Supplementary information for

Amalgamation of Export with Import Information: The Economic Complexity Index as a Coherent Driver of Sustainability

Benjamin Cakir* Isabelle Schluep † Philipp Aerni ‡

Isa Cakir §

February 10, 2021

Acknowledgment

The authors would like to thank Dr. Eugenio Paris for his support in programming, numerical analysis and his availability as a sparring partner.

*Mathematics Student at the University of Zürich

†Center of Corporate Responsibility and Sustainability (CCRS) at the University of Zürich

‡Center of Corporate Responsibility and Sustainability (CCRS) at the University of Zürich

§Center of Corporate Responsibility and Sustainability (CCRS) at the University of Zürich

Contents

1 Introduction **3**

2 How this manuscript is organized **4**

3 Modular Scheme of the Method of Reflection **6**
3.1 Information Matrix (or Fundamental Matrix) 6
3.2 Recalling the Method of Reflection 6
3.3 Modular Scheme of Modeling Process 8
3.4 Information Matrices and Engineering of S 10
3.4.1 Original Information Matrix 10
3.4.2 An Alternative Information Matrix and Engineering of S 11
3.4.3 Economic Performance Equivalent (EPE) (or Compatible) Import Information Matrix 17

4 Stochastic Counterpart of S and Interpretation of the ECI **20**
4.1 Settings 20
4.2 Role of e_2 for Large t 21

5 Coherence of Sustainability Index and ECI **33**
5.1 Motivation; Expert in Box 34
5.2 Some Coherency Conditions and ECI 35
5.3 Alternative Ranking Methods 43
5.4 Coherency of Ranking Methods After Amalgamation 43
5.5 Summary of Coherency Results 48

6 Integrating Import **50**
6.1 Amalgamation by Lie-Trotter Product Formula 51
6.2 Amalgamation by Lie-Trotter Formula on pre S-level 52

7 Appendix **53**
7.1 Random Walk on a Graph 53
7.2 Positivity 60
7.3 Amalgamation of Random Walks and Lie Trotter Approach . . 64

1 Introduction

The main objective of this manuscript is to establish self-contained technical manual on how to merge supplementary value chain information in a network in order to establish a holistic ranking standard in a coherent and organic manner. We propose a process of *amalgamation* of information matrices that is based on the *Lie-Trotter* product formula (see e.g. the original paper of Trotter [1] or [2] for more general cases) and analyze how this method adds value to the quality existing ranking framework. In order to focus on the ideas, this manuscript is mainly based on the original Economic Complexity Index (ECI) which was developed by C. Hidalgo and R. Hausmann in 2009 (see [3] and [4]) to measure the accumulated productive know-how of a country and to deduce from it its future growth prospects. It is based on the assumption that the availability of capabilities in a country can be inferred from export data taking into account the country's diversification and ubiquity of its exported products in a global network. Compared to models in neoclassical economics (see e.g. [5] for Heckscher–Ohlin model) the ECI represents a step forward adding additional insight about the sustainability of countries. Since then there have been multiple variations of methodological suggestions with similar predictive power (see e.g. [6] and [7]).

We assume that first step for improvement of ECI is to understand mechanic of the ECI from different perspectives and embed the ECI framework in its stochastic counterpart which allows interpretations in statistical, probabilistic as well as in information-theoretical terms. As one of the results, we present a relationship between *Shannon Entropy* and the ECI. The second step is to investigate to what extent the ECI can be viewed as a coherent sustainability driver and what the alternatives are. These preparations will enable us to detect the relevancy of the ECI as a coherent driver of economic sustainability and how we can integrate additional value chain information in a coherent and tractable manner.

It is also worth mentioning that the suggested amalgamation process of global import data with export data allows the incorporation of internal economic complexity (see [8], [9], [10] and [11] internal complexity) into the ECI framework in a comparable way.

Althoug we used the original ECI as guidance, we expect that our main objective can be explored further and applied to other ranking methods for similar purposes (compare subsection 3.4). However, rigorous treatment of possible generalizations is out of scope of the present manuscript.

2 How this manuscript is organized

In order to ensure the tractability, in section 3, we first recover the method of reflection introduced in [3] and thereafter embed it in a modular scheme. We translate the bipartite graph representing the exported products of countries into a weighted graph inducing a Markov Chain on a Network of countries and products respectively which are special Random Walks on a Graph. In this section, we demonstrate how to switch from one playground to the other in a coherent general way. The subsection 3.4 is devoted to prototypes of weight matrices and their associated Random Walks on a Graph driven by the prototype behaviors of market players in an organically evolving market (see e.g. [12], [13] and [14] for further readings). But, in this subsection we also discuss how the information matrices should be prepared, what the assumptions are behind them, and *performance equivalent* versions of import information as supplementary value chain information.

In section 4, the stochastic counterpart of the method of reflection is introduced, where we treat the stochastic interpretation of the ECI and PCI, and give an elementary proof of the lemma allowing statistical interpretation under mild assumptions for a general random variable X (which we call pay-off function) on the state space of the Random Walk on a Graph (see lemma 4.1). The idea of the lemma 4.1 was already discussed for a certain case, with a rough formulation of the conditions (e.g. see equation 11 and the related discussions in [15]). The lemma 4.1 as well as a lemma in probabilistic terms (see lemma 4.2) deliver already appealing interpretations of e_2 and hence ECI under mild assumptions. In our proofs, we used "*ergodicity*", and "*positivity*" of eigenvalues of a Random Walk on a Graph generated by intersection matrices [see e.g. [16]], which is equivalent to the Method of Reflection for even time indices (compare the section 3.3). These crucial properties are treated rigorously in the Appendix. Especially, the positivity of the second eigenvalue plays an important role, which is according to our literature research in this context not mentioned explicitly. Finally, we demonstrate the sharp dominance of the *Shannon Entropy* of transition probabilities by e_2 (given their initial know-how states, see lemma 4.3), where the lemma 4.1 plays an important role.

Sustainability in a is sense a "more difficult risk management". In risk management, it took decades to pay attention to *coherency* of a risk measure which is now a well understood and crucial approach. In [17] the authors have formulated the so-called "axioms of coherence for acceptance sets or for risk measures" which are natural conditions from an economic point of view. The regulatory capital required for insurance companies is now based on the coherent risk measure within the Swiss Solvency Test (SST). Inspired by this discipline, we also drew our attention to the coherency issue from a sustainability point of view, which is an important eligibility criterion, closely related to the predictive power of a ranking instrument. In section 5, we formulate some coherency conditions of complexity as a systematic driver of sustainability rating. In subsection 5.2

we discuss the coherence of e_2 as well as an alternative (diagonal of weighting matrix S) by looking at prototype examples in standalone cases. Finally, we also show that under monotony condition e_2 behaves well (see lemma 5.1). The coherency of e_2 and diagonal of weighting matrix S with respect to the *amalgamated* version is discussed in the last subsection of section 5.

In section 6, we discuss some obvious approaches for merging of supplementary information and present our main result, which we call *amalgamation on a pre-S-level* or shortly *amalgamation*. We believe that if an organic evolution of an issue is in place, the amalgamation of supplementary information is an alternative that can improve the quality of a ranking system in general. However, in the present paper, our special focus will be on the amalgamation of import information with export information.

3 Modular Scheme of the Method of Reflection

Our main goal goal in this section is to introduce an engineering mindset based on an alternative information matrix, that can be used for generating alternative weight matrices allowing evolutionary interpretation.

The purpose of the next two subsection are to recover the original idea in the matrix language and explore the different possibilities of information matrices.

3.1 Information Matrix (or Fundamental Matrix)

The Method of Reflection in the original ECI is based on the information matrix covering all the information of a bipartite graph.

Given a bipartite graph (A, B, E) consisting of disjoint sets of numerated vertexes $A = \{1, 2, ...m\}$, $B = \{1, 2, ...n\}$ and edges E connecting the elements of A with the elements of B, we can define a matrix $(M(i, j))_{i\in A, j\in B}$ which captures all the information of the graph (A, B, E), by putting:

$$M(i,j) = \begin{cases} 1 & \text{if } i \in A \text{ and } \ j \in B \text{ is connected by an edge in } E \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

In this manuscript, the matrix M will be called zero one information matrix (or fundamental matrix.)

The more general information matrices with positive real numbers will be denoted later by Y.

3.2 Recalling the Method of Reflection

Hidalgo et al introduced the Method of Reflection from which a transition probability matrix, governing the dynamic of network is constructed (see e.g. [2] and [3]).

Given the countries C and the exported products of the countries P we can directly write the information matrix as follows:

$$M(i,j) = \begin{cases} 1 & \text{if } i \in C \text{ exports the product } j \in P \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

By using the the above information matrix M we define a recursive process starting from $t = 0$

$$k_{i,0} = \text{number of products exported by the country } i, \tag{3}$$

$$K_{j,0} = \text{number of countries exporting the product } j \tag{4}$$

$k_{i,0}$ and $K_{i,0}$ are called diversity of the country c and ubiquity of product j respectively.

For $t > 0$ we put:

$$k_{i,t} = \frac{1}{k_{i,0}} \sum_j M(i,j) K_{j,t-1} \tag{5}$$

$$K_{j,t} = \frac{1}{K_{i,0}} \sum_i M(i,j) k_{i,t-1} \tag{6}$$

Each even time index $2t$ represents a "complete" step in the above recursion process mixing the diversity and ubiquity. If we use recursive step 6 for $K_{j,t-1}$ and plug into 5 we obtain:

$$k_{\cdot,2t} = D_C * M * D_P * M^{tr} * k_{\cdot,2(t-1)}. \tag{7}$$

Similarly we obtain

$$K_{\cdot,2t} = D_P * M^{tr} * D_C * M * K_{\cdot,2(t-1)}. \tag{8}$$

Here D_C and D_P are diagonal matrices with:

$$D_C(i,i) = \frac{1}{k_{i,0}} \tag{9}$$

and

$$D_P(j,j) = \frac{1}{K_{j,0}}. \tag{10}$$

We define transition probability matrices of countries rsp. products as follows:

$$P_C = D_C * M * D_P * M^{tr} \tag{11}$$

$$P_P = D_P * M^{tr} * D_C * M \tag{12}$$

With a simple time transformation $\hat{k}_{\cdot,t} = k_{\cdot,2t}$ and $\hat{K}_{\cdot,t} = K_{\cdot,2t}$ the above recursions 7 and 8, can be written in equivalent form as follows:

$$\hat{k}_{\cdot,t} = P_C * \hat{k}_{\cdot,t-1} \tag{13}$$

$$\hat{K}_{\cdot,t} = P_P * \hat{K}_{\cdot,t-1}. \tag{14}$$

In what follows we will omit the hats on $\hat{k}_{\cdot,t}$ and $\hat{K}_{\cdot,t}$ in on order to avoid notational overflow and use $k_{\cdot,t}$ and $K_{\cdot,t}$ keeping the time transformation in mind.

We would like to interpret $P_C(i,\tilde{i})$.

Let μ_i be the uniform probability measure on set of products exported by i so that $\mu_i(j) = \frac{1}{k_{i,0}}$ for all products j exported by i and $1_i(j) = 1$ if product j is exported by i and zero otherwise.

We can write

$$P_C(i,\tilde{i}) = E_{\mu_i}(1_i 1_{\tilde{i}} \frac{1}{K}) \tag{15}$$

Where K is a function of products and is defined by $K(j) = K_{j,0}$ which is the number of countries exporting the product j. $\frac{1}{K(j)}$ indicates the concentration of the product j with respect to countries takes values between zero and one. Since high value of $\frac{1}{K(j)}$ indicates low number of competitors, we can expect that the higher the value of $\frac{1}{K(j)}$ the higher is its "length" of its value chain (or better; aggregated steps from origin to the end). Hence $P_C(i,\tilde{i})$ can be interpreted as: "expected length" of a value chain of common products exported by i and $\tilde{i}$, given the products of i. In some sense $P_C(i,\tilde{i})$ can also be interpreted as "common knowledge" shared by $\tilde{i}$ and i, from the point of view of i.

Finally we would like note that if we work with zero one information matrix as described above, each country i, $P_C(i,i)$ dominates the other elements in the same row:

$$P_C(i,i) \geq P_C(i,\tilde{i}), \quad \forall \tilde{i}, \tag{16}$$

This inequality already indicates that we are dealing with special transition probability matrices.

In the next subsection, we would like to introduce a flexible and helpful scheme which allows us to model the complexity in more general and modular way.

3.3 Modular Scheme of Modeling Process

In this subsection, we would like introduce a scheme which allows to obtain other possibilities of weighting a graph.

In terms of Graph Theory we can switch from our bipartite weighted graphs $G_C = (C, E_C, S_C)$ resp. $G_P = (P, E_P, S_P)$ consisting of vertexes, edges connecting the vertexes (which can be identified as subset of the Cartesian product of vertexes) and weights of edges.

We put

$$E_C = C\text{x}C \tag{17}$$

resp.

$$E_P = P\text{x}P \tag{18}$$

and

$$S_C = M * D_1 * M^{tr} \tag{19}$$

resp.

$$S_P = M^{tr} * D_2 * M. \tag{20}$$

In the above scheme, the information matrix M does not have to be a zero-one matrix. We only require that it has positive elements. Also D_1 and D_2 don't have to be generated from information matrix internally. We only require that it is a diagonal matrix with strictly (in order to avoid degenerated cases) positive elements at the diagonal.

Now we will define P_C resp. P_P left diagonals operating on S_C resp. S_P.

$$D_C(i,i) = \frac{1}{\sum_{\tilde{i}} S_C(i,\tilde{i})} \tag{21}$$

and

$$D_P(j,j) = \frac{1}{\sum_{\tilde{j}} S_P(j,\tilde{j})}. \tag{22}$$

Now we can put:

$$P_C = D_C * S_C \tag{23}$$

and

$$P_P = D_P * S_P, \tag{24}$$

As a conclusion, the above process suggests that once the general information matrix (or fundamental matrix) Y and a meaningful diagonal Matrix D_C resp D_P are given, we can switch to weighted graphs $G_C = (C, E_C, S_C)$ resp. $G_P = (P, E_P, S_P)$ by putting

$$S_C = Y * U * Y^{tr} = Y * U^{\frac{1}{2}} * U^{\frac{1}{2}} * Y^{tr}. \tag{25}$$

Followed by normalizing steps 21 and 22, respectively, in order to get the Markov Chains 23 and 24.

However, the usual matrix operation with a general information matrix as above doesn't necessarily make sense. As mentioned before, our goal in this section is to introduce an engineering approach that allows different weight matrices generating different transition probability matrices accordingly which can be interpreted.

For this purpose, we will put:

$$S_C = A \circ^f A^{tr}. \tag{26}$$

Our choice of notation "$\circ^f$" instead of matrix multiplication "$*$" at the right hand side of the equation 26 is deliberate. Roughly speaking, it allows a symmetric binary function $f(a,b)$ between matrix elements preserving the desired properties such as positivity and which can be interpreted in probabilistic terms. The operation $\circ^f$ imitates the matrix multiplication operation with general symmetric binary function generalizing the multiplication of real numbers:

$$A \circ^f A^{tr} = \sum_{k \leq n} f(A(i,k), A^{tr}(k,\hat{\imath})) \tag{27}$$

Here

$$A = Y_N * U^{\frac{1}{2}} \tag{28}$$

and Y_N in 28 is generated from Y by an additional moralizing step ensuring theoretical consistency so that it makes sense in probabilistic terms complies with the binary function f.

The precise formulation for different prototype probabilistic assumptions will be treated in the next subsection.

Note that in the original approach, for the construction of Markov Chains on state spaces C resp. P, the building blocks such as ubiquities of the products or diversities of the countries are generated by a zero-one matrix M and therefore they are purely internal. However, depending on the application, it might also be more meaningful to use other diagonal matrices exogenous nature for weighting purposes in the above definition of S_C resp. S_P.

Since our conclusions are analog for ECI and PCI, for the rest of this manuscript, we will omit the subscripts C and P in S and P and simply write S and P instead of S_C and P_C or S_P and P_P unless there is a necessity within a context.

3.4 Information Matrices and Engineering of S

Our plan is to introduce an evolutionary mindset. In order to motivate our approach, we will first recall the original information matrix of ECI followed by an alternative with its associated diagonal matrix in the same spirit. The goal of this second approach is to avoid losing information by possible intermediate steps. Thereafter, we will introduce different weight matrices under different behavior assumptions of the market players. Finally, we discuss how to generate the *economic performance equivalent* (EPE) information matrix of import trade volume of the countries.

3.4.1 Original Information Matrix

In page 6 of [6] $M(i,j)$ was generated as follows (only with a slightly different notation):

Let $X_{i,j}$ be the random variable representing the exported trade volume of country i in product j and $x_{i,j}$ its observed realization. Then the expected export of product j by a country i can be estimated by:

$$\widehat{E(X_{i,j})} = \frac{\sum_{\hat{i} \in C} x_{\hat{i},j} \sum_{\hat{j} \in P} x_{i,\hat{j}}}{\sum_{\hat{i},\hat{j}} x_{\hat{i},\hat{j}}} \tag{29}$$

or

$$\widehat{E(X_{i,j})} = \frac{\sum_{\hat{i} \in C} x_{\hat{i},j}}{\sum_{\hat{i},\hat{j}} x_{\hat{i},\hat{j}}} \sum_{\hat{j} \in P} x_{i,\hat{j}} \tag{30}$$

The first term on the right-hand side of the last equation estimates the probability that the product j is exported (share of j with respect to global export volume) and the second term estimates the expected exported volume of the country i.

Now we can put:

$$M(i,j) = \begin{cases} 1 & \text{if } \frac{x_{i,j}}{E(\widehat{X_{i,j}})} \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{31}$$

The term $\frac{x_{i,j}}{E(\widehat{X_{i,j}})}$ in 64 most commonly refers to an index, called the Balassa index or RCA (revealed comparative advantage).

