

A Simple Model of Global Value Chains

Sourish Dutta

1 Background (Taglioni & Winkler 2016)

Nowadays most of the production processes are composed of different tasks performed by several geographically separated economic agents with its profound impact similar to what observed regarding the division of labour. This system of global production sharing, which is known as Global Value Chains (GVCs), is eventually forming an extensive network of economic values or value added that explains the changing nature of international trade & development. In this context, the gross nature of imports and exports data is such that it cannot provide any direct information about the foreign and domestic value added to an economy. Analysis of value-added trade requires much more sophisticated data like TiVA and WIOD. Nonetheless, we can get some indirect indications about the GVC phenomenon through a set of analytic refinements based on gross trade data and its informed classifications.

The first consideration when investigating a country's potential in GVCs is what the country exports and imports. Looking at the top 50 to 100 export and import products of a country, classified at the most disaggregated level (at least Harmonized System [HS] 6-digit or Standard International Trade Classification 5-digit products) is a good starting point. For most countries, the top 50 exports and imports are likely to cover at least 50 percent of the total trade value in each direction, and the top 100 cover at least 75 percent.

A country's distribution of exports tends to follow a lognormal, power, or Zipf's law distribution. Zipf's law, originally applied to language, states that given some universe of items, the frequency of any item is inversely proportional to its rank in the frequency table. That is, a few items account for the bulk of the given universe; the contribution of most items is marginal. Exports and imports loosely follow such asymmetric distribution laws. Therefore, the marginal additional information that can be gathered from import and export products beyond the top 50 to 100 is generally small.

1.1 Objectives

The whole purpose of this part of my research is to quantify the degree of value transformation taking place domestically between imports and exports due to an economy's GVC integration in the manufacturing sector. The consideration of using gross imports & exports data to draw a quantitative inference about an economy's GVC potential leads to a different conceptualisation of the existing production process regarding knowledge, learning, and upgrading. Traditional trade theories assume that the whole production process of a product is taken place in one country and marketed in another. However, the notion of GVC trade is different – quantifying how much of an economy's export value is contributed by imports and how much it is adding value to its exports. The basic concept is "import to export" or I2E. First, we have to look at what an economy imports and exports. Looking at the top imported and exported products of the economy, classified at the most disaggregated level (Harmonised System [HS] 6-digits) is a good starting point. Second, we have to compare statistically those product-level imports with exports regarding values, volumes, and unit prices. If exports and imports follow a similar distribution or having a short distance between two distributions, this may imply that relatively little transformation is taking place in the domestic segment of the economy's manufacturing GVCs.

The logic behind these objectives is that we can easily point out the core factor of global inequality, i.e. an asymmetric distribution of knowledge. In fact, due to Paul Romer's paper on endogenous technological change in 1990, economists finally figured out that knowledge, not capital, not technology by itself, not labour or land, was the ultimate factor of production that allowed economic growth. Therefore, understanding the structure & dynamics of knowledge distribution is crucial in the age of global production networks. In this context, the processes of learning and upgrading (innovation) are the fundamental forces of diffusing productive knowledge as a means of generating economic values among the different economic actors in the various production activities. Many authors emphasise the importance of understanding knowledge (local & global) creation and utilisation in the process of innovation for value addition. They indeed identified information as the medium to move knowledge from one place to another, and its interpretation (decoding) depends upon the degree of user's (individual or organisation) set of capabilities. That is why there is a high chance of uneven distribution of knowledge in a few pockets of the world economy.

Here the fundamental idea is that more emphasis should be given to the study of statistical distributions than to other quantities. An analysis through statistical approach is unavoidable since production networks are extensive systems, whose collective behaviour cannot be made out of the

elementary features. However, it would be logical to think the flow of value in global value chains through information (medium) from knowledge as input to learning as the process to upgrading as output (innovation) in the system of production sharing. Therefore, concentration or diffusion of the economy's economic value (stored information) among different products can be interpreted (decoded) with the analysis of structure and dynamics of value distribution between exports and imports.

