ARMG PUBLISHING "Thinking ahead" sciendo
Financial Markets, Institutions and Risks, Volume 7, Issue 1, 2023
ISSN (online) – 2521-1242 ISSN (print) – 2521-1250

The Synergic Entropy

An efficient frontier output derived from merged input units boosted by synergy and constrained by critical input

http://doi.org/10.21272/fmir.7(1).39-70.2023

Henrique De Carvalho Videira https://orcid.org/0000-0001-9362-2244

Federal University of Rio de Janeiro, Central Bank of Brazil, Centro, Rio de Janeiro-RJ, Brazil

Email: henriquevideira@live.com, henrique.carvalho@coppead.ufrj.br

Abstract. *The theory equates the maximum output deviations (efficient frontier) caused by combined inputs with affinity-synergy in a system, which leads to a parametric volatility whose curve can be compared to data envelopment analysis (DEA). The input is a cumulative variable (e.g.: merged assets), and the output is a flow variable (e.g.: combined incomes). Rather than being purely stochastic, volatility is estimated by a novel parameter for risk named synergy, which is constrained by critical input (scarce resources). The output acceleration derived from the mergers among inputs, boosted by synergy, is the main foundation of the approach, which particular case gives Shannon and Boltzmann-Gibbs entropies. Tests are done in the 11 USA Sectors over their quarterly financial statements, proving that synergy is significant for financial statements, whereas typical betas only present significance in stock market data. A practical application is a novel discount rate for valuation using synergy, whose results for each sector are stable and coherent with perceived risk. Systems that rely on causal relations between output and multiple inputs can be regressed under novel parameters, rather than reckoning exclusively in optimization procedures.*

Keywords: Parametric Volatility; Synergy; Efficient Frontier; Critical Input; Maximum Entropy; Risk Analysis

JEL Classification: C53, G01, G17, G18, G32, G34.

Type of manuscript: research paper

Received: 15.01.2023 **Accepted:** 18.02. 2023 **Published:** 31.03.2023

Funding: There is no funding for this research.

Publisher: Academic Research and Publishing UG (i. G.) (Germany)

Cite as: Videira, H. De C. (2023). The Synergic Entropy. An efficient frontier output derived from merged input units boosted by synergy and constrained by critical input. *Financial Markets, Institutions and Risks, 7*(1), 39-70. http://doi.org/10.21272/fmir.7(1).39-70.2023

Introduction

In Finance and Operations, entropy is commonly regarded as a measure of randomness, uncertainty, and risk. Conversely, the synergic entropy does not rely on this physics approach at first, but rather is conceived depending on the output function and new parameters as synergy, critical input (scarce resources), and output acceleration. Thus, the current approach is born inside Finance and Operations, using synergy derived from combinations among inputs as the main engine. However, as the design relies on output deviations, randomness, and mergers of inputs, some commonalities with entropy should appear in the final solution. As expected, the particular case of the synergic entropy gives Shannon's (1948) and Boltzmann-Gibbs entropies, which acknowledges synergy as closely related to entropy – the combined effect from mergers among inputs that is greater than the ordinary sum of the individual effects.