The above definition means that, if a country i has revealed comparative advantage with respect to the product j, then $M_{i,j}$ will be one and zero otherwise which is intuitively appealing. On the other hand, if we make a theoretical observation, by looking at special cases we might have surprises. For example, if the exported product profile of a country i is a *multiple* of a country $\hat{i}$, than they will have the same values in their associated rows in M, implying the same ranking in ECI. Let us assume two countries with same population size and exported products, but the first one having two times more export volume. In this case, the same ranking might be distorted, since due to its presumable efficiency, the first country is likely to have a more efficient and innovative production process than the second one. As far as we understand the original method, we think that this critical point (at least theoretically) should be kept in mind, although from an empirical point of view it seems to represent minority cases.

3.4.2 An Alternative Information Matrix and Engineering of S

As mentioned before, although the original method with zero-one matrix (in other words $Y = M$) is proven to be useful, one can have the view to let data talk as much as possible and avoid losing any information or model mismatch risk to a certain extend. Among others, we would like to introduce an alternative way to rule out possible distortion, mentioned at the end of the last section.

Let us denote the export volume of country i in product j by $x_{i,j}$. The question is what is a meaningful way of defining the weights of S. Since we would like to compare different sizes of countries, it makes sense to rule out the linear performance impact due to size. Hence, we will start with the export volume of each country i in product j *per capita* which will be denoted by $y_{i,j}$ defining the elements of the information matrix Y.

As mentioned before, the reciprocal of components of the row vector in the original method $K_{0,\cdot} = (K_{0,1}, K_{0,2}, ..., K_{0,n})$ can be interpreted as a concentration measure of products with respect to countries. Low ubiquity implies high concentration of products (with respect to countries). Relying on the spirit of this interpretation, we can replace the vector $K_{0,\cdot}$ with the row vector $g(Y)$ by:

$$g(Y) = (\frac{(\sum_i y_{i,1})^2}{\sum_i y_{i,1}^2}, \frac{(\sum_i y_{i,2})^2}{\sum_i y_{i,2}^2}, ..., \frac{(\sum_i y_{i,n})^2}{\sum_i y_{i,n}^2}) \tag{32}$$

Note that if put $Y = M$ is zero one matrix as in the original method, we would get the previously defined ubiquity of the products back in this special case so that $g(Y) = K_{0,\cdot}$.

Now we can define a diagonal matrix as a building block of S.

$$D(j,j) = \frac{1}{g(Y)_j} \tag{33}$$

$D(j,j)$ defined as above is known as normalized Herfindahl-Hirschman-Index (HHI-Index).

Now, assume that we have two countries i and $\tilde{i}$ with exactly the same exported products and volumes per capita given by the rows i and $\tilde{i}$ of the Matrix Y respectively. If we scale the exported volumes from i by a positive *multiple*, we will change the concentration of all its products (which are by assumption same products of $\tilde{i}$).

In the original method, common knowledge of countries i and $\tilde{i}$ shared with respect to specific exported product j are the same if either they both export that product or they both don't export that product which is decided by the revealed comparative advantage. In other words if $M(i,j) = M(\tilde{i},j)$ which is pragmatic. However, we can have a more granular approach. In constructing $S(i,\tilde{i})$ based on zero-one matrix information M, the matrix operation standard (which is the "sum of the reciprocal of the ubiquity of joint products i with $\tilde{i}$) and has a conceivable interpretation (which is the "expected length of joint know-how$i$ with $\tilde{i}$"). The question is if we leave the zero-one word, is it possible to define a matrix operation for construction of S so that its elements can be interpreted as "expected length of joint know-how$i$ with $\tilde{i}$"?

Let us sketch the idea of how we would like to proceed in order to achieve a positive answer to this question. If we would put $S = Y * D * Y^{tr}$ directly we cannot interpret the elements of S without ambiguity.

A work around is to write

$$S = A \circ^f A, \tag{34}$$

with

$$A = Y_N * D^{\frac{1}{2}}. \tag{35}$$

where

$$Y_N = Y * D_N. \tag{36}$$

Here, D_N is an additional normalizing diagonal matrix with which we can interpret the elements $Y_N(i,j)$ of Y as the probability that i has one unit of indivisible manufacturing step under different prototype probabilistic assumptions reflecting the behavior of market players. D_N will be determined by backward engineering in such a way that the mentioned interpretation based on a consistent probability space that exists.

In this way, the elements $S(i,\hat{i})$ will have a clear interpretation as "common knowledge shared by i and $\hat{i}$" under the given probabilistic assumption reflected by its associated binary function.

Let us start with the above program.

Assume that we have a general network universe that covers all the value chain of existing as well as future products that are waiting to be discovered (or innovated, depending on the point of view). Assume that "indivisible" labor steps contributing to the global value of a product j conducted by some country can be decomposed in small indivisible outputs $H(j) = \{u_{j,1}, u_{j,2}, ..., u_{j,K(j)}\}$, which might very well include multiple copies in terms of know-how, if there are countries producing the same intermediate output. By unifying the indivisible copy of equivalent outputs we obtain set of "know-hows" $V(j) = \{v_{j,1}, v_{j,2}, ..., v_{j,N(j)}\}$, where each $v_{j,2}$ has different output. The set $V(j)$ can be interpreted as an image set of some function defined on $H(j)$ mapping into $V(j)$ in such a way that the equivalent outputs have the same image. Which is somehow idealized and assumes that equivalent outputs require exactly the same knowledge and rules out different ways of manufacturing. We chose this pragmatic mindset in order to have a shortcut and avoid notational overflow. Under these assumptions, the set of know-hows $W(i,j)$ of each country i on a product j can be interpreted as a subset of $V(j)$ and the union of all know-hows $W(i,j)$ over all countries must be equal to $V(j)$:

$$\cup_{i \leq m} W(i,j) = V(j). \tag{37}$$

Assume that the elements $Y(i,j)$ of Y are proportional to the probability that country i owns an indivisible unit of know-how $v_{j,l}$ in $V(j)$ with a proportionality factor N_j depending only on j :

$$\mathbf{P}[W(i,j)] = Y(i,j)N(j) =: Y_N(i,j), \ \forall i \leq m. \tag{38}$$

The above assumption means that a share of a country i on product j per capita times N_j is the probability that the country i owns any indivisible know-how $v_{j,l} \in V(j)$. This assumption holds in idealized global market conditions (e.g. high efficiency in the labor market as well as the market for goods), and reasonable to start with as an approximation.

The assumption 38 can be written in compact form as

$$\mathbf{P}[W(i,j)] = Y * D_N(i,j) \tag{39}$$

Where D_N is nxn is a diagonal matrix which scales each column j with N_j :

$$D_N(j,j) = N_j, \tag{40}$$

in order to have consistent probability space on the set of existing know-how for the product j.

The factors N_j will depend on the underlying probabilistic assumption and will be defined in the following.

The question is how to determine the element of the diagonal matrix D_N in the definition of matrix Y_N. This depends on the choice of the underlying probabilistic dependency structure of countries given a product which can be expressed by the symmetric binary function f defined on the set of the pairs of positive real numbers. In following we will introduce three prototypes binary functions $f_l(Y_N(i,j), Y_N^{tr}(j,\tilde{i}))),\ l = 1,2,3$ indicating the probability that countries i and $\tilde{i}$ have common knowledge on product j

$$\mathbf{P}[W(i,j) \cap W(\tilde{i},j)] = f_l(Y_N(i,j), Y_N^{tr}(j,\tilde{i})),\ l = 1,2,3, \tag{41}$$

each under its associated probabilistic assumption:

1) *Independence assumption* ($S^{(1)}$)

$$f_1(Y_N(i,j), Y_N^{tr}(j,\tilde{i})) = \begin{cases} Y_N(i,j) & \text{if } i = \tilde{i} \\ Y_N(i,j)Y_N^{tr}(j,\hat{i}) & \text{otherwise} \end{cases} \tag{42}$$

2) *Absolute dependency* ($S^{(2)}$)

$$f_2(Y_N(i,j), Y_N^{tr}(j,\tilde{i})) = \min(Y_N(i,j), Y_N^{tr}(j,\tilde{i})) \tag{43}$$

3) *Absolute complementary* ($S^{(3)}$)

$$f_3(Y_N(i,k), Y_N^{tr}(j,\tilde{i})) = \begin{cases} Y_N(i,j) & \text{if } i = \hat{i} \\ 0 & \text{otherwise} \end{cases} \tag{44}$$

The Independence assumption says that in the evolution of the know-how process of countries, the market players of the countries accumulate their know-how independently. The second assumption implies the dominance of the country exporting more per capita compared to the countries importing less. In other words, if i exports more of j than $\tilde{i}$, then i knows everything that $\tilde{i}$ knows about j. The last assumption means that the knowledge about the products are complementary which would in some sense imply absolute coordination of exporting countries. We think that reality is a mixture of the above assumptions even in a more granular form.

The diagonal matrix D_N in the definition of Y_N depends on the above assumption and can be determined with the boundary condition:

$$\mathbf{P}[\cup_{i \le m} W(i,j)] = \mathbf{P}[V(j)] = 1, \tag{45}$$

in order ensure that we have in fact a probability measures on $V(j)$ under the above prototype assumptions.

The last two can be determined straight forward:

Absolute dependency case:

$$N_2(j,j) = \frac{1}{\max_{i \le m}(Y(i,j))} \tag{46}$$

Absolute complementary case:

$$N_3(j,j) = \frac{1}{\sum_{i\leq m}(Y(i,j))} \tag{47}$$

independence case is bit more involving.
Using the following equation for independent events A_1 and A_2 recursively:

$$\mathbf{P}[A_1 \cup A_2] = \mathbf{P}[A_1] + \mathbf{P}[A_2] - \mathbf{P}[A_1]\mathbf{P}[A_2] \tag{48}$$

we can show that the boundary condition 45 requires the solution of following polynomial equation:

$$1-x\sum_{i\leq m}(Y(i,j)) + x^2 \sum_{i_1<i_2} Y(i_1,j)Y(i_2,j) - x^3 \sum_{i_1<i_2<i_3} Y(i_1,j)Y(i_2,j)Y(i_3,j)+... = 0 \tag{49}$$

which is equivalent to:

$$(1 - Y(1,j)x)(1 - Y(2,j)x)(1 - Y(3,j)x)... = 0 \tag{50}$$

The above polynomial has n solutions from which only one of them makes sense (the other solutions would lead to an event with a larger probability than one). Namely $x = \frac{1}{\max_{i\leq m}(Y(i,j))}$ which leads to put:

$$N_1(j,j) = \frac{1}{\max_{i\leq m}(Y(i,j))} \tag{51}$$

Until now we took care of probabilities of sharing one unit of know-how without assigning a value to them. Any know-how has some degree of intellectual complexity or severity to access. In order to incorporate this issue, we will use the nxn diagonal matrix D where each diagonal element $D(j,j)$ indicates the concentration of product j with respect to countries. The result we are targeting is that each element $S(i,\hat{i})$ indicates the "expected value of common know-how" shared by i and $\hat{i}$. The intermediate step for this purpose is to define a matrix $A(i,j)$ indicating the expected value of know-how on product j owned by the country i :

$$A = Y_N * D^{\frac{1}{2}}. \tag{52}$$

The equation above says that the expected value of know-how on product j owned by the country i is the probability that country i owns one unit of know-how, times the severity of the know-how which is similar to the well-known formula "expected loss=Probabilty of Default times the Severity (or Loss given Default) in credit risk. Please note that taking the square root of D (instead of D) makes sense if we follow the path of the construction of D (which is composition of quadratic elements). Formally we have:

$$A(i,j) = \mathbf{E}[D^{\frac{1}{2}}(j,j)1_{W(i,j)}] \tag{53}$$

Where $1_{W(i,j)}$ is the indicator function which takes the value 1 on the set $W(i,j)$ and zero otherwise.

Now we can define matrix operation $\circ^f$, so that $S(i.\hat{\imath}) = A \circ^f A^{tr}(i,\hat{\imath})$ is the "expected value of knowledge shared by i and $\hat{\imath}$. This can be achieved directly by putting:

$$S = A \circ^f A^{tr} = \sum_{k \leq n} f(A(i,k), A^{tr}(k,\hat{\imath})), \tag{54}$$

Note that in all three cases $S = A \circ^f A^{tr}$ can be decomposed as

$$S = \sum_{j \leq n} A(.,j) \circ^f A^{tr}(.,j), \tag{55}$$

where $A(.,j)$ is the jth column of A and each element $A(i,j)$ means the expected value of the know-how of country i on the product j.

It can be verified directly that each addend $A(.,j) \circ^f A^{tr}(.,j)$ can be represented as co-variance of some random variables implying the positivity of $A(.,j) \circ^f A^{tr}(.,j)$ and hence, the positivity of S in all three cases. This property holds also for S based on mixed assumptions, where for each product the probability that a country i owns one unit of the nondivisible value chain is a convex combination of the prototype assumptions.

Finally, by using the normalizing step 21 and 22 we can obtain the desired transition probability matrices governing the associated Markov Chains.

Intuitively, none of the above three assumptions occur in reality in their formulated extreme form as above and standalone they are not realistic. The reality is rather a mixture of them depending on e.g. average age of the innovations or evolution of know-how spread. In the evolution of manufacturing of products, know-how share, outsourcing, and other activities will drive the weights of the above prototypes for the final S. Of course, from a statistical point of view, when determining the parameter of mixture over-fitting should be avoided. For the calibration of the mixture, a machine learning approach can be an alternative.

Let us demonstrate the transition probability matrices under the assumptions $S^{(1)}$, $S^{(2)}$ and $S^{(3)}$ by a simple special example which might help to understand the implications.

Example 1

Let us assume that we have three countries and two products where each raw (export per capita) is multiple of the others :

$$Y = \begin{pmatrix} & p_1 & p_2 \\ c_1 & 8 & 4 \\ c_2 & 4 & 2 \\ c_3 & 2 & 1 \end{pmatrix}$$

Both products have the same consecrations represented by diagonal matrix D:

$$D = \begin{array}{c} \\ p_1 \\ p_2 \end{array} \begin{pmatrix} p_1 & p_2 \\ d & 0 \\ 0 & d \end{pmatrix}$$

We obtain following transition probabilities $P^{(1)}$, $P^{(2)}$, and $P^{(3)}$ ::

$$P^{(1)} = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \end{array} \begin{pmatrix} c_1 & c_2 & c_3 \\ p_1 & p_2 & p_3 \\ p_1 & p_2 & p_3 \\ p_1 & p_2 & p_3 \end{pmatrix}$$

Where $p_1 > p_2 > p_3$

$$P^{(2)} = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \end{array} \begin{pmatrix} c_1 & c_2 & c_3 \\ p_1 & p_2 & p_3 \\ q_1 & q_1 & q_2 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Where $p_1 > p_2 > p_3$ and $q_1 > q_2$
and

$$P^{(3)} = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \end{array} \begin{pmatrix} c_1 & c_2 & c_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Please note that above if follow the path of original process wit M we would obtain

$$P = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \end{array} \begin{pmatrix} c_1 & c_2 & c_3 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Comparing to the transition probability matrix P generated by the original process, the $P^{(1)}$, $P^{(2)}$, and $P^{(3)}$ contain additional information depending on the assumption. Especially (coming back to our comment at the end of subsection 3.4.1,) $P^{(2)}$ seems to be sensible with respect to our remark at the end of 3.4.1.

3.4.3 Economic Performance Equivalent (EPE) (or Compatible) Import Information Matrix

In the last two subsections, we have discussed the building blocks of ranking based on two possible information matrices for ranking purposes. Namely, the original zero-one matrix M and the trade volumes of exported products per capita Y. M is purely based on the trade volumes of exported products X, whereas Y needs the population size in addition. Please recall that our main objective is to merge export to import information which should add value to the quality of the ranking system based on solely export. A naive approach would be to imitate everything we did in the last section with an import. Assume that we are given a import matrix X^{im}, where its each element $x^{im}_{i,j}$ is the the trade volume of imported product j by country i. If we copy the recipe given by the formulas 30 and 64 directly assuming positive $x^{im}_{i,j}$, then due to incompatibility (with respect to what we expect in direction) merging process will not make sense. On the other hand, if we change the sign of X^{im}, we will not have Markov Chain not to mention positivity of associated S.

However, we can generate compatible M^{im} by using the same estimator given by the formula 30 for positive signed import trade volumes, but changing the direction of inequality (!) in formula 64. Or equivalently, we copy the recipe given by the formulas 30 and 64 directly assuming positive $x^{im}_{i,j}$, but replace the ones with zero and zeros with ones.

Based on this approach, $M^{im}(i,j) = 0$ means that j is a significant import product for the country i. A high number of zeros in the column j means high demand (or lower better lower demand/supply relation) for product j which at the same time means a low number of ones in column j. Hence $M^{im}(i,j) = 0$ (!) with a low number of ones in the jth column should rather imply an unfavorable ranking of i (i is buying something which everybody wants to buy). In the case of export, $M(i,j) = 0$ with a low number of ones in column j should also imply a rather unfavorable ranking for i (due to the low ubiquity of j which i doesn't own). This shows the compatibility of suggested M^{ex} with export zero one matrix in this case. Similarly, let $M^{im}(i,j) = 0$, but with a high number of ones in the jth column. This means j is a significant import of product for i which nobody wants and therefore, i is adding comparably higher value in its manufacturing. Hence, we expect a rather favorable ranking in this case. In the case of the export zero-one matrix M, not exporting comparably high ubiquity product should rather be more favorable, which sows compatibility also in this case. The compatibility in other cases (e.g $M^{im}(i,j) = 1$, with low number of zeros) can be verified similarly.

If we leave the zero one world and would like to have a compatible import information matrix, again we cannot just use the negative signed version of Y^{im} per capita. However, given the positive version of Y^{im}, we can first determine the maximal import trade volume per capita with in the importing countries:

$$y^{max,im}_j = \max_i (y^{im}_{i,j}). \tag{56}$$

Next step is define compatable information matrix for import, which can be achieved by matrix $Y^{im,com}$ define by:

$$y^{im,com}_{i,j} = y^{max,im}_j - y^{im}_j \tag{57}$$

Thereafter the application of the whole program in last section delivers compatible positive symmetric matrix S^{im} .