2 Conceptual Framework (Hausmann et al. 2014)

We can draw an analogy between the production process and the game of Scrabble. In Scrabble, players utilise plenty number of tiles carrying single letters of the alphabet. Players then assemble their letters in order to make words. The value of the words is calculated by the sum of the value of each letter, with rarer letters (e.g. Q, X, Z) having a higher value than common letters (e.g. A, E, S, T). Players then play to produce the most valuable words out of the letters available to them. In an economy, firms attempt to make products similar to how Scrabble players make words. To make a product, firms must bring together a variety of very different and very specific inputs and activities. Each of these inputs can be thought of as one capability needed for production, just as a letter in Scrabble represents one capability needed to make a word possible. The list below summarises some economic phenomena consistent with the metaphor.

Scrabble Game

- The more letters available, the more words players are able to make.
- The rarer the letters and the longer the word, the more valuable it is.
- Which words a player can spell in the future is determined by the letters currently available.
The value of adding any particular letter is determined by the set of existing letters.

Economic Activity

- Economies tend to diversify, rather than specialise, as the number of available capabilities grows.
- Economies that are able to utilise rare capabilities or combine many capabilities to make complex products tend to be richer and/or grow more quickly.

- Economies tend to find new products to produce that leverage their existing capabilities. The new industries that are most likely for an economy to discover are influenced by the industries currently existing there.

The fundamental notion of the scrabble theory of economic development is that economic development is the fruit of shared (collective) learning in an economy. This concept consists of three properties of knowledge and learning: limited capacity to reason i.e. bounded rationality, a limited capacity to learn individually, a limited capacity to transmit experiential knowledge. These constraints are the principal drivers of team formation among individuals to unite their knowledge. It is the basis of economic complexity, where each production process is a combination of numerous complementary capabilities. Economies that are participating in various production networks accumulate adequate capabilities and can combine with productivity.

The idea of limited learnability has the inference of why sometimes knowledge does not move freely and concentrate in few pockets. Therefore, it is hard for individuals to learn, collectives (firms, industries, countries) on the other hand can learn by gathering people. Hence, the diffusion of knowledge does not move necessarily because of ideas flowing into people heads, but because of people with ideas assemble in certain places. Capabilities embodied in the brain of people will move where they can combine with other capabilities productively. This conceptualisation thus indicates why knowledge accumulates slowly, and why its distribution is uneven across spaces.

3 Model of Economy's Collective Learning (Gomez-Lievano 2018)

Let an economy's economic activities be indexed by p (i.e. industry specific output or product p) and production units (i.e. firms) by f . Here our objective is to find out the expression of probability that a production unit or firm f is capable to produce product in the economy: $Pr(X_{f,p} = 1)$, where $X_{f,p}$ is a variable that takes value 1 if the firm f is capable to produce p successfully in the economy, and 0 if not.

It is natural to say that a firm f will be cable to produce a industry-specific p , depending upon whether it is capable to assemble the appropriate knowledge and capabilities. Assume this refers combining C_p different and complementary capabilities. It is impossible to specify these capabilities in general. The crucial part is that capabilities are all the ingredients needed to produce the product p . These include, in principle, design, administrative services, manufacturing, assembling, marketing, distribution and customer services. Thus it would be reasonable to think of capabilities

as professional or job occupations, although they can also include public services that a production process may need as a necessary requirement.

The parameter C_p represents, in this context, the complexity (degree of production) of the economic activity associated with the production of p . The more capabilities are needed, the larger the value of C_p , and the more complex the activity. Notice that this approach differs from the conventional production process assumed in economics, whereby the emphasis is on the substitutability of a few production factors (i.e. capital and labour). Instead, we are assuming that (i) there is no substitutability between capabilities, and that (ii) the number of factors is not two, but $C_p \gg 1$. The reason we need to think probabilistically in this model comes from this assumption about the large multiplicity of capabilities.

