. Cristelli M, Tacchella A, Pietronero L (2015) The Heterogeneous Dynamics of Economic Complexity. *PLOS ONE* 10(2):e0117174.

17. Cristelli M, Tacchella A, Pietronero L (2014) An Overview of the New Frontiers of Economic Complexity. *Econophysics of Agent-Based Models* (Springer International Publishing), pp 147–159.

18. Richerson PJ, Boyd R (2004) *Not By Genes Alone* (University of Chicago Press) doi:10.7208/chicago/9780226712130.001.0001.

19. Henrich J (2016) *The Secret of Our Success* (Princeton University Press) doi:10.1515/9781400873296.

20. Romer P (1989) *Endogenous Technological Change* (National Bureau of Economic Research) doi:10.3386/w3210.

21. Romer P (1989) *Human Capital And Growth: Theory and Evidence* (National Bureau of Economic Research) doi:10.3386/w3173.

22. Jones C, Romer P (2009) *The New Kaldor Facts: Ideas Institutions, Population, and Human Capital* (National Bureau of Economic Research) doi:10.3386/w15094.

23. Feldman MW, Cavalli-Sforza LL (1986) TOWARDS A THEORY FOR THE EVOLUTION OF LEARNING. *Evolutionary Processes and Theory* (Elsevier), pp 725–741.

24. Cavalli-Sforza L, Feldman M, Chen K, Dornbusch S (1982) Theory and observation in cultural transmission. *Science* 218(4567):19–27.

25. Feldman MW, Cavalli-Sforza LL (1984) Cultural and biological evolutionary processes: gene-culture disequilibrium.. *Proceedings of the National Academy of Sciences* 81(5):1604–1607.

26. Richerson PJ, Boyd R (2008) Being human: Migration: An engine for social change. *Nature* 456(7224):877–877.

27. Badcock C, Boyd R, Richerson PJ (1988) Culture and the Evolutionary Process.. *Man* 23(1):204.

28. Boyd R, Richerson PJ, Henrich J (2013) The Cultural Evolution of Technology. *Cultural Evolution* (The MIT Press), pp 119–142.

29. Griesemer JR, Boyd R, Richerson PJ (1986) Culture and the Evolutionary Process. *The Condor* 88(1):123.

30. Muthukrishna M, Henrich J (2016) Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371(1690):20150192.

31. Muthukrishna M, Doebeli M, Chudek M, Henrich J (2017) The Cultural Brain Hypothesis: How culture drives brain expansion underlies sociality, and alters life history. doi:10.1101/209007.

32. Mesoudi A (2017) Pursuing Darwin's curious parallel: Prospects for a science of cultural evolution.. *Proc Natl Acad Sci U S A.*

33. Kempe M, Lycett SJ, Mesoudi A (2014) From cultural traditions to cumulative culture: parameterizing the differences between human and nonhuman culture.. *J Theor Biol* 359:29–36.

34. Legare CH (2017) Cumulative cultural learning: Development and diversity.. *Proc Natl Acad Sci U S A.*

35. Kline MA, Boyd R (2010) Population size predicts technological complexity in Oceania. *Proceedings of the Royal Society B: Biological Sciences* 277(1693):2559–2564.

36. Vaesen K (2012) Cumulative Cultural Evolution and Demography. *PLoS ONE* 7(7):e40989.

37. Henrich J (2004) Demography and Cultural Evolution: How Adaptive Cultural Processes Can Produce Maladaptive Losses—The Tasmanian Case. *American Antiquity* 69(02):197–214.

38. Hausmann R, et al. (2011) *The atlas of economic complexity: Mapping paths to prosperity* (Puritan Press).

39. Yildirim MA, Coscia M (2014) Using Random Walks to Generate Associations between Objects. *PLoS ONE* 9(8):e104813.

40. Kemp-Benedict E (2014) An interpretation and critique of the Method of Reflections. *MPRA Paper* 60705.

41. Capocci A, Servedio VDP, Caldarelli G, Colaiori F (2005) Detecting communities in large networks. *Physica A: Statistical Mechanics and its Applications* 352(2-4):669–676.

42. Malliaros FD, Vazirgiannis M (2013) Clustering and community detection in directed networks: A survey. *Physics Reports* 533(4):95–142.

43. Chauhan S, Girvan M, Ott E (2009) Spectral properties of networks with community structure. *Physical Review E* 80(5). doi:10.1103/physreve.80.056114.

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

45. Hausmann R, Hidalgo CA (2011) The network structure of economic output. *Journal of Economic Growth* 16(4):309–342.

46. Newman MEJ (2006) Finding community structure in networks using the eigenvectors of matrices. *Physical Review E* 74(3). doi:10.1103/physreve.74.036104.

47. Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. *Physical Review E* 69(2). doi:10.1103/physreve.69.026113.

48. Leicht EA, Newman MEJ (2008) Community Structure in Directed Networks. *Physical Review Letters* 100(11). doi:10.1103/physrevlett.100.118703.

49. Coscia M, Cheston T (2017) Institutions vs. Social Interactions in Driving Economic Convergence: Evidence from Colombia. *SSRN Electronic Journal*. doi:10.2139/ssrn.2939678.

50. Coscia M, Hausmann R (2015) Evidence That Calls-Based and Mobility Networks Are Isomorphic. *PLOS ONE* 10(12):e0145091.

51. Hausmann R, Pritchett L, Rodrik D (2004) *Growth Accelerations* (National Bureau of Economic Research) doi:10.3386/w10566.