In summary, the resulting merging should imply e.g. "buying cheap and selling expensive" gives more favorable ranking than "buying expensive and selling expensive". This expectation from a ranking system requires not only compatibility of import information matrix with the export information matrix, but also at the same time coherency of ranking method, which we will treat in section 5.

4 Stochastic Counterpart of S and Interpretation of the ECI

The transformation of symmetric weight matrices S_C or S_P on a graph (first playground) to transition probability matrices P_C and P_P respectively on the one hand and characterization of Random Walks on a Graph (second playground) on the other hand show that we don't lose any information (up to constant scaling factor) by switching from one playground to the other. In this section, we will examine whether we can gain additional insight by switching to Random Walk on Graph as a stochastic counterpart of S_C or S_P. In this sense, we will answer the question of why it makes sense to rank the complexity of the countries with the eigenvector corresponding to the second largest eigenvalue in stochastic as well as in information-theoretical terms.

In Appendix, we justified that Markov Chain governed by P is reversible (Random Walk on a Graph) and has positive eigenvalues. Using this fact we will introduce the Lemmas 4.1 resp. 4.2. delivering the rank behavior of stochastic processes driven by the Random walk on Graph resp. behavior of the transition probabilities $P^t(\imath, \tilde{i})$ for large t, where second eigenvector plays a dominating role. By combining these lemmas, we will establish the relationship between Shannon entropy and second eigenvector (lemma 4.3) which says that the ECI(besides some pathological cases), is nothing else than the rank of the asymptotic Shannon entropy (each country corresponds to a row which has an entropy at time t).

4.1 Settings

Given the state space Z and the set of all paths

$$Z^{\mathbf{N}} = \{\omega = (\omega_t)_{t\in\mathbf{N}} \mid, \omega_t \in Z\} =: \Omega, \tag{58}$$

we will say that the canonical process $(S_t)_{t\in\mathbf{N}}$ on $(\Omega, \mathbf{P})$ (endowed with an filtration) defined by

$$S_t(\omega) = \omega_t \tag{59}$$

is a time homogeneous Markov Chain governed by P if

$$\mathbf{P}[S_{t+1} = \tilde{i} \mid S_t = i, S_{t-1} = i_{t-1}, ...S_0 = i_0] = P(i, \tilde{i}),\ \forall t \in \mathbf{N} \tag{60}$$

Above equation says that S is a stochastic process without a memory.

Often there is necessity to specify the initial distribution $\mu = (\mu(1), \mu(2)...\mu(m))$ of a Markov Chain S_t which we will shortly denote with ω_t^μ instead of S_t^μ (unless it causes confusion) indicating:

$$\mathbf{P}[\omega_0^\mu = i] = \mu(i) \tag{61}$$

By using the well known Chapman-Kolmogorov equation, we can directly verify that a Markov Chain ω_t^μ with a initial distribution μ governed by a transition probability matrix P satisfies the following equation:

$$\mathbf{P}[\omega_t^\mu = \tilde{i}] = \sum_{i \le m} \mu(i) \sum_{i_1 < \ldots < i_t} P(i_1, i_2)P(i_2, i_3)...P(i_{t-1}, i) = P^t(i, \tilde{i}). \qquad (62)$$

where $i_1 = i, i_t = \tilde{i}$.

Hence, in terms of matrix operations

$$\mathbf{P}[\omega_t^\mu = \tilde{i}] = (\mu * P^t)_{\tilde{i}} \qquad (63)$$

Where P^t is the tth power of P and $(v)_{\tilde{i}}$ denotes the $\tilde{i}$th component of the raw vector v.

For special choice of initial distribution δ_i defined by

$$\delta_i(\tilde{i}) = \begin{cases} 1 & \text{if } \ i = \tilde{i} \\ 0 & \text{otherwise} \end{cases} \qquad (64)$$

we can directly verify that

$$\mathbf{P}[\omega_t^{\delta_i} = \tilde{i}] = (P^t(i, .))_{\tilde{i}} = P^t(i, \tilde{i}) \qquad (65)$$

Any function defined on Z induces a stochastic process f_t define by

$$f_t(\omega) = f(\omega_t). \qquad (66)$$

We can directly verify that the expected value of f_t given the initial distribution μ can be obtained in with the following matrix operations:

$$\mathbf{E}[f_t^\mu] = \mathbf{E}_\mu[f_t] = \mu * P^t * f. \qquad (67)$$

4.2 Role of e_2 for Large t

Let us first list our assumptions in this subsection:

1) The Markov Chain governed by P is *ergodic* (which ensures a unique *stationary* probability measure) and *reversible*, hence it is a Random Walk on a Graph. The definitions and justification of these assumptions in our case are presented in the Appendix.

2) In order to exclude the degenerated cases, we will assume that P is invertible and all its eigenvalues $(\lambda_i)_{i \le m}$ are all different:

$$\lambda_i \neq 0 \ \forall i \text{ and } \lambda_i \neq \lambda_{\tilde{i}} \text{ for } i \neq \tilde{i} \qquad (68)$$

3) Further, we will assume that the vector e_2 has no zero components.

$$e_{2,i} \neq 0 \ \forall i \leq m \tag{69}$$

The last two assumptions are mild from practitioner's point of view. For example, if we work with zero one information matrices, with large number of countries and large number of products comparing to number of countries (which is satisfied by the number of countries and products considered in ECI), we expect a negligible likelihood of cases contradicting the above assumptions, which is even less restrictive if we work with real valued information matrices.

4) The stochastic counterpart of the method of reflection has appealing properties leading to coherent interpretations. An important fact which will play a key role is that the eigenvalues P are positive so that we have.

$$1 = \lambda_1 > \lambda_2 > \lambda_3 ..... > \lambda_m \geq 0 \tag{70}$$

The justification of this fact in our case is given in the Appendix.

and with assumption 2) we obtain:

$$1 = \lambda_1 > \lambda_2 > \lambda_3 ..... > \lambda_m > 0 \tag{71}$$

Given a pay-off function $X : Z -> \mathbf{R}$ of the states, we can define a stochastic pay-off process by putting:

$$X_t^\mu(\omega) = X(\omega_t^\mu), \tag{72}$$

X can be represented as a row vector $X = (X(1), X(2), ..., X(m))$ and we can write:

$$P^t * X^{tr} = (\mathbf{E}[X_t^{\delta_1}], \mathbf{E}[X_t^{\delta_2}] .... \mathbf{E}[X_t^{\delta_m}])^{tr}, \tag{73}$$

The method of reflection computes the time dependent expected pay-off column vector $u(t)$ defined by

$$u(t) = (\mathbf{E}[X_t^{\delta_1}], \mathbf{E}[X_t^{\delta_2}] .... \mathbf{E}[X_t^{\delta_m}])^{tr} \tag{74}$$

with a special choice of pay-off function defined by $X(i) = k_{i,0}$.

Let us denote the unique stationary measure with π.

For any initial distribution μ we have

$$\lim_{t \to \infty} X_t^\mu = Y \tag{75}$$

in distribution (!), where Y is a random variable with $P[Y = i] = \pi(i)$.

Hence, we have

$$\lim_{t \to \infty} u(t) = \text{constant} \cdot (1, 1, ..., 1)^{tr} = \text{constant} \cdot e_1 \tag{76}$$

independent of the initial state of the Markov Chain.

The constant can be determined by

$$\text{constant} = E_\pi(X) = \sum \pi(c_i)X(c_i) \tag{77}$$

Let us represent $u(t)$ in terms $(e_i)_{i\leq m}$:

$$u(t) = \alpha_1 e_1 + \sum_{i\geq 2} \lambda_i^t \alpha_i e_i, \tag{78}$$

where λ_i are eigenvalues of eigenvectors e_i in descending order.

The original reasoning of Hidalgo et al for the choice e_2 as a ranking standard was the domination of variance of the system by the variance of e_2 .

But what does this mean?

We interpret the above reasoning by observing that the prediction at time t of the performance vector $u(t+1)$ at next period $t+1$ can be approximated by using second largest eigenvalue and its associated eigenvectors as follows:

$$u(t+1) \approx u(t) + \alpha_2(\lambda_2^{t+1} - \lambda_2^t)e_2 \tag{79}$$

with a negligible error term for large large t.

Hence, the variance of the prediction based on the information at time t is dominated by the second eigenvalue and its associated eigenvector:

$$\text{Var}[u(t+1) - u(t)] \approx \alpha_2^2(\lambda_2^{t+1} - \lambda_2^t)^2 Var[e_2] \tag{80}$$

for large t. Where the variance is taken with respect to uniform measure on Z.

In Hidalgo et al the pay-off is defined by special choice. Namely,

$$X(i) = k_{i,0} \tag{81}$$

The observations give hope that normalized and centralized second eigenvector associated with second largest eigenvalue is a suitable complexity index of the countries.

Statistically, assume that we have a large portfolio of independent individuals each starting from some initial state having the same pay-off function but *iid* Markov Chains driving the dynamic of their pay-offs. Depending on the distribution of the initial states of the individuals, we can easily determine the pay-off function on portfolio level which leads to a similar conclusion that the variance of pay-off on a portfolio level is asymptotically dominated by the second eigenvector of the Markov Chain. This statistical observation gives additional hope but doesn't answer the question "what are we ranking?". In statistical terms, we can say even more which allows us an additional insight for statistical interpretation. Namely, there exist t_0 such that the rank of expected pay-off process defined as above is strictly (not only asymptotically!) dominated for all t larger than t_0.

At this stage, we would like to spend few words in order to avoid confusions which might be caused by the following small abuse of notation. We used the notation i for countries but also for the elements of *static* state space Z of Markov Chain generated by the information matrix in order to avoid rotational overflow. As an element of Z, i represents the "observed" accumulated know-how of i (given by the exported products of i) at time $t = 0$. Hence, in our context $P^t(i, \tilde{i})$ represents the probability that the country i has the accumulated know-how $\tilde{i}$ (which is the initial know-how state of $\tilde{i}$) at time t, given the initial accumulated know-how of i (which is also denoted by i).

Now we can start with our program of interpretations.

Let $R : \mathbf{R}^m \setminus E \to \{1, 2....m\}$ be a ranking function which simply ranks the components of the vectors in standard coordinate system according to relation "$\geq$" in $\mathbf{R}$, excluding the elemets of $E := \{x \in \mathbf{R}^m : \exists i, j \leq m \text{ with } x_i = x_j\}$ which is only a zero set with respect to Lebesgue measure on $\mathbf{R}^m$.

We have the following lemma:

Lemma 4.1

Let $X : Z : -> \mathbf{R}$ be a random variable with

$$X = \sum_{i \leq m} \alpha_i e_i \tag{82}$$

with $\alpha_2 \neq 0$ and $u(t)$ as in 74

Let $e_{i,j}$ be the jth components of e_i in standard coordinate system.

Assume that $e_{2,j}$ are pairwise different (which means $e_2 \in \mathbf{R}^m \setminus E$) and put:

$$\min\{\mid e_{2,j} - e_{2,k} \mid : j,\ k \leq m,\ j \neq k\} = \epsilon > 0. \tag{83}$$

Then there exist $t_0 > 0$ such that:

$$R(u(t)) = R(\alpha_1 e_1 + \alpha_2 \lambda_2^t e_2) = R(\alpha_2 \lambda_2^t e_2) = R(\alpha_2 e_2) \tag{84}$$

for all $t > t_0$.

Proof. First note that $u(t)$ can be expressed in terms of matrix operations:

$$u(t) = P^t * X^{tr}, \tag{85}$$

where P is the transition probability matrix resulting from subsection 3.3 for modular scheme.

For the proof the lemma we first observe the following simple properties of ranking function R.

i) By multiplying with strictly positive constant c and adding any constant eigenvector $\alpha e_1 = \alpha(1, 1...., 1)^{tr}$ to a vector we do not change the rank of a vector (rank neutrality).

$$R(x) = R(cx + \alpha e_1) \tag{86}$$

ii) Put

$$\rho(x) = \min\{\mid x_j - x_k \mid : j,\ k \leq m,\ j \neq k\} \tag{87}$$

and assume that $\rho(x) = \delta > 0$. Then for any vector $a \in \mathbf{R}^m$ such that $\max\{| a_j |: j \leq m,\} < \frac{\delta}{2}$

$$R(x) = R(x + a) \tag{88}$$

By assumption we have $\lim_{t\to\infty} = \frac{\lambda_i^t}{\lambda_2^t} = 0$ for all $i \geq 3$ implies that there exists a t_0 so that:

$$| \frac{\lambda_i^t \alpha_i}{\lambda_2^t \alpha_2} e_{i,j} | < \frac{\epsilon}{2(m-2)} \tag{89}$$

for all i with $3 \leq i \leq m$ and $t \geq t_0$
Hence,

$$| \sum_{3\leq i\leq m} \frac{\lambda_i^t \alpha_i}{\lambda_2^t \alpha_2} e_{i,j} | < \frac{\epsilon}{2}, \tag{90}$$

hence

$$| \sum_{3\leq i\leq m} \lambda_i^t \alpha_i e_{i,j} | < \frac{\epsilon \mid \lambda_2^t \alpha_2 \mid}{2}, \tag{91}$$

The jth component $u(t)_j$ can be expressed as

$$u(t)_j = \sum_i \lambda^t \alpha_i e_{i,j} = \alpha_1 + \lambda_2^t \alpha_2 e_{2,j} + \sum_{3\leq i\leq m} \lambda^t \alpha_i e_{i,j} \tag{92}$$

Hence, by 83, $x(t) = \alpha_1 e_1 + \lambda_2^t \alpha_2 e_2$ and $a(t) = \sum_{3\leq i\leq m} \lambda_i^t \alpha_i e_i$ satisfy the condition of the observation ii) with $\delta = \epsilon \mid \lambda_2^t \alpha_2 \mid$, for all $t \geq t_0$

Now, by corollary 7.11.1 in Appendix we know that all eigenvalues have positive sign. Therefore, we can assume without loss of generality $\text{sign}(\lambda_2^t \alpha_2) = +1$ and using i) and ii) we obtain:

$$R(u(t)) = R(x(t) + a(t)) = R(x(t)) = R(\alpha_2 e_2) \tag{93}$$

Which is the claim of the lemma.

□

Lemma 4.1 says that the rank of expected pay-off does not change after t_0 and will be $R(\alpha_2 e_2)$.

Now, assume that we have a pay-off of a bet given by the random variable X. For any t, the ith component of $(P^t * X)_i$ is the expected pay-off of the know-how at time t given that we bet on i at time $t = 0$. As t tends to infinity, the expected values will all converge to the same value which makes the bet at infinity "indifferent" or independent of our choice at $t = 0$. However, for any initial betting state i, as t gets large, we will have that the rank of expected pay-off will be the same as the rank of the second eigenvector corresponding to the second largest eigenvalue in spite of the indifference at infinity which is some how surprising.

The next lemma 4.2 gives us further inside with a different method. Note that that lemma 4.2 follows from lemma 4.1: We could take the vectors in the canonical basis as pay-off functions, and than conclude 4.2 in a few steps. But with the following proof, we want to introduce the reader already to the so called eigenvalue representation. This representation needs reversibility ergodicity positivity, which we can assume to be satisfied by P.

It is well known that second eigenvector e_2 is related to the optimal cut problem(see e.g. [18] and [19]) to separate the vertexes (which are countries in our case) into two classes C^+ and C^-. Without being specific on this issue let us put:

$$C^+ = \{i : \operatorname{sign}(e_{2,i}) = +\} \tag{94}$$

$$C^- = \{i : \operatorname{sign}(e_{2,i}) = -\} \tag{95}$$

We will see that the above classification makes sense from probabilistic point of view.

Lemma 4.2

Let i and $\tilde{i}$ be two countries. Then the following statements are equivalent:

i) The rank of i is higher than i :

$$e_{2,i} > e_{2,\tilde{i}} \tag{96}$$

ii) There exist $t_0 \geq 0$ so that for all h in C^+ we have

$$P^t(i,h) > P^t(\tilde{i},h),\ \forall t \geq t_0 \tag{97}$$

iii) There exist $t_0 \geq 0$ so that for all h in C^- we have

$$P^t(i,h) < P^t(\tilde{i},h),\ \forall t \geq t_0 \tag{98}$$

Proof. We will only prove the equivalence of i) with ii) (the proof of the equivalence of i) with iii) is analog). For this purpose we will use the "eigenvalue representation" of transition probabilities:

$$\frac{P^t(i,h)}{\pi(h)} - 1 = \sum_{k\geq 2} \lambda_k^t e_{k,i} e_{k,h} \tag{99}$$

$$\frac{P^t(\tilde{i},h)}{\pi(h)} - 1 = \sum_{k\geq 2} \lambda_k^t e_{k,\tilde{i}} e_{k,h} \tag{100}$$

Where π is the unique stationary measure of P.

The proof of above eigenvalue representation can be found in chapter12 of [20].

By multiplying both sides of above equations with $\frac{1}{\lambda_2^t e_{2,h}}$ we obtain:

$$\frac{1}{\lambda_2^t e_{2,h}}(\frac{P^t(i,h)}{\pi(h)}-1)=\frac{1}{\lambda_2^t e_{2,h}}\sum_{k\geq 2}\lambda_i^t e_{k,i}e_{k,h} \tag{101}$$

$$\frac{1}{\lambda_2^t e_{2,h}}(\frac{P^t(\tilde{i},h)}{\pi(h)}-1)=\frac{1}{\lambda_2^t e_{2,h}}\sum_{k\geq 2}\lambda_i^t e_{k,\tilde{i}}e_{k,h} \tag{102}$$

Since $\lambda_i \geq 0$ and $\lambda_2 > \lambda_i, \forall i > 2$, (see Appendix), the right hand side of above equations converge to $e_{2,i}$ and $e_{2,\tilde{i}}$ respectively. But, by assumption $e_{2,i} > e_{2,\tilde{i}}$ which implies that there exist t_0 so that

$$\frac{1}{\lambda_2^t e_{2,h}}(\frac{P^t(i,h)}{\pi(h)}-1)>\frac{1}{\lambda_2^t e_{2,h}}(\frac{P^t(\tilde{i},h)}{\pi(h)}-1) \tag{103}$$

for all $t \geq t_0$.