Let c_f be the probability that the firm f has any random capability of the C_p capabilities required for the p production. In other words, the number of capabilities the firm is expected to have is, on average, $c_f C_p$. This probability can be interpreted as a measure of the firm's knowledge. For example, the firm will manufacture $p =$ a machine, which let us assume requires $C_p = 10$ different capabilities, and c_f represents the fact that it has the capacity to do easily designing and manufacturing, so $c_f = 2/10 = 0.2$. The larger the parameter c_f is, the better equipped is the firm in producing the product p individually, the less it needs a group of firms supporting it as a consequence, and less dependent it will be of the country it is in and on what the country offers to it. Notice, however, that while c_f can be interpreted as how many different things it could know how to do individually. The probability that the firm f will be capable to produce successfully by itself is $c_f^{C_p}$. Since c_f is a number between 0 and 1, the more complex the economic activity, the probability it will be successful will decrease exponentially. But how does the country change the probability of the firm to be capable to run its production, $Pr(X_{f,p} = 1)$?

Of the C_p capabilities required to produce product p , suppose the economy provides C capabilities to the firm f (where $0 \leq C \leq C_p$). In other words, through general public and private services, which it is typically exposed to on a regular basis by locating in the economy, the firm f could in principle be capable to get and complete the remaining 8/10's missing capabilities outside its expertise, which it expects to require to run. Presumably, the larger the country, the more diverse, and the larger C will be, and the easier it will be to get those capabilities.

The problem, note, is that the firm will only get all the C_p capabilities it needs if the C capabilities offered by the economy contain the capabilities that the firm f does not have given c_f . The only situation in which the firm will be capable to run its production is if it requires none of the capabili-

ties the economy does not have. Locating in the economy, the firm f can be sure it has C of the C_p capabilities. These are a given, in a sense, and it does not need to worry about them. The challenge it faces is rather with getting the $C_p - C$ capabilities not offered by the economy, which it cannot take for granted. These are capabilities that it will need to bring to the production process, without the help of the economy. Necessarily, solving the challenge of lacking $C_p - C$ capabilities will depend on its knowledge. The firm f has a probability c_f of having any of these capabilities.

Let us compute the probability that it will be capable to get them all and operate its production process $Pr(\text{firm } f \text{ in the economy produces } p \mid \text{economy facilitates } C \text{ of the } C_p \text{ capabilities})$, which we can write more concisely as $Pr(X_{f,p} = 1 \mid C)$. According to reasoning above, this probability is equal to

$$Pr(X_{f,p} = 1 \mid C) = c_f^{C_p - C} \quad (1)$$

Equation (1) is the product $c_f c_f \dots c_f$ because it is the probability of having the first capability times the probability of having the second, and so on, until we have the probability of getting each of the missing capabilities not offered in the economy.

In reality, however, C is not a fixed number. They are changing over the period, and no one knows the economy as a whole completely. Hence., if the firm f is very unlucky it may get $C = 0$, or it can be super lucky and get $C = C_p$. To take this stochasticity into account, we can think instead of the probability that the economy provides any of the capabilities. Let us denote this probability by c . The expected number of capabilities required to produce p that the economy can offer on average is $E(C) = cC_p$. Thinking of C probabilistically, means thinking of C in this context as a binomially distributed random variable with parameters C_p and c .