Hence,

$$P^t(i,h) > P^t(\tilde{i},h),\ \forall t \geq t_0 \tag{104}$$

$\square$

Note that the proof actually shows:

$$\frac{P^t(i,h)}{\pi(h)}=1+\sum_{k\geq 2}\lambda_k^t e_{k,i}e_{k,h}=1+\lambda^t e_{2,i}e_{2,h}+o(\lambda^t) \tag{105}$$

and therefore

$$P^t(i,h)=\pi(h)+\lambda^t e_{2,i}e_{2,h}\pi(h)+o(\lambda^t) \tag{106}$$

which can be seen directly with the same type of arguments we used in the proof.

Let's now give an interpretation of the ECI by using lemma 4.2. The Markov Chain governed by P assumes static state space which is naive and therefore doesn't describe a real-world dynamic of information transition realistically. However, we can get a bit more motivation out of 4.2. let's speculate a bit about what $P^t(i,h)$ might roughly approximate in a more realistic transition-dynamic. Let's denote units of information in h by $I(h)$. First, we weigh the average value of this information from the perspective of i by $\frac{\pi(h)}{\pi(i)}$ (capacity of sources in h relative to the capacity of sources in i), analogue we define the average value of $I(h)$ from perspective of $\tilde{i}$. Let $f^t_{I(h),i}$ denote the observed frequency of information transition from h to i along transitions of length t countries, analogue we define $f^t_{I(h),\tilde{i}}$. We can also think of t as the length of the walks starting from any time $s \geq 0$. Due time homogeneity of the

Markov Chain $f^t_{I(h),i}$ respectively $f^t_{I(h),\tilde{i}}$, can be approximated by $P^t(h,i)$ respectively $P^t(h,\tilde{i})$ independent of starting time $s \geq 0$. Hence, we approximate the expected value of information transition from h to i by $\frac{\pi(h)}{\pi(i)}P^t(h,i)$, analogue we approximate the expected value of information transition from h to $\tilde{i}$. Note that since the chain P is a random walk on a finite graph, we have $\frac{\pi(h)}{\pi(i)}P^t(h,i) = P^t(i,h)$, $\frac{\pi(h)}{\pi(\tilde{i})}P^t(h,\tilde{i}) = P^t(\tilde{i},h)$. Hence, by lemma 4.2, for large t, $\frac{\pi(h)}{\pi(i)}P^t(h,i) > \frac{\pi(h)}{\pi(\tilde{i})}P^t(h,\tilde{i})$ for all h in C^+ if and only if i is higher ranked than $\tilde{i}$ by the ECI.

In the context of the discussion above, we think of $P(\tilde{i},i)$ as an approximation of the "direct" information transition-rate from $\tilde{i}$ to i, and on the basis of this approximation, we further approximate the information transition-rate from $\tilde{i}$ to i (over the walk $i_1,...,i_t$), by $P(i_1,i_2)P(i_2,i_3)...P(i_{t-1,i_t})$. Hence considering all possible walks of countries, $P^t(\tilde{i},i)$ approximates the total information transition-rate over t-walks from $\tilde{i}$ to i.

Note that we always said "the rank of e_2". Obviously, before one can rank with e_2, one has to solve the sign problem first, i.e one has to decide which class is C^+ and which class is C^-. We will see later in section 5 different prototype examples where there is no reasonable choice of sign. In example 2, there are two classes and each element of one class has a equivalent counterpart in the other class with same absolute value of their associated e_2, but with an opposite sign. We will now introduce an alternative ranking method by means of Shannon Entropy which solves this problem at least to a certain extend.

Shannon entropy is an established theory when it comes to study information contend of objects formally (see [21] for further reading). To our knowledge, to rank vertices in a network is not a widely applied method. The following result shows that there is a general information theoretic meaning of e_2 in a large class of networks modeled by a Markov Chain (i.e ergodic,reversible,finite Markov chains) and it says that the ranking of Shannon Entropy of transitions probabilities of countries with an initial state i are given by e_2 for large t if the logarithm of stationary measure is not orthogonal to e_2 . More precisely we have the following lemma:

.

Lemma 4.3

Assume that assumption of the lemma 4.1 for the vector $X = (\log_2(\pi(1)),...,\log_2(\pi(m))^{tr}$ are satisfied. Then with an adequate choice of sign we have:

$$R((H(t,1),H(t,2),...,H(t,m))^{tr}) = R(e_2) \tag{107}$$

for all large enough t.

Here, $H(t,i)$ denotes the Shannon Entropy of the probability measure $P^t(i,.)$:

$$H(t,i) = -\sum_{h\le m} P^t(i,h)\log_2(P^t(i,h)) \tag{108}$$

Proof. It's enough to show that, for each fixed state i

$$H(t,i) = \alpha_1 + \alpha_2\lambda_2^t e_{2,i} + o(\lambda_2^t) \tag{109}$$

Where α_1 and α_2 are constants independent of t and $o(x)$ refers to Landau-Symbol (small o of x).

If the statement holds for the natural logarithm then it holds also for $\log_2$ as $c\log = \log_2$ for a constant c. Hence, in the following we can switch from $\log_2$ to natural logarithm which we shortly denote with log.

For each fixed i the eigevalue decomposition says:

$$P^t(i,h) = \pi(h) + \pi(h)\sum_{k\ge 2}\lambda_k^t e_{k,i}e_{k,h} \tag{110}$$

Since by assumption λ_2 is strictly larger than all the other eigenvalues except λ_1 we obtain:

$$P^t(i,h) = \pi(h) + \pi(h)\lambda_2^t e_{2,i}e_{2,h} + o(\lambda_2^t) = \pi(h)(1 + \lambda_2^t e_{2,i}e_{2,h} + \frac{o(\lambda_2^t)}{\pi(h)}) \tag{111}$$

Hence we can write,

$$\log(P^t(i,h)) = \log(\pi(h)) + \log(1 + \lambda_2^t e_{2,i}e_{2,h} + o(\lambda_2^t)) \tag{112}$$

By the first order Taylor approximation of $\log(1+x) = x + o(x)$ together with calculation rules with "small o" results in:

$$\log(P^t(i,h)) = \log(\pi(h)) + \lambda_2^t e_{2,i}e_{2,h} + o(\lambda_2^t) \tag{113}$$

Note that from the proof of lemma 4.1, for the pay-off function $X(i) = \log(\pi(i))$ we obtain

$$\sum_{h\le m} P^t(i,h)\log(\pi(h)) = \alpha_1 + \lambda_2^t\alpha_2 e_{2,i} + o(\lambda_2^t) \tag{114}$$

By combining 114 with 113 and using the calculation rules with "small o" results in:

$$\sum_{h\le m} P^t(i,h)\log(P^t(i,h))) = \alpha_1 + \lambda_2^t\alpha_2 e_{2,i} + o(\lambda_2^t) + \sum_{h\le m} P^t(i,h)(\lambda_2^t e_{2,i}e_{2,h} + o(\lambda_2^t)) \tag{115}$$

The last term on the right hand side of 115 can be written as:

$$\sum_{h\leq m} P^t(i,h)(\lambda_2^t e_{2,i}e_{2,h} + o(\lambda_2^t)) = \lambda_2^t e_{2,i}\sum_{h\leq m} P^t(i,h)e_{2,h} + \sum_{h\leq m} P^t(i,h)o(\lambda_2^t)) \tag{116}$$

Which can be summarized as:

$$\sum_{h\leq m} P^t(i,h)(\lambda_2^t e_{2,i}e_{2,h} + o(\lambda_2^t)) = \lambda_2^{2t}e_{2,i}^2 + o(\lambda_2^t)) \tag{117}$$

Finally by combining 117 with 115 we obtain:

$$\sum_{h\leq m} P^t(i,h)\log(P^t(i,h))) = \alpha_1+\lambda_2^t\alpha_2 e_{2,i}+\lambda_2^{2t}e_{2,i}^2+o(\lambda_2^t)+o(\lambda_2^t) = \alpha_1+\lambda_2^t\alpha_2 e_{2,i}+o(\lambda_2^t) \tag{118}$$

which proves 109

□

It is worth to spend some time on lemma 4.3. The lemma 4.3 says that if the α_2 of the $X := (\log_2(\pi(1)),...,\log_2(\pi(m))^{tr}$ is non-zero, then it ranks exactly same as e_2 for large enough t. Here by ranking with $(H(t,1),H(t,2),...,H(t,m))^{tr}$ we mean ranking in descending order. In other words lower components indicate more favorable ranking.

The question is does ranking with $(H(t,1),H(t,2),...,H(t,m))^{tr}$ have advantages comparing e_2. The answer is yes it has some advantages. Firstly we do not have the mentioned sign problem. The components of $(H(t,1),H(t,2),...,H(t,m))^{tr}$ are always positive and as just mentioned the lower values indicate better ranking. Secondly, if for example we have even number of counties and the transition probability matrix P is symmetric at the same time with symmetric diagonal, then there will be pairs of two equivalent countries. But in this cases α_2 of X is zero and e_2 would rank one member of each equivalent pair with positive and the other member with negative sign which is not coherent. But in this cases ranking of $(H(t,1),H(t,2),...,H(t,m))^{tr}$ (in descending order) jumps to ranking of positive valued version of e_3 automatically which is coherent.

We will now explicate why the notion "complexity" makes sense for e_2 to certain extend.

For a fixed i and t let $(\chi_n^i)_{n\in\mathbf{N}}$ be *iid* copies of Random Walks on a Graph with an initial state i governed by the transition probability matrix P and let Z^∞ be the set of all strings $s = (\tilde{i}_1, \tilde{i}_2, ...)$ in Z^∞ realized by $(\chi_n^i)_{n\in\mathbf{N}}$ at time t. Further let S^n be projection defined on Z^∞ :

$$S^n(s) = (\tilde{i}_1, \tilde{i}_2, ..., \tilde{i}_n) \tag{119}$$

The original Kolmogorov Complexity (plane-comlexity) is related to e_2 by the following theorem (theorem 148 in [22]):

Theorem 4.4

Let $t > 0$ be fixed and C be the original Kolmogorov Complexity defined on $S^n(Z^\infty)$ for all n. Then
$\lim \frac{C(S^n(s))}{n} = H(P^t(i,1), ..., P^t(i,m))$ almost suerly

From above theorem and the lemmas 4.1 and 4.3, it follows that for large n and large t, Kolmogorov Complexities $C(\omega(n)^i)$ are ranked according to second eigenvector e_2, if α_2 of the vector $X = (\log_2(\pi(1)), ..., \log_2(\pi(m))^{tr}$ is non-zero.

We saw that if e_2 is not orthogonal to $X = (\log_2(\pi(1)), ..., \log_2(\pi(m))^{tr}$ ranking of countries with $H(t,.)$ for large enough t is equivalent to ranking by e_2 which is in practice almost always the case. Although $H(t,.)$ or e_2 can be called a complexity measures, in section 5 we will see that they can interpreted as coherent driver of sustainability only to a certain extend.

The question is can we do better than the alternative complexity measure introduced by the lemma 4.3. In order to answer this question observe that the proven equation 109 in the proof lemma 4.3 can be written as

$$H(t,i) - \alpha_1 = \alpha_2 \lambda_2^t e_{2,i} + o(\lambda_2^t). \tag{120}$$

For each fixed t, by summing the above equation over $s \le t$ we obtain:

$$\sum_{s \le t} H(t,i) - t\alpha_1 = \sum_{s \le t} o(\lambda_2^s) \tag{121}$$

But since $0 \le \lambda_2 < 1$, the right hand side of the above equation dominated by a convergent geometric series. Hence, the limits $\lim_{t\to\infty} \sum_{s\le t} H(t,i) - t\alpha_1$ exists. Given the countries i and $\tilde{i}$, with different product profile, in practice, the likelihood of having same limits for these countries becomes marginal for large dimensional matrices with large enough product/country ratio (tested numerically). Hence, assuming that the mentioned limits for countries with different product profile are different , with simple we conclude the rank of $\sum_{s\le t} H(t,.)^{tr} - t\alpha_1$ remains same for all t larger than some t_0. These observations together with the rank neutrality property with respect adding a constant vector as well as scaling by positive number implies following lemma:

Lemma 4.5

Let $AH(t,.)$ be the average entropy at t :

$$AH(t) = \frac{\sum_{s\le t} H(s,.)}{t} \tag{122}$$

If the limits $\lim_{t\to\infty} \sum_{s\le t} H(t,i) - t\alpha_1$ are pairwise different then there exist t_0 so that

$$R(H(t_1,.)^{tr}) = R(H(t_2,.)^{tr}), \ \forall t_1, t_2 \ge t_0 \tag{123}$$

By the above lemma we gain another alternative ranking by $R(H(t_0,.)^{tr}$ for large t_0 which compensates the lack of memory of a Markov Chain to a certain extend.

Finally, we would like to note that an interesting and further research object is the speed of rank convergence or more precisely a criteria for how large t_0 related to the lemmas 4.1, 4.2, 4.3 and 4.5 should be.

5 Coherence of Sustainability Index and ECI

Does a rating model do, what we expect it to do?

If somebody claims that what he wants to do will have sustainable *positive net-impact*, then there are at least three questions which comes in mind:

Who are the stakeholders or for whom should it have net-positive impact?

For how many generations or how long is the positive impact going to last?

What is the quality of his performance?

These interdependent and difficult questions are easier to answer when we manage the risk of a company. Because the main stakeholders are the owners of the company and risk management focuses mainly on shareholder's view which is broadly accepted, since the mainstream consciousness is that the owners are at risk. Hence the boss ("as a risk-taker") will tell the risk manager what is sustainable and what is not. Due to regulatory paradigm shift in the last decades, also clients gained more and more weight as stakeholders.

It took decades to measure the risk of insurance companies coherently (e.g. with coherent risk measure such as Expected Shortfall), instead of traditional non-coherent risk measures such as Variance (see e.g.Artzner, Philippe, et al. [14]). Meanwhile, the concept is a well-accepted mindset and in Switzerland, the regulatory capital required of insurance companies is determined periodically based on *coherent risk measure*. Empirical evidence of sustainability ranking with respect to key sustainability indicators is important. But its coherency is not less important. Because as in risk management the "more" incoherent a model is, the more it allows "model arbitrage" (which related to non-intended incentives), distortions and might even give undesired incentives, which also play role when it comes to its predictive power. Sustainability management is in this sense"more difficult risk management". Because comparing to "boss tells you what is sustainable and what is not sustainable," there is a long way to go for global *consensus* about the above questions when it comes to measuring of net-positive impact in a context related to the above questions. It is not only about "I exist" (or survive) but also about long-term coexistence in a broader view. Hence, we think that a sustainable measurement instrument should be a learning tool in a framework that can improve *coherently* and give a better quality answers when the global consensus improves. This mindset will be more clear in motivating subsection 5.1.

ECI is to a certain extent eligible index giving an inside for the accumulated productive know-how justified from a theoretical point of view (see lemma 4.1, lemma 4.2 and 4.3 for statistic, probabilistic and information-theoretical properties) and has empirically justified explanatory power for macroeconomic factors such as growth rate of GDP as well. However, since our objective is to merge available value chain information of countries improving the quality of sustainability index, it makes sense to examine the coherency behavior of ECI and compare it to an alternative. This question will be treated in subsection 5.2. In subsection 5.3 we give a shortlist of alternatives without further exploration. In the last subsection, we will discuss the coherency behavior of e_2 and an alternative with respect to amalgamated versions.

5.1 Motivation; Expert in Box

Please note that our intention is by far not to generate an ideology or whatsoever. It is rather about what it means modeling of a sustainability index (or ranking) which we hope that it can grow in a sustainable manner itself.

Quoting George Box "all the models are wrong, but some are useful". In our opinion, the degree of usefulness of the sustainability index related to its *predictive power*. But not less important than its predictive power is what kind of *incentives* the model gives and to what degree it allows a *model arbitrage* which are at same time hybrid notions to predictive power. Especially, a notion such as *sustainability index* is expected to map a proxy of *consensus* in the society which we mentioned at the beginning of this section.

Imagine a diligent expert who is commissioned with a challenging and noble task to test a developed sustainability ranking. Presumably, his first goal would be to understand, to which degree and why a model is not "coherent" to what he expects. If the answer to "why" convinces him then he can even learn from the model and its outputs add value to his expertise. But if the surprises don't convince him (which often the case), he would like to understand to what extent in order to be able to engineer an improved model. For the achievement of this objective a list of coherence condition which is free of contradictions itself could help. This list can be classified as *minimal conditions* and *proxy conditions for common sense.* Out of such a mixture of conditions we can assume statements logical implications of this mixture and statement which cannot be decided by the mixture of these conditions.

Let us display the class of possible coherence scenarios and how an expert could feel about between his consistency conditions the model output.

Output vs. coherence statements	**Expert's assessment**
In line with coherence list	Ok
Logical inconsistencies (minimal requirement)	Incoherence of first order
Not inline with common sense	Incoherence of second order
Not decidable	Indifferent

Incoherencies to some extend can be expected. However, if they dominate and especially occur systematically in case of first-order (which can be even classified as an inconsistency), then it becomes a serious quality issue also related to its predictive power. The second case cannot be classified as an inconsistency, but depending on the requirement, it can be classified as strong or weak incoherence and from our point of view, it makes sense to keep an eye on it in the engineering of a sustainability index. Let us illustrate this point with an example. Assume that in our observed data we have countries who add value to their economy by an intelligent manufacturing process which increases their diversity over time with negligible expense making these countries also highly competitive with respect to mentioned products (not necessarily new) or helping them to improve their internal economy. Now, assume contrary to our expectation (which is probably higher sustainability index) we have a negative change of the countries sustainability index based only zero one information matrix delivering *less favorable ranking*. The impact of such scenarios on the correlation between the "sustainability index" (which we are testing) and GDP will be rather negative, which could also be an issue concerning the predictive power of the " sustainability index" since the productivity of a nation is a significant driver of GDP. However, this expectation of an expert can be classified rather as common sense and subject to consensus in society or within the experts and the severity of "inconsistencies" depends on how strong the conditions are. Although at this stage we cannot rule it out completely, that the sustainability index which the expert is testing is intelligent and able to handle some kind of causal relationship in some constellations and such an upgrade of the country (which increases his diversity) in some cases might even be justified depending on the constellation of the zero-one information matrix. But we rather think that the mentioned upgrade can be justified rather by additional information sources and therefore implausible which is incoherence of second order.