To correctly compute the probability that the firm f will be capable to manage its production process we need to average equation (1) over all the possible number of capabilities the economy may offer:

$$\begin{aligned} Pr(X_{f,p} = 1) &= \sum_{C=0}^{C_p} Pr(X_{f,p} = 1 \mid C) Pr(C) \\ &= \sum_{C=0}^{C_p} c_f^{C_p - C} \binom{C_p}{C} c^C (1-c)^{C_p - C} \\ &= \sum_{C=0}^{C_p} c^C (c_f (1-c))^{C_p - C} \\ &= (c + c_f (1-c))^{C_p} \\ &= (1 - (1 - c_f)(1 - c))^{C_p} \end{aligned} \quad (2)$$

For clarity, let us denote this probability as an explicit function of the parameters involved, $Pr(X_{f,p} =$

1) $\equiv \phi(C_p, C_f, c)$. We can generalise this model and imagine that the country has a "field" spread in the xy -coordinates, $c(x,y)$. This field is an abstraction of the economy's milieu, it represents the probability that the economy provides one of the ingredients for productions to occur, and we can assume that people interact with it as they live and work in the economy. It should capture the elements from all the types of interactions to which people are exposed (both the social and built environment). In this view, the economy functions as though it is permeated across space by a "cultural field", and $c(x,y)$ quantifies the magnitude of the social, economic, and cultural repertoire available at a particular location. When the value of the field is high, it means that this location in the economy has a high concentration of "diverse economic factors". A high value of $c(x,y)$ will therefore increase the probability that an individual firm will find the right elements to engage in a given economy's phenomenon (e.g., to be successful in producing a product). Equation (2) assumes firms interact with the average intensity of the field, $c \equiv \langle c(x,y) \rangle$, where $\langle . \rangle$ is a spatial average. The mean field approach allows us to isolate a core mechanism in the model: that economic phenomena occur because firms are able to "execute" a recipe (e.g., a production process) if the environment is favourable, that is, if the economy complements the individual firm. How complex a given production process is, how capable is an individual firm, and how suitable is the country for executing the production process are the three fundamental quantities that determine the overall statistics of production sharing system, as well as other measures of economic output.

Using the properties of exponential and logarithms, equation (2) can be simplified to yield the following expression:

$$\phi(C_p, c_f, c) \approx e^{-C_p(1-c_f)(1-c)} \quad (3)$$

The exponent is the product of three quantities, respectively associated with the production, the firm, and the economy. These are the drivers of production sharing system in the economy, which is why we refer to them as the drivers of economy's economic complexity. Thus, the idea of the field c is therefore to quantify what really matters about the context in which firms are embedded. Some contexts are more complementary than others to some firms, which is to say, some contexts are more diverse than others. Our claim that the drivers behind the prevalence of an economic phenomenon in the economy are the firms, the economy, and the phenomenon itself may seem trivial. However, we are arguing that one can quantify how the production process, the individual firm and the country each contribute to the occurrence of the economic phenomenon, and we specify the mathematical way in which they interact. Thus, we are proposing an integrated understanding of economic phenomena that improves on network-based models because we account not only for

the properties of an economy (i.e., the network), but also for the characteristics of the phenomenon and the attributes of firms (i.e., the nodes in the network).

3.1 Three Drivers of Economy's Collective Learning

What is equation (3) telling us? Two aspects stand out. First, the exponent in equation (3) is the negative of a log-probability and, as a consequence, it has units of entropy (i.e., "nats" or "bits"). Given that entropy quantifies the information required to describe a stochastic system, it is reasonable to refer to this (negative) exponent as a "net complexity". For us, net complexity captures the complementarity between the firm, the economy and the activity. With this terminology, we say that production rates are determined by this complementarity. We claim, however, that net complexity has consequences beyond just terminology.

The second observation about equation (3) is that net complexity is decomposable into three factors: the drivers of economic complexity. Let us recall the meaning of the terms in equation (3):

1. C_p : This is the number of capabilities required to produce the industry-specific product p . We can refer to it as the complexity of the product.
2. $1 - c_f$: This is the firm-specific probability of lacking any of the capabilities required in production processes. Conversely, c_f can be referred to as a measure of individual knowledge.
3. $1 - c$: This is the economy-specific probability of lacking any of the capabilities required in production processes. Conversely, c can be referred to as a measure of collective knowledge and represents a measure of input availability, which in turn represents a measure of the economy's diversity.

Changes in these terms change the probability of production. The fact that net complexity is decomposable leads to a strategy for estimating its components. This is because some industries involve production processes that are inherently more difficult to execute than others, some countries foster industrial activities less than others, and some firms are less capable of contributing to industrial processes than others. They are, a priori, independent quantities, but we expect them to be correlated in reality, as firms with complex production processes are likely to choose to locate in large countries, which are the places where high skill firms sort themselves into. Our model explicitly shows how these three drivers of net complexity, once given, interact and determine the statistical properties of production.

3.2 Impacts of Drivers on Economic Performance

Not surprisingly, increasing any of these three terms will decrease the probability that firm f will produce p . But the crucial observation is that they involve exponential changes. That is to say, small changes in any of these three terms can in principle have (exponentially) large effects on the success of firms' production processes, especially if the value of those variables is already high. But let's study the partial rates of change separately, in order to compare them:

Technological improvement of production process of p :

$$\frac{\frac{\partial \phi(C_p, c_f, c)}{\partial (-C_p)}}{\phi(C_p, c_f, c)} = (1 - c_f)(1 - c) \quad (4)$$

Individual learning for firm f :

$$\frac{\frac{\partial \phi(C_p, c_f, c)}{\partial (C_p c_f)}}{\phi(C_p, c_f, c)} = (1 - c) \quad (5)$$

Collective learning for the economy:

$$\frac{\frac{\partial \phi(C_p, c_f, c)}{\partial (C_p c)}}{\phi(C_p, c_f, c)} = (1 - c_f) \quad (6)$$

The partial derivatives have the term C_p because we want them to reflect changes in the number of capabilities. Hence, $\partial(-C_p)$ represents the reduction of the number of capabilities required to produce p , $\partial(C_p c_f)$ represents the increase in the average number of capabilities known by the individual firm f , $\partial(C_p c)$ represents the increase in the average number of the capabilities present in the economy. In other words, the probability the firm f will be successful to produce p , $\phi(C_p, c_f, c)$, will increase according to equation (4) if the complexity of product p decreases (through technological improvements), equation (5) tells us that it will increase if the individual firm f 's knowledge increases (individual learning), and equation (6) that it will increase if the collective knowledge of an economy increases (collective learning).

Let us study in detail the magnitude of these rates of change. On the one hand, we have C_p which is supposed to be large, $C_p \gg 1$. On the other hand, c_f and c are probabilities and are therefore between 0 and 1. However, since an economy is a collective, the probability it provides an input is larger than the probability an individual has it, so $c \gg c_f$. Conversely, $1 - c \ll 1 - c_f$. Consequently, we have that $0 < (1 - c)(1 - c_f) < (1 - c) \ll (1 - c_f)$. The implication is that these rates have the following order:

$$0 < \frac{\frac{\partial \phi(C_p, c_f, c)}{\partial (-C_p)}}{\phi(C_p, c_f, c)} < \frac{\frac{\partial \phi(C_p, c_f, c)}{\partial (C_p c_f)}}{\phi(C_p, c_f, c)} \ll \frac{\frac{\partial \phi(C_p, c_f, c)}{\partial (C_p c)}}{\phi(C_p, c_f, c)} \quad (7)$$

Thus, the effect of a technology improvement is smaller than the effect of individual learning which is much smaller than the effect of collective learning.

Effect of technological improvement < Effect of individual learning « Effect of collective learning.

Increasing the collective knowledge of an economy has a significant effect on the probability of firm f being capable to run production process which produces product p . Presently, we lack a detailed theory about the dynamical laws of these drivers and how they relate to one another. Intuitively, thus, it is easy to see that increases in collective knowledge have a reinforcing effect and suggest a virtuous cycle: a place with a relative large body of collective knowledge will attract more people and facilitate more complex economic activities, which themselves will increase the body of collective knowledge in that place. This process will thus propel a run-away cycle of collective learning that will concentrate economic activities and wealth in a relatively compact region of space: a nation. The more complex the activities, the more concentrated they will be across places. This explains why complex innovations tend to happen in large diverse nations.