Finally, there will be also cases that cannot be assessed as "coherent" or "not coherent" which might need adjustment of condition list. This will be clearer in the next subsection.

5.2 Some Coherency Conditions and ECI

Our plan in this subsection will be to formulate coherency conditions based on the information matrix. By doing so and for the sake of simplicity we will work with zero one information matrices. Please also note that our approach in this subsection doesn't give a final answer for this challenging issue. Nor we claim that the below coherency conditions are free of redundancy. In this subsection, we rather cover the quality of the sustainability index with respect to the economic & risk management points of view. In this subsection, in order to focus on the ideas, we will formulate coherence conditions on the zero-one information matrices.

Let us assume that $I(M, i)$ is a one-period sustainability ranking of countries based on the trade information matrix and start with what we would expect from such an index from the logical point of view.

C1 (*Condition of Equivalence1*): Assume we are given a permutation σ of countries and M_1 and M_2 are two information matrices with same dimensions such that.

If $M_1(i, j) = M_2(\sigma(i), j),\ \forall i \leq m, j \leq n$ then

$$I(M_1, i) = I(M_2, \sigma(i)), \forall i \leq m \tag{124}$$

C2 (*Condition of Equivalence 2*): Assume we are given a permutation ρ of products such that:

$M_1(i,j) = M_2(i,\rho(j)),\ \forall i \leq m, j \leq n$ then

$$I(M_1, i) = I(M_2, i),\ \forall i \leq m \tag{125}$$

C1 says that if we permute the countries and permute the rankings accordingly than the countries should have the same rankings. C1 implies for example that if two countries have exactly same values in their associated rows, then their ranking will be the same:

If $M(i,.) = M(\hat{i},.)$ then

$$I(M, i) = I(M, \hat{i}). \tag{126}$$

C2 says that permutation of the products has neutral effect on the ranking of the countries.

The conditions C1 and C2 can be seen as a minimal requirement and the second eigenvector e_2 satisfies both of them. However, die situation becomes critical if we require the following coherency conditions which is from our point of view can be also classified as a minimal requirement.

C3 (*Condition of Equivalence 3*): A formal formulation of this condition is subject to our working paper and we rather prefer an example which illustrates what we mean with the condition of equivalence.

Example 2

Let us assume we have the following information matrix with six countries and seven products.

$$M = \begin{pmatrix} & p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 \\ c_1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ c_2 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ c_3 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ c_4 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ c_5 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ c_6 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

where ith row represent exported product of country i and takes the value one at product j if j is exported.

Let us display the second eigenvector.

$$e_2 = \begin{pmatrix} -0.56 & -0.41 & 0 - 0.15 & 0.15 & 0.41 & 0.56 \end{pmatrix}$$

From the above information matrix we would expect the same ranking for the first and the last country and we see no reason why the first and the last country (which is the sixth country) should be rated differently from economic and/or risk management point of view. They have both equal number of products with a isomorphic ubiquity profile. The ubiquity of the first product exported by the first country and the ubiquity of the last product exported by the last country are one. The other exported product by first and exported product by the last country also have the same ubiquity. We can observe analog relation e.g. between the second and the fifth country. More formally, the above information matrix can be split in to two isomorphic (weight preserving and one-to-one) subgraphs $C = \{c_1, c_2, c_3\}$ and $\hat{C} = \{c_4, c_5, c_6\}$ which are connected over c_3 and c_4 symmetrically. However, second eigenvector reflects this natural expectation only in absolute value. This is also why our statement 98 in the probabilistic lemma 4.2 is restricted to countries (denoted by h) in the class C^+ which is fixed beforehand depending on the choice of sign of e_2. On the other hand, observe that in this example, the Entropy vector $H(t, .)$ ranks the countries according to positive version of e_3 for all t larger than some t_0 which is the same ranking as the absolute value of e_2 (see also the comments after the proof of 4.3). Note that in this example $AH(t)$ for large t ranks same as $H(t)$ for large t. Where as the diagonal of S ranks the fist and last country as highest and the others countries as indifferent which is a bit unsatisfactory from complexity (or similarity) point of view.

C4 (*Strong Diversity Condition*): In terms of sensitivity, if we add a product to a country i without an expense, then its position should not become less favorable than before by pairwise comparison to other countries. Formally let i be fixed and $\hat{M}$ be the information matrix which has only one additional product on behalf of i in comparison to M. Then

$$\{\hat{i} : I(M, i) \geq I(M, \hat{i}\} \subseteq \{\hat{i} : I(\hat{M}, i) \geq I(\hat{M}, \hat{i})\}. \tag{127}$$

By using 126 as consequence of C1 and applying C4 successively, we can conclude that if a country beside its low ubiquity products, can also compete with all the other countries also with respect to large ubiquity products due to its sub-processes as side-products with a negligible expense, then this country should not have less favorable ranking than others. Or simply a country which exports everything should not have a less favorable ranking then others.

Now let us assume that we have again six countries and seven products but with different product mixes:

Example 3

$$M = \begin{pmatrix} & p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 \\ c_1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ c_2 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ c_3 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ c_4 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ c_5 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ c_6 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Let us display the the second eigenvector.

$$e_2 = \begin{pmatrix} 0.06 & 0.57 & 0.57 & -034 & -0.34 & -0.34 \end{pmatrix}$$

If we accept the condition C4, one can ask why the first country should have a rank lower than country 2 and 3, given the sign as above. If we change the sign of e_2 we have lower rank of first country comparing to forth, fifth and sixth country. Hence in any case, the country that exports everything gets considerably lower ranking than some other countries. Note that this example is different from the previous example 2 and cannot be corrected by e_3 or $H(t,.)$ for large t.

Ranking with the diagonal of S resp. $AH(t,.)$ for large t is monotone deceasing resp. increasing which is coherent.

Although ranking by diagonal of S resp. $AH(t,.)$ for large t are highly correlated, in some cases we observed slight incoherence of $AH(t,.)$ for large t with respect to diversity condition.

If the above condition is too strong for an expert, then he can try to asses above example with the weaker diversity condition:

C5 (*Week Diversity Condition*): If we add a product to a country i without an expense, then its ranking should not be less favorable than before. This condition doesn't exclude that the position between i comparing to some other country $\hat{i}$ may be less favorable by adding a additional product. It only says that overall ranking is at least as good as before. Clearly C4 implies C5. However, the implication of C4 by C5 seems to be very challenging and we could neither prove it nor find a counter example until now. Hence, an expert who has not overcome this challenge jet cannot say that the C5 is rejected by e_2 which rather falls rather in the category of "Not decidable" from his point of view.

What about ubiquity? Does a decline of ubiquity of products owned by a country results in a better ranking?

For the above question we like formulate a possible version of complexity condition in terms of sensitivity:

C6 (*Condition of Ubiquity*): Let j be a product exported by i and $\hat{i}$. Assume that for some reasons $\hat{i}$ decides not to export the product j next period anymore. Assume also that the product profile of all the other countries remains the same. In this case the ubiquity of j declines and we postulate that any country h which did not export j and had lower rank than i will still have lower rank than i after the decision of $\hat{i}$.

Does the ECI satisfy the ubiquity condition C6 above?
The answer is no which can be observed by the following counterexample:

Example 4

For the first period put:

$$M = \begin{array}{cc} & \begin{array}{ccccccc} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & \end{array} \\ \begin{array}{c} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{array} & \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

the corresponding eigenvector of the random walk is ,

$$e_2 = \begin{pmatrix} 0.45 & 0.19 & -0.06 & -0.29 & -0.5 & -0.65 \end{pmatrix}$$

We strongly believe that choice of sign is adequate in the first period!
Now we remove the product 4 from the country 1 and obtain:

$$M = \begin{array}{cc} & \begin{array}{ccccccc} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & \end{array} \\ \begin{array}{c} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{array} & \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

$$e_2 = \begin{pmatrix} 0.67 & 0.1 & -0.2 & -0.44 & -0.34 & -0.43 \end{pmatrix}$$

The ubiquity of the product 4 decreased after removing it from country 1. The country 4 still has the product 4, but its rank became lower than the rank of 5 (5 didn't have the product 4, and had a lower rank than 4 in the previous period). If we change the sign of e_2 than for example the rank of country 2 would be lower than the rank of 6 , although country 2 still has the product 4 and the country 6 didn't have the product 4 in the previous period). However the diagonal of S resp $AH(t,.)$ for large t are monotone deceasing resp. increasing which can be seen as coherent in this example.

Note that since in case of increasing resp. reducing the diversity of a country (without using a new product), has an effect on the ubiquity of the products, there is relationship between the notion of diversity and complexity. We would like to go one step further without exploring this issue further.

As mentioned before, e_2 satisfies the coherency conditions C1 and C2. But, we think from economic point of view, satisfying this natural conditions cannot be enough for its comparable high correlations with GDP. Next lemma gives a one good reason why this is the case.

If the inclusion relation $\subseteq$ is a total order between the products of the countries, ECI delivers coherent results if sign of e_2 is chosen adequately. This means e_2 delivers coherent ranking if we can find some permutations of countries so that we can achieve that the first country exports all the products that the second country export, the second country exports all the products that the third country exports, third country exports all the products that the forth country exports,.., and so on.

Formally:

Lemma 5.1

Assume that the products $(c_i)_{i\leq m}$ of countries $(i)_{i\leq m}$ satisfy

$$c_1 \supseteq c_2 \supseteq c_3 ... c_{m-1} \supseteq c_m. \tag{128}$$

Then with an adequate choice of sign we have

$$e_{2,1} \geq e_{2,2} \geq e_{2,3} ... e_{2,m-1} \geq e_{2,m}. \tag{129}$$

Proof. It turns out that this intuitively natural lemma is tricky. The idea is first to show that if a transpose X^{tr} of a pay-off function X is monotone decreasing (as the diagonal does), then $P * X^{tr}$ is also monotone decreasing. Thereafter, we can take any monotone decreasing pay-off function X which has nonzero α_2 in its eigenvalue decomposition. Finally, lemma 4.1 ensures that ranking of $\alpha_2 e_2$ is the same as the ranking of $P^t * X^{tr}, \forall t \geq t_0$ which is the ranking of the diagonal of P because of monotony of $P^t * X^{tr}$ which proves the claim.

In order to focus on the ideas we will assume that inclusions are strict (this restriction is not really necessary, it makes the proof easier to follow):

$$c_1 \supset c_2 \supset c_3 ... c_{m-1} \supset c_m \tag{130}$$

and prove

$$e_{2,1} > e_{2,2} > e_{2,3} ... e_{2,m-1} > e_{2,m}. \tag{131}$$

Please first observe that under the above condition we have:

$$P(i,k) = P(i,i), \ \forall \ k \leq i \tag{132}$$

$$P(i,i) > P(i+1,i+1), \ \forall i \tag{133}$$

132 can be verified directly from 130.
For 133 put

$$d_i = k_{i,0} \tag{134}$$

and

$$u_j = K_{j,0}. \tag{135}$$

Then by 130 we have

$$P(i,i) = \frac{a+b}{d_i} \tag{136}$$

and

$$P(i+1,i+1)=\frac{a}{d_{i+1}}, \tag{137}$$

where

$$a=\sum_{j\in c_{i+1}}\frac{1}{u_j} \tag{138}$$

$$b=\sum_{j\in c_i\setminus c_{i+1}}\frac{1}{u_j} \tag{139}$$

By 130 we also have

$$\frac{1}{u_j}<\frac{1}{i},\ \forall j\in c_{i+1}, \tag{140}$$

and

$$\frac{1}{u_j}\geq\frac{1}{i},\ \forall j\in c_i\setminus c_{i+1}, \tag{141}$$

above last two inequalities together with 138 resp.139 imply

$$a<\frac{d_{i+1}}{i}, \tag{142}$$

resp.

$$b\geq\frac{d_i-d_{i+1}}{i}, \tag{143}$$

Hence by 142 we have,

$$d_{i+1}a+\frac{d_{i+1}}{i}(d_i-d_{i+1})>d_{i+1}a+a(d_i-d_{i+1})=ad_i \tag{144}$$

Dividing the above expression by d_id_{i+1} we obtain

$$\frac{a}{d_i}+\frac{d_i-d_{i+1}}{id_i}>\frac{a}{d_{i+1}} \tag{145}$$

Hence, by 143 we have

$$\frac{a}{d_i}+\frac{b}{d_i}>\frac{a}{d_{i+1}}, \tag{146}$$

which implies 133

Note that since the rows $P(i,.)$ are probability measures 132 and 133 imply

$$P(i,k)=P(i+1,k)+\epsilon,\ \forall\ k\leq i \tag{147}$$

for some $\epsilon>0$ and

$$P(i,k)=P(i+1,k)-\delta_k,\ \forall\ k>i \tag{148}$$

for some positive δ_k with $\sum_{i<k\leq m} \delta_k = i\epsilon$.
Now assume that we are given strictly monotone decreasing X^{tr} :

$$X_1^{tr} > X_2^{tr} ... > X_m^{tr}. \tag{149}$$

Then we have

$$(P * X^{tr})_i = \sum_{k\leq i} P(i,k)X_k^{tr} + \sum_{i<k\leq m} P(i,k)X_k^{tr}. \tag{150}$$

Hence,

$$(P * X^{tr})_i = \sum_{k\leq i}(\epsilon + P(i+1,k))X_k^{tr} + \sum_{i<k\leq m} (P(i+1,k) - \delta_k)X_k^{tr} \tag{151}$$

for some positive δ_k with $\sum_{i<k\leq m} \delta_k = i\epsilon$.
But because of monotony of X^{tr} we also have

$$\sum_{k\leq i} \epsilon X_k^{tr} + \sum_{i<k\leq m} -\delta_k X_k^{tr} > 0 \tag{152}$$

Hence by combining 151 and 152 we obtain

$$(P * X^{tr})_i = \sum_{k\leq m} P(i,k)X_k^{tr} > \sum_{k\leq m} P(i+1,k)X_k^{tr} = (P * X^{tr})_{i+1} \tag{153}$$

which proves that $P * X_i^{tr}$ is decreasing in i.

□

Definition 1

We will say that the ranking I is coherent with respect to monotony if

$$c_1 \supseteq c_2 \supseteq c_3 ... c_{m-1} \supseteq c_m \tag{154}$$

implies that ranking according to I is monotone deceasing.

Note that in case of above monotony 154, it can be verified directly that the diagonal elements $S(i,i)$ of the weight matrix S are monotone decreasing (compare section 3). Hence, we have the following simple lemma:

Lemma 5.2

Ranking the countries with the diagonal of S is coherent with respect to monotony.

5.3 Alternative Ranking Methods

In this subsection we would like to note some alternatives to e_2 that might be for interest. The degree of coherence is an important driver of correlations between two alternative measures. Two highly coherent alternative measures are likely to have high correlation, so that the final choice will depend on other properties.

1) *Ranking according to rank of diagonal elements of* $S(i,i)$
2) *Ranking according to rank of* $H(t,.)$ *for large* t
3) *Ranking according to rank of* $AH(t,.)$ *for large* t

As mentioned before, there are different types of merging of supplementary information. In the following we will analyze Lie-trotter approach (amalgamation) by comparing the behavior of the amalgamated Random Walks on pre S-Level (induced by the export and import data) with respect to e_2 and above alternatives on different prototype examples. As already mentioned, we think know-how accumulation is well measured by the ECI (see lemmas 4.1, 4.2 and 4.3 for statistic, probabilistic and information theoretical interpretations). However, since we would like to be sure that the suggested amalgamation is an improvement that it is important to analyze the coherency behavior of alternative rankings after amalgamation.

5.4 Coherency of Ranking Methods After Amalgamation

In this subsection we would like to demonstrate how different ranking methods behave numerically after amalgamation of S_{ex} with S_{im}. Note that due to commutativity of amalgamation on pre S-Level (see subsection 6.2 for the definition of amalgamation on pre S-Level), all the conclusion in this subsection don't depend on the order of amalgamation.

For the *Coherence after Amalgamation with respect C3* we would like to present the following example:

Example 5

Put

$$M_{im} = \begin{pmatrix} & p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 \\ c_1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ c_2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ c_3 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ c_4 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ c_5 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ c_6 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

and

$$M_{ex} = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{array} \begin{pmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Let us display e_2 of Random Walk induced by the amalgamated $S_{ex,im}$ $(= S_{ex} *_{LT} S_{ex}\ )$ and diagonal elements v_D of $S_{ex,im}$.

$$e_2 = \begin{pmatrix} 0.43 & 0.44 & 0.35 & -0.35 & -0.44 & -0.43 \end{pmatrix}$$

and

$$v_D = \begin{pmatrix} 0.34 & 0.34 & 0.40 & 0.40 & 0.34 & 0.34 \end{pmatrix}$$

From coherence point of view, the first and the last country should be ranked equally with respect to M_{ex} and M_{im} before amalgamation. Moreover, both export and import information matrices segregate the countries in two classes. In one of the classes we have the first three and in the other class we have the last three countries respectively, where each element of one class has an equivalent counterpart in the other class and the equivalent pairs with respect to export and import are the same. Hence, due to symmetry of the of the problem, the amalgamated version should also segregate the countries in the same manner. Similar to prototype example 2 before amalgamation, the absolute vales of e_2 instead of e_2 is coherent with respect to C_3 which implies that $H(t,.)$ for large enough t ranks the amalgamated version in coherent manner. Note also that in this case, the displayed diagonal v_D as well as $AH(t,.)$ for large t meet also our coherency expectations with respect to C_3.

Coherence of Amalgamation with respect to diversity and monotony

The last prototype example and other numerical observations show that behavior of e_2 in case of amalgamated Random Walk show similar surprises as it was the case before amalgamation. We have shown in 5.1 in case of monotony of information matrices $e2$ behaves coherently if we rank them standalone. On other hand intuitively, if we amalgamate the a monotone export information matrix M_{ex} with an in some sense "neutral" M_{im}, it shouldn't have an impact on the ranking. Let us examine this with the following prototype example:

Example 6

$$M_{ex} = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{array} \begin{pmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$M_{im} = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{array} \begin{array}{ccccccc} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}$$

M_{ex} is monotone and due to special structure of M_{im}, we expect that M_{im} should not cause much distortion of any reasonable ranking of M_{ex}. Alternatively we can argue as follows: M_{im} defines two homogeneous equivalent classes; first three and last three countries. The equivalent pairs are (c_1, c_6), (c_2, c_5) and (c_3, c_4). Firstly, after amalgamation of monotone deceasing M_{ex} to M_{im} (which is due to commutativety the same as amalgamation of M_{im} to M_{ex}) should deliver monotone decreasing ranking within the classes. Secondly, due to equivalence of c_3 with c_4 before amalgamation, c_3 should be rank higher than c_4 after amalgamation of monotone decreasing M_{ex}. Analog we can argue that c_2 should be higher ranked than c_5 and c_1 should be higher ranked than c_6. Combining these two facts imply that the after amalgamation of monotone deceasing M_{ex} to M_{im} any reasonable ranking I should be monotone decreasing: $I(c_1) > I(c_2) > I(c_3)...I(c_6)$.

This expectation is satisfied by e_2 with the right choice of sign, regardless of "incoherence" ranking of e_2 applied on M_{im} in standalone case(!) as well as V_D :

$$e_2 = \begin{pmatrix} 0.321 & 0.316 & 0.308 & -0.477 & -0.485 & -0.488 \end{pmatrix}$$

and

$$v_D = \begin{pmatrix} 0.252 & 0.248 & 0.242 & 0.2020 & 0.2026 & 0.2019 \end{pmatrix}$$

The very light incoherence of diagonal at forth and fifth country is due to numerical error (warning by our software in Amalgamation process).

It is interesting to note that the "incoherence" ranking of e_2 applied on M_{im} before amalgamation becomes coherent after amalgamation with the monotone information matrix M_{ex}.

Finally we would like to note that $H(t,.)$ and $AH(t,.)$ applied on this example deliver similar results.

Next we would like to demonstrate the balancing behavior of ranking alternative rankings with the following two information matrices:

Example 7

$$M_{ex} = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{array} \begin{pmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$M_{im} = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{array} \begin{pmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

In this case again e_2 is only in absolute value "coherent". As in standalone similar cases $H(t,.)$ or $AH(t,.)$ for large t rates "coherently".

Let us look at v_D :

$$v_D = \begin{pmatrix} 0.27 & 0.23 & 0.22 & 0.22 & 0.23 & 0.27 \end{pmatrix}$$

v_D in this case preserves the desired symmetry and ranks the more balanced ones higher (first and last countries have a share in all products, whereas the middle ones exclude at least one product).

Without incorporation of import information, the last country would have the lowest rank. But after incorporation of import information it will be upgraded which is reasonable from economic sustainability point of view.

Please also note that this example is at the same time an additional prototype with respect to equivalence condition C3 which e_2 satisfies only in absolute value.

Remark 1

Finally, as amalgamated version of lemmas 5.1 and 5.2 we would like to note that e_2 of $P_{ex,im} := P_{ex} *_{LT} P_{im}$ as well as diagonal of $S_{exp.im} := S_{ex} *_{LT} S_{im}$ with monotone export and import information matrices with equivalent order relationship with respect to export and import information matrices behave numerically well. Here, equivalent order relationship means:

$$c_{ex,i} \supset c_{ex,\tilde{i}} \tag{155}$$

if and only if

$$c_{im,i} \supset c_{im,\tilde{i}} \tag{156}$$

where $c_{ex,i}$ and $c_{im,i}$ denote the exported and the imported (with respect to economic performance equivalent scheme!) products of country i respectively.

Our numerical investigations supports the idea. However, rigorous proof of this statement is a bit involving. A possible idea for example in case of amalgamated version of 5.1 is the following: First note that $P_{ex}^{\frac{1}{n}}$ resp. $P_{im}^{\frac{1}{n}}$ share same second eigenvector as P_{ex} resp. P_{em}. Next step would be to show that $P_{ex}^{\frac{1}{n}}$ resp. $P_{im}^{\frac{1}{n}}$ inherit the proven property that P_{ex} resp. P_{em} map the monotone decreasing pay-off functions into monotone decreasing pay-off functions as indicated in the proof of lemma 5.1. Taking this property to the limit in the definition of amalgamation would then deliver the proof.

5.5 Summary of Coherency Results

In this section we will summarize the coherency of e_2 as well as our alternative rankings. Since the conditions C1 and C2 are satisfied by all the alternatives, we will restrict our self on the other coherency conditions.

Coherency before Amalgamation
Following table gives an overview of coherency of ranking alternatives before amalgamation.

Coherency of Alternatives (before Amalgamation)

Ranking by	**C3**	**C4 or C5**	**C6**	**Monotony**
e_2	on example 2 not satisfied	on example 3 not satisfied	on example 4 not satisfied	yes (see lemma 5.1)
$H(t,.)$	coherent on example 2 (see comments to lemma 4.3)	analog to e_2 (see lemma 4.3)	analog to e_2 (see lemma 4.3)	analog to e_2 (see lemma 4.3)
$AH(t,.)$	analog to $H(t,.)$ (see comments to lemma 4.3)	coherent on example 3	coherent on example 4	analog to e_2 and $H(t,.)$
Diagonal of S	not equivalent to H(t,.) on example 2	equivalent to $AH(t,.)$ on example 3	coherent on example 4	yes (see lemma 5.2)

Please note that ranking by $H(t,.)$ or $AH(t,.)$ for large t should be conducted in descending order. In other words, the lower $H(t,i)$ or $AH(t,i)$ (for large t) are, the more favorable are the ranking according to $H(t,i)$ or $AH(t,i)$ for large t. The order of ranking methods in the above table reflects degree of coherency by we have observed in our numerical analyzes and theoretical considerations. As a summary we can say that $H(t,.)$ ranks slightly more coherent than e_2. Moreover, $AH(t,.)$ and diagonal of S behave quite harmonically to a large extend and rank lightly more coherent than e_2. and $H(t,.)$.

Coherency after Amalgamation
Amalgamation of import to export changes the game. The following table gives an overview of the behavior of different ranking methods Random based on amalgamation on pre S-Level.

	Coherency of Alternatives (after Amalgamation)		
Ranking by	**C3**	**Balancing Impact of Amalgamation of Import**	**Monotony**
e_2	analog to before amalgamation (see examples 5 and 7)	on example 7 only in absolute value satisfied	numerically satisfied (see also remark1)
$H(t,.)$ for large t	coherent on the examples 5 and 7	analog to e_2 (see lemma 4.3)	(see lemma 4.3)
$AH(t,.)$ for large t	analog to $H(t,.)$ (see comments to lemma 4.3)	coherent	coherent
Diagonal of S	yes	yes	yes

It is interesting to note that our further combination of prototype examples where e_2 behaves incoherent before amalgamation, behaves coherently after amalgamating monotone information matrix on which e_2 behaves coherently. On the other hand, prototype example 7 shows that e_2 might be risky if we would like to ensure balancing effect (or compensating "week" export with "strong" import). However this risk can be removed to a large extend by $H(t,.)$ for large enough t which is in some sense almost equivalent ranking method.

Finally we would like note that we also conducted simulations of rankings in order to measure the correlations of above ranking methods. It turns out that coherence of ranking methods plays important role for the correlation of two ranking methods and can be interpreted as systematic driver of a ranking systems. For example we can confirm that ranking by $AH(t,.)$ for large t is higher correlated with the diagonal of S than e_2 with diagonal of S:

$$\mathrm{cor}(R(e_2), R(\text{diagonal of } S)) < \mathrm{cor}(R(AH(t,.),), R(\text{diagonal of } S)) \tag{157}$$

As a summary of this section and due to introduced examples in this manuscript other numerical investigations we did, we conclude that regardless of eligibility of e_2 as complexity ranking method to a certain extend, the diagonal of standalone as well as the amalgamated version of weight matrix S seem to be more preferable from the coherency point of view. But on the cost of interpretations given by the lemmas 4.1, 4.2, 4.3 and the lemma 4.5. On the other hand, the alternatives $H(t,.)$ is a bit more and $AH(t,.)$ even more elaborate and numerical point of view, more complex to handle.

6 Integrating Import

The exported products in a closed system of trading countries give deep insight about what empowers them. However two countries exporting the same product with the same market share might very well have different value chain paths. Hence, an assessment based on only on their exported products might in some cases deliver distorted ranking. The question is not only what the countries export, it is also what they import and their internal trade. However, among others due to challenging data collection problem, we would like to postpone direct merging of internal trade information. Also, we think merging the import with the export information delivers already important insight about the internal trade.

As mentioned in 3.2, $P(i, \tilde{i})$ as "expected length" of a value chain of common products exported by i and $\tilde{i}$.

The question is how both information matrices M^{exp} and M^{im} or more generally Y^{exp} and Y^{im} should be mixed meaningfully, in order to obtain an improvement of the quality of ECI or other explanatory indicators of sustainability which is still holistic.

There are various possibilities. A direct approach is to look at the trade balance sheet and define an analog information matrix.

Formally (this is only in order to demonstrate the difficulties and it is not our suggestion):

$$M_{c,p}^{ex\ +\ im} = \begin{cases} 1, & \text{if } \frac{y_{i,j}^{ex} - y_{i,j}^{im}}{y_{i,j}^{ex}} \geq \text{benchmark} \\ \frac{1}{2}, & \text{if } \mid \frac{y_{i,j}^{ex} - y_{i,j}^{im}}{y_{i,j}^{ex}} \mid < \text{benchmark} \\ 0, & \text{otherwise} \end{cases} \tag{158}$$

Thereafter using the scheme in subsection 3.3 we can generate Random Walk on Graph and work and adapt the idea of ECI by taking second eigenvector or any other "eligible" ranking method.

Among others one drawback of above approach is that it requires a sensitive external "benchmark" parameter.

Other possibilities could be import content of export for each country and product which among requires tedious data preparation.

An approach which could avoid above drawbacks to construct a Markov Chain where each random steps is defined by first applying the export Random Walk P_1 and than by import Random Walk P_2 successively(see subsection 3.4.3 for compatible import information matrix) respectively. More precisely:

$$P_{ex,im} = P_1^{\alpha} * P_2^{1-\alpha}. \tag{159}$$

Here $0 \leq \alpha \leq 1$ and determines the weight of export and import respectively.

However, this obvious idea has its also drawbacks. The matrices P_1 and P_2 don't necessarily commute and the dynamic of resulting Markov Chain is dominated by P_1 or P_2 depending on the order of matrix multiplication and therefore the resulting second eigenvector depends strongly on this order, which we cannot decide without an ambiguity at this stage. Although in some cases there might be good reasons for the dominance of one Random Walk by the other one, we believe economic complexity or sustainability of a country grows organically and export and import trade of country are interactively growing process in which their order in the value chain cannot be identified without additional information or assumptions. Therefore our target is to model a sustainability index based on export and import data in an unbiased manner. We will see later that if the domination of export or import should be an issue, then it can be controlled by the parameter α anyway. Moreover, although Markov Chain governed by $P_1^{\alpha} * P_2^{1-\alpha}$ process is ergodic, it is not necessarily a Random Walk on Graph with positive eigenvalues. This is an another serious obstacle for interpretations in statistic probabilistic or information theoretical terms.

Next we will introduce a more appareling approach which is source of our main idea in the last subsection of this section, without further exploration of further alternatives.

6.1 Amalgamation by Lie-Trotter Product Formula

The above discussion leads to natural limiting approach which is known as Lie–Trotter product formula .

P_1 and P_2 are similar to a positive diagonal matrices D_1 and D_2 respectively:

$$P_1 = \phi_1 * D_1 * \phi_1^{-1} \tag{160}$$

and

$$P_2 = \phi_2 * D_2 * \phi_2^{-1} \tag{161}$$

Where ϕ_1 are orthogonal transformation matrices.
We can define

$$P_{1,n} = P_1^{\frac{1}{n}} = \phi_1 * D_1^{\frac{1}{n}} * \phi_1^{-1} \tag{162}$$

$$P_{2,n} = P_2^{\frac{1}{n}} = \phi_2 * D_2^{\frac{1}{n}} * \phi_2^{-1} \tag{163}$$

Put

$$P_{ex,im,n} = (P_{1,n} * P_{2,n})^n \tag{164}$$

$$P_{im,ex,n} = (P_{2,n} * P_{1,n})^n \tag{165}$$

If $P_{1,n}$ and $P_{2,n}$ are not commutative the $P_{ex,im,n}$ and $P_{im,ex,n}$ are not necessarily equal. However their limes coincide:

$$P_{ex,im} := \lim_{n\to\infty} P_{ex,im,n} = \lim_{n\to\infty} P_{im,ex,n} =: P_{im,ex} \tag{166}$$

In order to see this observe that

$$P_{ex,im,n} = P_{1,n} * P_{im,ex,n-1} * P_{2,n}. \tag{167}$$

which implies the commutativity of limes process due to

$$\lim_{n\to\infty} P_{1,n} = \mathrm{Id} \tag{168}$$

and

$$\lim_{n\to\infty} P_{2,n} = \mathrm{Id} \tag{169}$$

This means instead of merging discrete manner by merging infinite decimal manner we gain a commutativety of amalgamation process of export with import which does not carry ambiguity of order of amalgamation.

The process can be carried out also for P_1^{α} and $P_2^{1-\alpha}$

This order independent amalgamation will be denoted by $P_1 *_{LT} P_2$. However, although numerically, e_2 of amalgamation of import with export delivers acceptable results from coherency point of view, it does not necessarily represent a Random Walk on a Graph if P_1 and P_2 are not commutative. This motivates to introduce our favorite suggestion in the next section.

6.2 Amalgamation by Lie-Trotter Formula on pre S-level

If we use the binary operation $*_{LT}$ on a "S level" fallowed by normalizing the resulting matrix according to subsection 3.3 we to obtain a Random Walk on a Graph if the elements of amalgamated S are positive. More precisely we have the following theorem:

Theorem 6.1

let S_1 resp. S_2 be symmetrical weight matrices induced by the export and compatible import information matrices respectively. Then the amalgamated $S_{ex,im} = S_1 *_{LT} S_2$ is symmetric and has a positive eigenvalues (see Appendix). Moreover, if the elements of $S_{ex,im}$ are positive then the Markov Chain $P_{ex,im}$ obtained by normalization of $S_{ex,im}$ according to subsection 3.3 is reversible and represents a Random Walk on a Graph with positive eigenvalues.

The proof of symmetry and the positivity of eigenvalues of $S_{ex,im}$ and $P_{ex,im}$ can be found in Appendix. The second condition "positivity of the elements of $S_{ex,im}$" is from practical point of view mild. By generating random information matrices in a time consuming process, we observed that the likelihood of $S_{ex,im}$ having negative elements converges to zero as die size of information matrix tends to infinity and becomes negligible if we deal with information matrices with sizes comparable with the number of countries and products involved in ECI ranking.

7 Appendix

For a satisfactory interpretation of ECI or any sustainability index the stochastic counterpart of a graph according to 3.3 must satisfy some properties. In our context these properties can be listed as ergodicty, reversibility and that the second eigenvector of its transition probability matrix is positive. Although the literature is very reach in the area of graph theory and stochastic, we could not find self-contained a reference which is tailored for our purpose. The first objective of this section is to provide the reader with the mathematics of these properties for the non-amalgamated case. Our second objective to show that our main finding (which is the amalgamation of supplementary value chain information) also satisfy these properties. By doing so, we will exclude all " pathological" cases in order keep the formalism simple and focus on ideas.

7.1 Random Walk on a Graph

Random Walks on Graph are special Markov Chains with appealing properties such as ergodicity and reversibility which plays a important role for our interpretation of ECI. The Markov Chains governed by transition probabilty matrices according to 3.3 are special Random Walks on a Graph (positivity of their eigenvalues). However, in case of non-comutativity, we can only show that the product of $P_1 * P_2$ is ergodic. Reversibility and positivity is not necessarily satisfied. Nevertheless, idea behind behind behind the proof will be useful for our main result in the last subsection for amalgamation. Good news is that the nth root of Random Walks in this class remains in this class. We will also show that the reversibility and positivity of eigenvalues of $P_1 * P_2$ can be also expected if P_1 and P_2 are commutative.

Definition 2

Markov Chain on a state space Z governed by the transition probability matrix P is called ergodic if following conditions are satisfied:

$$\text{i. (connectivity): } \forall i, j \in Z : P^t(i, j) > 0 \text{ for some } t \tag{170}$$

$$\text{ii. (aperiodicity): } \forall i, j \in Z : \gcd\{t : P^t(i, j) > 0\} = 1 \tag{171}$$

Here the term gcd denotes the gratest common divisor.

In the literature of probability theory, the property connectivity is called irreducibility. Since we would like establish self contained bridge between probability theory and graph theory, the terminology connectivity seem us more appealing.

Intuitively, above definition tells that transition probability matrix of an ergodic Markov Chain mixes a large sample towards a unique stationary distribution which we will formalize next.

Definition 3

A a probability measure π on state space Z is called stationary with respect to transition probability matrix P if

$$\mu * P = \mu \tag{172}$$

Distribution of an ergodic Markov Chains converges to attracting (or stable) stationary distribution which is unique. Formally:

Theorem 7.1
Let Markov Chain on a state space Z governed by the transition probability matrix P be ergodic. Then there exists unique stationary distribution π so that

$$\lim_{t\to\infty} \mu P^t = \pi \tag{173}$$

for all initial distributions μ on the state space Z.

Note that uniqueness of stationary distribution does not imply the ergodicity in our sense, which we can demonstrate with the fallowing example.