Of course, the comparison in equation (7) has several problems and hinges on highly simplifying assumptions. For example, the comparison assumes that a linear (infinitesimal) change in the three variables is comparable among them. In other words, it does not take into account the cost of these changes. But one can play a bit with the equations, make some assumptions, and it is easy to see that this result holds for a wide range of situations. A second intuition is that the effects of collective learning is differentially distributed across people and economic activities. The combined effect can be summarised by noticing that countries with a large body of collective knowledge will make difficult activities easier. And the difficulty can be because the activity is itself very complex, or because the firm lacks several capabilities, or both. Equation (6) shows these two effects clearly given it is a function of both C_p and c_f . Nevertheless, with these comparisons one can hypothesise which of the three drivers has the largest impact on economic performance.

Are the drivers of economic complexity proposed by this model tangible? Their intangibility may arise from the unspecified nature of a "capability", and the fact that we have not defined a method to measure the size of sets of capabilities.

3.3 The Distributional Implications

Our model assumes that a large set of inputs must be combined to generate an output. The output will not be produced, however, if any of the inputs is missing. This is a specific form of a production function called the Leontief production function. By using a Leontief function we are assuming that the presence of an output follows a logic of complementarity. Mathematically, complementarity meant taking products of probabilities. Thus, it can also be said that outputs follow a logic of multiplicativity. This logic allowed us to calculate the consequences of our model, summarized in equation (3).

We can claim that the logic of multiplicativity between inputs for determining the presence of an output, in turn, implies that the presence of an input follows a logic of additivity between outputs. To see this, let us put both logics in words. On the one hand, we say that to observe a unit of output of a given product X we need input A and B and input C , and and so on, such that we list all the inputs required to produce X . On the other hand, this allows us to say that if we observe the presence of input A it is because output X is present or output Y is present or output Z is present, and so on, as we list all the products that require input A as a necessary factor of production. Thus, concatenating conditions with and's is akin to multiplying probabilities while concatenating conditions with or's is akin to adding probabilities. Consequently, outputs follow a logic of multiplicative processes while inputs follow a logic of additive processes.

3.4 Measuring the Drivers of Economic Complexity

Taking negative logarithms twice in (3) yields

$$-\ln(-\ln(\phi(C_p, c_f, c))) = -\ln(C_p) - \ln(1 - c_f) - \ln(1 - c) \quad (8)$$

The net complexity is decomposed linearly in this way into its main components. If one has the means to estimate the probability on the left-hand side, $\widehat{\phi}(C_p, c_f, c)$, then one can regress this "double-logged-likelihood" against three additive fixed-level effects from the production, the firm, and the economy,

$$-\ln(-\ln(\widehat{\phi}(C_p, c_f, c))) = \alpha_p + \beta_f + \gamma + \varepsilon_{p,f} \quad (9)$$

where $\varepsilon_{p,f}$ is the error one would try to minimise in the estimation of the fixed effects.

4 Empirical Underpinning

The empirical analysis is based on collective characterisation & quantification of production sharing system. This principle comprises of

1. Structure of value distribution (System's Composition): The aim of this study is to discover law-governed relations between the products forming this production sharing system. In the process of this research we will be identifying the structures peculiar to this sharing system.
2. Dynamics of value distribution (System's Functioning): But structure is not enough to make a system. This production sharing system consists of something more than structure: it is a structure with certain properties and stability (including their movement in time, their sequence etc).

When the production sharing structure is understood from the standpoint of its properties, it is understood as a system.

References

- Gomez-Lievano, A. (2018), 'Methods and concepts in economic complexity', *arXiv preprint arXiv:1809.10781* .
- Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M. & Simoes, A. (2014), *The atlas of economic complexity: Mapping paths to prosperity*, MIT Press.
- Taglioni, D. & Winkler, D. (2016), *Making global value chains work for development*, World Bank Publications.