Example 8
Put

$$P = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

The first state is and absorbing state (and does not mix properly) and we have

$$\lim_{t\to\infty} \mu P^t = (1, 0)$$

for all initial distributions μ on the state space $\{1, 2\}$

But the connectivity is not satisfied since $P^t(1, 2) = 0,\ \forall t \geq 0$.

This kind Markov Chains are automatically ruled out in our case because of special weight matrices define by the "intersection" of an information matrix M according to subsection 3.3.

Next we would like to formulate a useful a sufficient condition for aperiodicity which is in some sense a uniform (or strong) connectivity.

Lemma 7.2
Markov Chain governed by P is aperiodic if there exists t_0 so that

$$P^t(i, j) > 0\ \forall t \geq t_0, \forall i, j. \tag{174}$$

The proof of above lemma is straight forward and it implies directly the following corollary.

Corollary 7.2.1
Markov Chain on a state space Z governed by the transition probability matrix P is ergodic if there exist t_0 so that

$$P^t(i, j) > 0\ \forall t \geq t_0, \forall i, j. \tag{175}$$

We would like exclude any country in our data which does not trade anything, which means that each raw of the information matrix Y has at leas one non-zero element. Then Markov Chain governed by P and which is generated by M according subsection 3.3 for modular scheme satisfies:

$$P(i,i) > 0 \ \forall i. \tag{176}$$

This property will help us to prove the ergodicity of Markov Chains governed by the product of a product of transition probabilities. More precisely:

Lemma 7.3

Let P_1 resp. P_2 be connective Markov Chains generated by the information matrices according modular scheme we introduced in section 3.3. Then P_1, P_2 and $P_1 * P_2$ are all ergodic.

Proof. First we introduce a convenient notation which we will also use later for amalgamation of Random Walks on a Graph.

Fix a i and j and put

$$A[0,t](i,j) = \{I(t) = \{i_0, i_1, ..., i_t\} : i_0 = i, i_t = j\} \tag{177}$$

and for $I(t) = \{i_0, i_1, ..., i_t\}$ in $A[0,t](i,j)$

$$P_1(I) = P_1(i_0,i_1)P_1(i_1,i_2)P_1(i_2,i_3)...P_1(i_{t-1},i_t). \tag{178}$$

Then we have

$$P_1^{t_1+1}(i,j) = \sum_{I(t+1)\in A[0,t+1](i,j)} P_1(I(t+1)) = \sum_{I(t)\in A[0,t](i,k)} P_1(I(t))P_1(k,j) \tag{179}$$

But the right hand side of the above equation satisfies:

$$\sum_{k\in Z, I(t)\in A[0,t](i,k)} P_1(I(t))P_1(k,j) \geq \sum_{I(t)\in A[0,t](i,j)} P_1(I(t))P_1(j,j) \tag{180}$$

This means, once $P_1^t(i,j) > 0$ then $P_1^{t+s}(i,j) > 0$ for all $s \geq 0$. By corollary7.2.1we conclude that P_1 is ergodic.

For the ergodicity of the product $P_1 * P_2$ first observe that

$$P_1 * P_2(u,v) = \sum_k P_1(u,k)P_2(k,v) \geq P_1(u,v) * P_2(v,v), \forall u,v \in Z \tag{181}$$

On the other hand we have

$$(P_1 * P_2)^t(I) = P_1 * P_2(i_0,i_1)P_1 * P_2(i_1,i_2)P_1 * P_2(i_2,i_3)...P_1 * P_2(i_{t-1},i_t). \tag{182}$$

Inserting the inequality 181 into implies 182 we obtain

$$(P_1 * P_2)^t(I) \geq P_1(I(t)) \prod_{i_k \in I(t)} P_2(i_k, i_k)\ \forall I(t) \in A[0,t](i,j). \tag{183}$$

Hence,

$$(P_1 * P_2)^t(i,j) = \sum_{I(t) \in A[0,t](i,j)} P_1 * P_2(I(t)) \geq \sum_{I(t) \in A[0,t](i,j)} P_1(I(t)) \prod_{i_k \in I(t)} P_2(i_k, i_k). \tag{184}$$

By assumption we have $P_2(u,u) > 0$ for all $u \in Z$, and hence $(P_1 * P_2)^t(i,j) > 0$ once $P_1^t(i,j) > 0$. On the other hand, from the proven ergodicity of P_1 we have $P_1^{t+s}(i,j) > 0,\ \forall s \geq 0$ once $P_1^t(i,j) > 0$. Hence, $(P_1 * P_2)^{t+s}(i,j) > 0,\ \forall s \geq 0$. We conclude that $P_1 * P_2$ is ergodic.

□

Definition 4

A Markov chain on a state space Z governed by the transition probability matrix P is reversible with respect to probability measure measure π on Z if the following detailed balance condition is satisfied.

$$\pi_i P(i,j) = \pi_j P(j,i) \tag{185}$$

or in compact form

$$D(\pi) * P = P^{tr} * D(\pi) \tag{186}$$

Which means that the matrix $D(\pi) * P$ is symmetric.

More intuitive version of reversibility is the Kolmogorov Criterion for reversibility which states that probability of any path with the same start and end state can be reverted, so that the forward and the backward predictions are indifferent.

Definition 5

Ergodic and reversible Markov Chain on a state space Z (which is finite in our case) will be called Random Walk on a Graph.

Lemma 7.4

Let let Z be state space and $G(Z) = (Z, E(Z), S(Z))$ be weighted and connected graph with respect to S. Moreover, let P transition probability matrix generated by (S, D) so that $P = D * S$ where D is the normalizing diagonal matrix as introduced in section 3.3 for modular scheme.

Then the Markov Chain governed by P is ergodic and reversible with respect to a unique stationary measure π.

Proof. The ergodicity of the P is the consequence of lemma 7.3 and implies existence of unique stationary π which we can write explicitly:

$$\pi_i = \frac{D(i,i)^{-1}}{\sum_{j\le m} D(j,j)} \tag{187}$$

Clearly π is probability measure on Z. In order to see that it stationary measure observe that

$$(\pi * D)_i = \frac{1}{\sum_{j\le m} D(j,j)} \ \forall i, \tag{188}$$

which means that $(\pi * D)_i$ is a constant independent of i

Therefore,

$$(\pi * P)_i = (\pi * D * S)_i = \frac{\sum_{j\le m} S(j,i)}{\sum_{j\le m} D(j,j)} = \frac{\sum_{j\le m} S(i,j)}{\sum_{j\le m} D(j,j)} \tag{189}$$

The first equality follows from the definition of P as in 3.3, second equality follows from 188 and the last equality is given by the symmetry of S.

But, by construction of diagonal matrix D we have (see section 3.3)

$$\sum_{j\le m} S(i,j) = D(i,i)^{-1}. \tag{190}$$

Hence, we have

$$(\pi * P)_i = \frac{\sum_j S(i,j)}{\sum_{j\le m} D(j,j)} = \frac{D(i,i)^{-1}}{\sum_{j\le m} D(j,j)} = \pi_i, \tag{191}$$

which means that π is a stationary measure.

Moreover,

$$D(\pi) * P = D(\pi) * D * S = \frac{1}{\sum_{j\le m} D(j,j)} S, \tag{192}$$

which is symmetric from which the reversibility follows.

□

If Markov Chain governed by transition probability measure P is ergodic and reversible with respect π, then it can be represented as product of diagonal matrix with a symmetric matrix. This fact is formulated by the following lemma.

Lemma 7.5

Let P be ergodic and reversible transition probability matrix on finite state space Z. Then there exist a diagonal matrix $\hat{D}$ and symmetric matrix $\hat{S}$ so that $P = \hat{D} * \hat{S}$.

Proof. Per definition we have $D(\pi)*P$ is symmetric. The lemma fallows directly by putting:

$$\hat{S} = D(\pi) * P \tag{193}$$

and

$$\hat{D} = D(\pi)^{-1} \tag{194}$$

□

As a summary, Random Walks on a Graph can be characterized with ergodic Markov Chains having the representation $P = D * S$. The Random Walks on Graph generated according to section 3.3 for modular scheme are special Random Walks with convenient properties such as positivity of its eigenvalues (see next subsection for this property) which allows following useful lemma).

Lemma 7.6

If the transition probability matrix P of Random Walks on Graph is induced according to subsection 3.3 for modular scheme with, then there is unique $P^{\frac{1}{n}}$ with positive eigenvalues, hence also positive according to our definition. Moreover, if the elements of $P^{\frac{1}{n}}$ are positive, then of $P^{\frac{1}{n}}$ represents a Random Walk on Graph.

Proof. Let $P = D^{-1} * S$ with a positive symmetric matrix S according to subsection 3.3 for modular scheme.

We can write $P^{\frac{1}{n}}$ explicitly as follows:

$$P^{\frac{1}{n}} = \theta * \hat{D}^{\frac{1}{n}} * \theta^{-1}. \tag{195}$$

Where

$$\theta = D^{-\frac{1}{2}} * \phi \tag{196}$$

Here, ϕ is unitary transformation diagonlizing the positive symmetric matrix $D^{-\frac{1}{2}} * S * D^{-\frac{1}{2}}$ with the diagonal matrix $\hat{D}$ so that:

$$\phi * \hat{D} * \phi^{-1} = D^{-\frac{1}{2}} * S * D^{-\frac{1}{2}} \tag{197}$$

We have

$$\theta\hat{D}\theta^{-1} = (D^{-\frac{1}{2}}*\phi)*\hat{D}*(\phi^{-1}*D^{\frac{1}{2}}) = D^{-\frac{1}{2}}*D^{-\frac{1}{2}}*S*D^{-\frac{1}{2}}*D^{\frac{1}{2}} = D^{-1}*S = P \tag{198}$$

This legitimizes the nth root in 195.

We have to show that $P^{\frac{1}{n}}$ can be represented as a product of diagonal matrix with a positive symmetric matrix.

$$P^{\frac{1}{n}} = D^{-\frac{1}{2}} * D^{\frac{1}{2}} * P^{\frac{1}{n}} = D^{-\frac{1}{2}} * D^{\frac{1}{2}} * (D^{-\frac{1}{2}} * \phi) * \hat{D}^{\frac{1}{n}} * (\phi^{-1} * D^{\frac{1}{2}}) \tag{199}$$

By the commutativity of diagonal matrices we can write

$$P^{\frac{1}{n}} = D^{-1} * D^{\frac{1}{2}} * (\phi * \hat{D}^{\frac{1}{n}} * \phi^{-1}) * D^{\frac{1}{2}} \tag{200}$$

It fallows that

$$P^{\frac{1}{n}} = D^{-1} * \hat{S}, \tag{201}$$

where

$$\hat{S} = D^{\frac{1}{2}} * (\phi * \hat{D}^{\frac{1}{n}} * \phi^{-1}) * D^{\frac{1}{2}}. \tag{202}$$

Since $\phi^{tr} = \phi^{-1}$, $\hat{S}$ must be symmetric with positive eigenvalues (see next subsection for positivity of eigenvalues).

Moreover, note that the matrices P and $P^{\frac{1}{n}}$ share the same right eigenvector $(1, 1, ...1)^{tr}$ with eigenvalue 1 (positive nth root of one is one). Applying this fact to $P^{\frac{1}{n}}$ implies that the sum of the rows are equal to one, which proves that $P^{\frac{1}{n}}$ is also a transition probability matrix if the elements of $P^{\frac{1}{n}}$ are positive.

□

As mentioned before, our main goal is to suggest an amalgamation of supplementary value chain information in a coherent manner with a desired properties. The fallowing lemma shows that if transition probabilities of Random Walks on Graph commute then our task becomes fairly easy.

Lemma 7.7

Let P_1 and P_2 be commutative transition probability matrices Random Walks on a Graph with the same state space Z so that

$$P_1 * P_2 = P_2 * P_1 \tag{203}$$

Then Markov Chain governed by

$$P = P_1 * P_2 \tag{204}$$

is also a Random Walk on a Graph.

Proof. Please first note that P_1 and P_2 share a common stationary measure π since by commutativity

$$P^n = P_1^n P_2^n = P_2^n P_1^n. \tag{205}$$

The above equation says that if π_1 respectively π_2 are stationary measures of P_1 and P_2 than as n tends infinity the distance $\mu * P^n$ to π_1 as well as π_2 tends to zero. Hence π_1 must be equal to π_2 for any initial distribution μ on Z.

Since P_1 is a random walk on graph with respect to π we can write

$$(D(\pi) * P_1 * P_2)^{tr} = P_2^{tr} * P_1^{tr} * D(\pi) = P_2^{tr} * D(\pi) * P_1 \tag{206}$$

On the other hand the commutatitivity of P_1 with P_2 implies the commutativity of P_1^{tr} with P_2^{tr} which implies

$$(D(\pi) * P_1 * P_2)^{tr} = P_1^{tr} * P_2^{tr} * D(\pi) = P_1^{tr} * D(\pi) * P_2 \tag{207}$$

But

$$(P_1^{tr} * D(\pi) * P_2)^{tr} = P_2^{tr} * D(\pi) * P_1 \tag{208}$$

Hence $D(\pi) * P_1 * P_2$ must be symmetric.

□

For simplicity assume that the algebraic multiplicity of eigenvalues of $\{\lambda_{1,i} : i \leq m\}$ and $\{\lambda_{1,i} : i \leq m\}$ are one and they are all strictly positive.

If $e_{2,i}$ are normalized eigenvector of P_1 with eigenvalue $\lambda_{2,i}$ then we have:

$$P_1 * P_2 * e_{2,i} = \lambda_{2,i} P_1 * e_{2,i}. \tag{209}$$

Using the commutativity we obtain

$$P_2 * P_1 * e_{2,i} = \lambda_{2,i} P_1 * e_{2,i}. \tag{210}$$

Hence the one dimensional invariant subspace of P_2 are also invariant under P_1 and vice versa.

We also have

$$e_{2,i} * P_2 * e_{2,i}^{tr} = \lambda_{2,i} > 0 \tag{211}$$

On the other hand by mentioned in-variance we must have

$$e_{2,i} * P_1 * e_{2,i}^{tr} = \lambda_{1,k} > 0 \tag{212}$$

for some $k \leq m$. This implies that besides sharing the same system of eigenvectors, they also share the same positive sign. Hence all the eigenvalues of $P_1 * P_2$ are positive. However,and some how disappointing is that, the set of commuting Random Walks of a fixed Random Walk is a poor set.

7.2 Positivity

For our interpretation of ECI (see lemma 4.1 and lemma 4.2 in subsection 4) the sign of the *second largest eigenvalue in absolute value*(which we will shortly phrase as "second largest eigenvalue") is crucial. The purpose of this section is among others to convince the reader that the sign of the second largest eigenvalue is positive. Since we want to compare different amalgamation method of supplementary value chain information matrices, we would like provide the reader also with some facts about the positivity of the products of Markov Chains.

There are different notions of positivity of a matrix. For our purpose we would like to be sure that the elements as well as the eigenvalues are positive. The former condition is needed in order to ensure the normalized version of S according to subsection 3.3 delivers transition probability matrix governing a

Random Walk on Graph and the later is needed for its asymptotic behavior. With a little abuse of terminology in our context, we will introduce following notion of positivity:

Definition 6

We will call mxm matrix A positive if

$$x^{tr} * A * x \geq 0 \ \forall x \in \mathbf{R^m} \tag{213}$$

It is easy to see that if a positive matrix has a diagonal form with real diagonal elements (similar to diagonal matrix) then all the eigenvalues $\lambda_i \ i \leq m$ are non-negative. Especially if A is a positive transition probability matrix we can write:

$$1 = \lambda_1 > \lambda_2 \geq \lambda_3 ... \geq \lambda_m \geq 0 \tag{214}$$

Products of positive matrices are not necessarily positive which can be shown by the following example:

Example 9

Put

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Then A cannot be diagonalized. But, A and A^2 are positive since

$$(x, y) * A * (x, y)^{tr} = x^2 + yx + y^2 \geq 0$$

and

$$(x, y) * A^2 * (x, y)^{tr} = x^2 + 2yx + y^2 \geq 0$$

But,

$$(x, y) * A^4 * (x, y)^{tr} = x^2 + 4yx + y^2$$

which is not positive.

However, in some cases the positivity is satisfied.

The next two lemmas are well known facts.

Lemma 7.8

Let A be a mxn matrix. Then $S = A * A^{tr}$ is a positive symmetric matrix and hence has positive eigenvalues. Moreover, if A is regular then all the eigenvalues are strictly positive.

Lemma 7.9

Let A be positive matrix with a diagonal form. Then there is unique positive nth root for every natural number n which can be explicitly written as:

$$A^{\frac{1}{n}} = \phi_A * (D_A)^{\frac{1}{n}} * \phi_A^{-1} \tag{215}$$

Where ϕ_A is the transformation matrix which diagonalizes A. Moreover, if A symmetric then ϕ_A is orthogonal symmetric transformation and hence $A^{\frac{1}{n}}$ is also symmetric.

A symmetric matrix is not necessarily positive even if all the elements are positive which the following examples shows.

Example 10
Put

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

A is not positive in our sense which can be checked by the test vector $x = (1, -1)$. Note that A is also not aperiodic (it has a period of 2).

Concerning the positivity, the situation does not change if we increase the weight of the diagonals of A a only bit in order to satisfy aperiodicity and hence ergodicity, which we would like to demonstrate with the following example.

Example 11

$$B = \begin{pmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{pmatrix}$$

Clearly, B is transition probability matrix. It has unique stationary measure $\pi = (0.5, 0.5)$ and it is ergodic since it satisfies the condition of corollary 7.2.1. It is also reversible which can be directly checked. Hence B represents a Random Walk. However since B has negative eigenvalue (because all the diagonals are dominated by the other elements in the same row), it cannot be represented by an intersection matrix according subsection 3.3 for modular scheme.

However, if the symmetric matrix is obtained according to modular scheme based on the intersection of information matrices, then all the eigenvalues are positive.

Lemma 7.10
Let M be zero one information matrix and D be a positive diagonal matrix with proper dimension. Moreover, let S be defined according to in subsection 3.3 for modular scheme:

$$S = M * D * M^{tr} \tag{216}$$

Then all the the eigenvalues of S are positive.

Proof. Put

$$A = M * D^{\frac{1}{2}} \tag{217}$$

Then we have

$$S = A * A^{tr} \tag{218}$$

Hence, by Lemma 7.8 S must have positive eigenvalues.

□

Lemma 7.11

Let D be positive diagonal matrix and S positive symmetric matrix then the matrices $X = D * S$ and $Y = S * D$ are also positive and have diagonal form.

Proof. We will only prove the statement of the lemma for$X$.

$$D^{-\frac{1}{2}} * X * D^{\frac{1}{2}} = D^{\frac{1}{2}} * S * D^{\frac{1}{2}} \tag{219}$$

$D^{\frac{1}{2}} * S * D^{\frac{1}{2}}$ is symmetric and positive which can be verified directly. Hence the matrix $D^{\frac{1}{2}} * S * D^{\frac{1}{2}}$ must have positive diagonal form. By similarity of X with $D^{\frac{1}{2}} * S * D^{\frac{1}{2}}$ we conclude that X and $D^{\frac{1}{2}} * S * D^{\frac{1}{2}}$ share the same roots which are positive and therefore X must positive eigenvalues.

□

Corollary 7.11.1

Let P transition probabilities of Markov Chains generated according to 3.3 for modular scheme:

$$P = D * S_1. \tag{220}$$

Then P is a transition probability matrix with positive eigenvalues.

Moreover, if we exclude the pathological cases by assuming the eigenvalues λ_i are all different then we can write

$$1 = \lambda_1 > \lambda_2 ... > \lambda_{m-1} > \lambda_m \geq 0 \tag{221}$$

The following two lemmas show that if we only require positivity instead of stronger assumption positive eigenvalues, then we can also weaken the assumptions.

Lemma 7.12

Let A and B mxm matrices and A positive. Then $X = B * A * B^{tr}$ is also positive.

Proof.

$$x * B * A * B * x^{tr} = y * A * y^{tr} \tag{222}$$

where $y = x * B$.

Hence, X is positive by positivity of A.

□

Lemma 7.13
Let A be positive with a diagonal form and S symmetric with strictly positive eigenvalues. Then the matrices $X = S * A$, $Y = A * S$ are also positive.

Proof. We will only prove that X is positive.

$$x * X * x^{tr} = x * S * A * x^{tr} = x * S^{\frac{1}{2}} * S^{\frac{1}{2}} * A * S^{-\frac{1}{2}} * S^{\frac{1}{2}} * x^{tr} \tag{223}$$

Put

$$y = x * S^{\frac{1}{2}} \tag{224}$$

Then we have

$$x * X * x^{tr} = y * S^{\frac{1}{2}} * A * S^{-\frac{1}{2}} * y^{tr} \tag{225}$$

By similarity A and $S^{\frac{1}{2}} * A * S^{-\frac{1}{2}}$ share the same roots. It follows that $S^{\frac{1}{2}} * A * S^{-\frac{1}{2}}$ positive. Hence, X is positive.

□

7.3 Amalgamation of Random Walks and Lie Trotter Approach

The behaivour of Random Walk on a Graph induced by an information matrix M, such as *diagonal* or *second eigenvector* of its transition probability matrix gives insight about how the countries evolve within a close system. But, coherency and meaningfulness of the key indicators of of a Random Walk on a Graph are reasoned by its properties such as *ergodicity*, *reversibility* and the *positivity* of the Random Walk on a Graph in our sense. Hence, we will pursue these properties in engineering of ranking system which amalgamates supplementary value chain information to ensure that these appealing properties are inherited. Moreover, an amalgamation should be unbiased if there is no justification that one system is dominated by an other system and reflect the organic charter of *complexity* or *sustainability.* For example if the transition probability matrices of export and import are *commutative* then our task becomes easy. Simple product of these matrices deliver directly a Random Walk on Graph satisfying desired properties. However in case of non-commutative transition probability matrices, the dynamic of the Markov Chain is biased and dominated by one of them depending on the order of matrix multiplication, which cannot be decided without ambiguity. Further drawback of this natural attempt is non-reversibility and non-positivity. In this section we will demonstrate that these challenges can be accomplished in a elegant manner. We first present a amalgamation method Random Walks on a Graph by Lie Trotter product of their transition probability matrices, which rationalizes the ambiguity of order away, inherits the ergodicity. However, our numerical results show that although deviation from reversibility and positivity of this method are fairly small, we cannot ignore it as numerical error. Luckily, we could discover a work around with which we can also ensure reversibility and positivity. This finding is presented at the end of this subsection is our main finding and which we think can be used in more general context.

Following theorem is a simple version of well known paper of Trotter for bounded generators of semi groups (see [7]).

Theorem 7.14

Let A_1 and A_2 two mxm matrices with well defined diagonal form, nth roots and logarithm. Then following limes exists

$$A_1 *_{LT} A_2 := \lim_{n\to\infty} (A_1^{\frac{1}{n}} * A_2^{\frac{1}{n}})^n. \tag{226}$$

The symbol $*_{LT}$ refers to Lie-Trotter Product.

Moreover, we have following properties of Lie-Trotter Product:

I) Commutativity:

$$A_1 *_{LT} A_2 = A_2 *_{LT} A_1 \tag{227}$$

II) Associativity: For mxm matrices A_1, A_2 and A_3 with well defined nth roots and logarithm we have:

$$(A_1 *_{LT} A_2) *_{LT} A_3 = A_1 *_{LT} (A_2 *_{LT} A_3) \tag{228}$$

We would like to note that in our case the Lie-Trotter Product can be written explicitly:

$$A_1 *_{LT} A_2 = \exp(\log(A1) + \log(A2)), \tag{229}$$

where

$$\log(A) = V^{-1} * \log(D) * V \tag{230}$$

Here V is the transformation matrix diagonalizing A :

$$A = V * D * V^{-1} \tag{231}$$

Note that if A is symmetric. then V can assumed to unitary transformation, so that $V^{-1} = V^{tr}$

It is wort to spend some time for the above theorem.

Firstly, the conditions of the above theorem are satisfied if all the eigenvalues are strictly positive which not a serious restriction in practice.

secondly, if A_1 and A_2 are commutative then "$*_{LT}$" coincides with the usual matrix product "$*$".

thirdly, If A_1 and A_2 are transition probability matrices then $A1 *_{LT} A_2$ is also a transition probability matrix if their nth roots have positive elements which is from practical point of view the case (see our comments after the proof of lemma 7.6). This follows from the fact that nth roots as well as the products of transition probability matrices are transition probability matrices which is also an inherited property at infinity. For the rest of this chapter we will assume this from practical point of view mild condition that the nth roots of the matrices which would like to amalgamate (P or S) have positive elements.

The question is, is there a sufficient condition so that "$*_{LT}$" product of random walks ergodic which we will answer by the following lemma.

Lemma 7.15

Let P_1 and P_2 be transition probability matrices of connecting random walks with strictly positive eigenvalues on the same state space Z so that, for every pair $(i, j) \in Z$ there exist a t_1 and t_2 with

$$P_1^{t_1}(i, j) > 0 \tag{232}$$

$$P_2^{t_2}(i, j) > 0 \tag{233}$$

Then $P_1 *_{LT} P_2$ is ergodic if P_1 and P_2 have a time indiscreet $RCLL$ ("right continuous with left limits") version.

Proof. Please note that we cannot use the lemma 7.3 for the ergodicity of product of Random Walks directly. We have to ensure that the the right hand side of inequality 184 in the proof of lemma 7.3 is bounded away from zero as n tends to infinity in definition of 226.

We will use the continuous time version of the time homogeneous Markov Chain governed transition probability matrix P so that:

$$P^s = \exp(s \log(P)) \tag{234}$$

which solves the following differential equation:

$$\frac{dU(s)}{ds} = \log(P) * U(s) \tag{235}$$

With initial condition

$$U(0) = \text{Id} = \lim_{s \to 0} P^s \tag{236}$$

The explicit representation of P^s according to 234 is justified since we assume that all the eigenvalues of P are strictly positive which implies the existence of real valued logarithm of the matrix P.

We have

$$\lim_{s \to 0} \frac{\text{Id} - P^s}{s} = \log(P) \tag{237}$$

Especially,

$$\lim_{s \to 0} \frac{\text{Id} - P^s}{s}(i, i) = \log(P)(i, i) \tag{238}$$

The idea is to look at the continuous version of the Markov Chain governed by P_1 and and P_2 in discrete time intervals. For fixed natural numbers n and t let

$$H_{tn} = \{l + \frac{k}{n} : \ 0 \leq k \leq n, \ l \leq t-1\} \tag{239}$$

equidistant points of interval $[0, t]$.

For each $i \in Z$ the probability $L^{\delta}(i) = P(X_s = i, \ \forall s \in [0, \delta) \mid X_0 = i)$ of lingering at i in an infinite decimal interval $[0, \delta]$ can be approximated by

$$Ł^{\delta}(i) \approx 1 + \delta \log(P_2)(i, i). \tag{240}$$

Here, $(X_s)_{s \in [0,\infty)}$ is a $RCLL$ ("right continuous with left limits"; see e.g. page 4 and 5 of [11]) version of Markov Chain governed by P. Using the above approximation and the time homogeneity property, for each $i \in Z$, we can approximate the lingering probability $L^t(i)$ of the above Markov Chain at i up to t given that it starts from i at initial time 0:

$$L^t(i) \approx (P^{\frac{1}{n}})(i, i))^n)^t \approx (1 + \frac{1}{n} \log(P)(i, i))^{nt} =: L^t_n(i). \tag{241}$$

We have

$$\lim_{n \to \infty} L^t_n(i) = \exp(t \log(P)(i, i)). \tag{242}$$

Hence, for each path $(i_{t_k})_{t_k \in H_{tn}}$ we obtain

$$\prod_{t_k \in H_{tn}} P_2^{\frac{1}{n}}(i_{t_k}, i_{t_k}) \geq (((P_2^{\frac{1}{n}})(u, u))^n)^t. \tag{243}$$

Where u minimizing state of the function $P_2^{\frac{1}{n}}(u, u)$:

$$P_2^{\frac{1}{n}}(u, u) = \min\{P_2^{\frac{1}{n}}(i, i) : \ i \in Z\} = \min\{\exp(\frac{1}{n} \log P)(i, i) : \ i \in Z\} \tag{244}$$

Note that since Z is finite state space, for large enough n, the above minimizing state u dose not depend on n.

Put

$$\alpha = \min\{\lim_{s \to 0} \frac{1 - P_2^s(ii)}{s} : \ i \in Z\} = \min\{\log(P_2)(i, i) : \ i \in Z\} \tag{245}$$

By assumption we have $P_2(i, i) > 0, \ \forall i \in Z$. Hence, we must have $\alpha > -\infty$ which implies that $\exp(\alpha) > 0$.

Letting $t \to \infty$ we obtain

$$\lim_{n \to \infty} \prod_{t_k \in H_{tn}} P_2^{\frac{1}{n}}(i_{t_k}, i_{t_k}) \geq \lim_{n \to \infty} (((P_2^{\frac{1}{n}})(u, u))^n)^t = \exp(t\alpha) > 0 \tag{246}$$

Now we can use the same idea as in lemma 7.3 and obtain

$$((P_1^{\frac{1}{n}} * P_2^{\frac{1}{n}})^n)^t(i,j) = \sum_{I(t)\in A[0,nt](i,j)} P_1^{\frac{1}{n}} * P_2^{\frac{1}{n}}(I(t)) \geq P_1^t(i,j)\exp(t\alpha) > 0. \tag{247}$$

The right hand side of above inequality does not depend on n. Hence, once $P_1^{t_0}(i,j) > 0$ we conclude $(P_1 *_{LT} P_2(i,j))^t\ \forall t \geq t_0$. Which implies that $P_1 *_{LT} P_2$ is ergodic.

□

Although we can only claim that the log of transition probabilities Random Walks represent intensity matrices (which is necessary for having a time indiscreet version) only approximately and not rigorously, our numerical results show that the deviance from reversibility and positivity is insignificant. However they cannot be argued away as numerical error.

But by the weighting matrices S_1 and S_2 fallowed by normalizing ensures all the desired properties if the positivity and connectivity of the elements amalgamated S satisfied which is our key finding. Note that non-positivity and non-connectivity of the elements amalgamated S is rare and numerically it becomes marginal as the size of information tends to infinity.

Theorem 7.16

Let P_1 resp P_2 two mxm transition probability matrices driven by modular scheme from zero one information matrices M_1 and M_2 respectively as in 3.3, so that

$$P_1 = \hat{D}_1 * S_1 \tag{248}$$

$$P_2 = \hat{D}_2 * S_2 \tag{249}$$

where

$$S_1 = M_1 * D * M_1^{tr} \tag{250}$$

and

$$S_2 = M_2 * D * M_2^{tr}, \tag{251}$$

with $\hat{D}_1$ resp. $\hat{D}_2$ normalizing diagonal matrices of S_1 resp. S_2.

Further, let S denote the Lie Trotter Product of S_1 with S_2 :

$$S = S_1 *_{LT} S_2. \tag{252}$$

Then S is symmetric and have positive eigenvalues. Moreover, if as is connective and the elements of S are positive then the P induced by S according subsection 3.3 is ergodic, reversible, hence Random Walk on a Graph.

Proof. For the symmetry and positivity of the eigenvalues of S note that $\log(S_1)$ and $\log(S_2)$ are symmetric by the definition of log, which can be directly checked by the formula 231. Since sum of symmetric matrices are symmetric and symmetry preserving property of exp we conclude that $S = S_1 *_{LT} S_2 = \exp(\log(S_1) + \log(S_2))$ is symmetric and has a symmetric square root which can be expressed as:

$$S^{\frac{1}{2}} = \exp(\frac{1}{2}(\log(S_1) + \log(S_2)). \tag{253}$$

But

$$S = S^{\frac{1}{2}} * S^{\frac{1}{2}} = S^{\frac{1}{2}} * (S^{\frac{1}{2}})^{tr} \tag{254}$$

Hence by lemma 7.8 it fallows that the amalgamated S is symmetric matrix with positive eigenvalues by 7.4 and corollary 7.11.1.

□

References

[1] Trotter, Hale F. "On the product of semi-groups of operators." Proceedings of the American Mathematical Society 10.4 (1959): 545-551.

[2] Chernoff, Paul R." Product formulas, nonlinear semigroups, and addition of unbounded operators": 162-177". Vol. 140. American Mathematical Soc., 1974.

[3] Hidalgo, César A., and Ricardo Hausmann. "The building blocks of economic complexity." Proceedings of the national academy of sciences 106.26 (2009): 10570-10575.

[4] Hausmann, Ricardo, et al. "The atlas of economic complexity: Mapping paths to prosperity." Mit Press, 2014.

[5] Deardorff, A. V. The general validity of the Heckscher-Ohlin theorem. AER 1982, 72(4), 683-694.

[6] Albeaik, S., Kaltenberg, M., Alsaleh, M., Hidalgo, C. A. (2017). Improving the economic complexity index. arXiv preprint arXiv:1707.05826

[7] Albeaik, S., Kaltenberg, M., Alsaleh, M., Hidalgo, C. A. 729 new measures of economic complexity (Addendum to Improving the Economic Complexity Index). arXiv 2017. arXiv preprint arXiv:1708.04107

[8] Buccellato, T. The Competences of Firms are the Backbone of Economic Complexity. SSRN 2016 (August 22). Available online: http://dx.doi.org/10.2139/ssrn.2827468 (accessed on 29 November 2020).

[9] Reynolds, C.; Agrawal, M.; Lee, I.; Zhan, C.; Li, J.; Taylor, P.; Mares, T.; Morison, J.; Angelakis, N.; Roos, G. A sub-national economic complexity analysis of Australia's states and territories. Regional Studies 2018, 52(5), 715-726.

[11] Brummitt, C. D.; Huremović, K.; Pin, P.; Bonds, M. H.; Vega-Redondo, F. Contagious disruptions and complexity traps in economic development. Nature Human Behaviour 2017, 1(9), 665-672. https://doi.org/10.1038/s41562-017-0190-6.

[12] Olson, M. (2012). The Logic of Collective Action [1965]. Contemporary Sociological Theory, 124.

[13] Schelling, T. C. (1958). The strategy of conflict. Prospectus for a reorientation of game theory. Journal of Conflict Resolution, 2(3), 203-264.

[14] Schelling, T. C. (2006). Strategies of commitment and other essays. Harvard University Press.

[15] Caldarelli, G., Cristelli, M., Gabrielli, A., Pietronero, L., Scala, A., Tacchella, A. (2012). A network analysis of countries' export flows: firm grounds for the building blocks of the economy. PloS one, 7(10), e47278.

[16] Ryser, H. J. "Set intersection matrices." Journal of Combinatorial Theory, Series A 32.2 (1982): 162-177.

[17] Artzner, Philippe, et al. "Coherent multiperiod risk adjusted values and Bellman's principle." Annals of Operations Research 152.1 (2007): 5-22.

[18] Fiedler, Miroslav. "Algebraic connectivity of graphs." Czechoslovak mathematical journal 23.2 (1973): 298-305.

[19] Mealy, P., Farmer, J. D., and Teytelboym, A. (2018). A new interpretation of the economic complexity index. Alexander, A New Interpretation of the Economic Complexity Index (February 4, 2018).

[20] Wilmer, E. L., David A. Levin, and Yuval Peres. "Markov chains and mixing times." American Mathematical Soc., Providence (2009).

[21] Shannon, C. E. (1948). A mathematical theory of communication. The Bell system technical journal, 27(3), 379-423.

[22] Shen, A., Uspensky, V. A., Vereshchagin, N. (2017). Kolmogorov complexity and algorithmic randomness (Vol. 220). American Mathematical Soc..

[23] Haber,H. E. (2018). Notes on the matrix exponential and logarithm. URL http://scipp. ucsc. edu/ haber/webpage/MatrixExpLog. pdf.