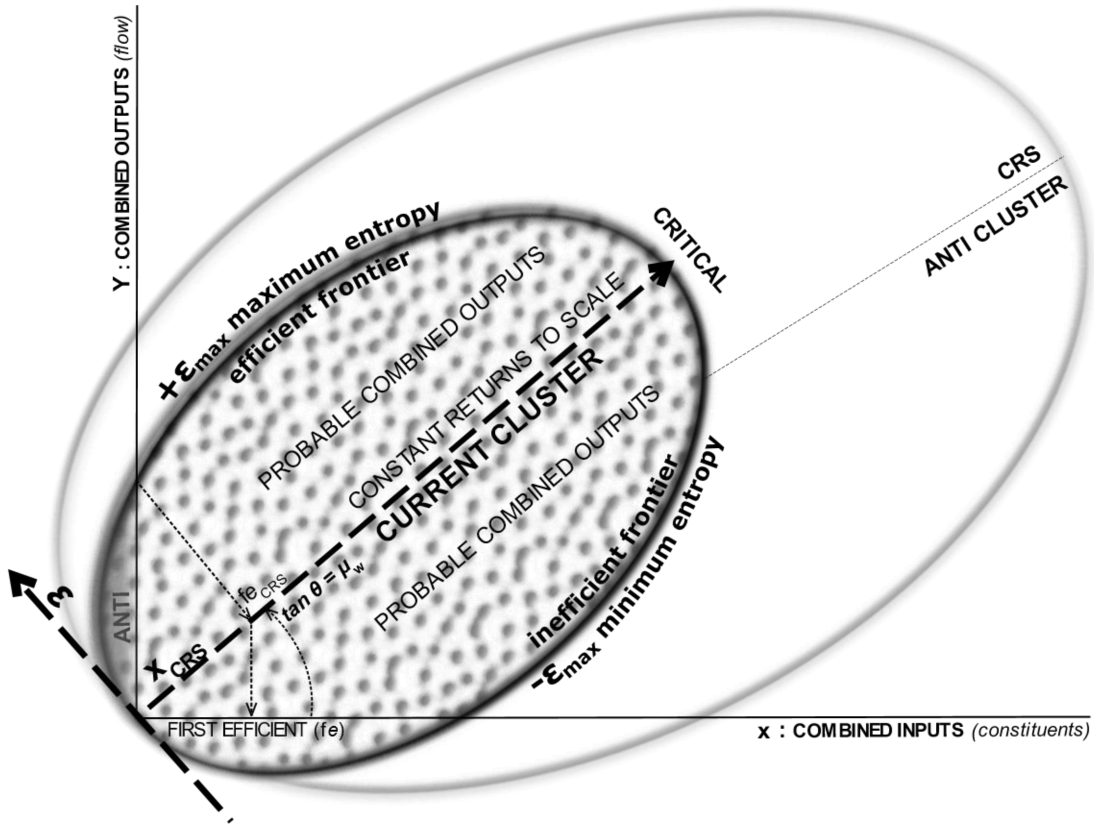

Figure 1. Theory Schematics

Source: compiled by the author.

X axis (input) is the value from possible mergers among individuals or constituents in pairs, trios, and so forth. Y axis (d_output/dt) is the Flow derived from these combined constituents, which probable values are shown in gray dots. The reference axes in the differential equation are x_{crs} (constant returns to scale) and ε (entropy), whose maximum entropy is $+\varepsilon_{max}$ and the minimum entropy is $-\varepsilon_{max}$. Maximum entropy ($\pm\varepsilon_{max}$) *is also regarded as the $\pm$efficient frontier, which is a set formed by outputs sorted by maximum and minimum ordered derivatives with respect to x_{crs}*. The regions given by the label *anti* are either located at an outer cluster or not achievable in the current configuration (shaded area in the second quadrant). The projection of the intersection between the Y axis and the efficient frontier in x_{crs} is fe_{crs}, which projection in x is the first efficient unit of the cluster (fe) – the minimum number of constituents to ignite the cluster. Critical input is a constraint that gives the maximum number of constituents (inputs) available to gather in a cluster.

The first insight in the novel theory is to defy the mainstream behavior of volatility as being stochastic. Rather than being purely stochastic, volatility could eventually be semi-deterministic under new parameters, such as synergy and critical input. Thus, the main question that arises is how to forge this semi-deterministic volatility, aiming at synergy as a new parameter for efficiency and risk.

The second insight that emerges to solve this puzzle is that volatility can be bounded by maximum and minimum outputs, in a fashion similar to the efficient frontier given by DEA (Charnes et al. (1978)). Therefore, all the probable outputs are inside a container provided by the efficient and the inefficient frontier, which are the maximum and minimum entropy, respectively. This insight that provides an envelopment for volatility is the foundation of the schematics shown in Figure 1, a novel approach based on the output function. The third insight that comes to light is that these boundaries can be built by combining the most efficient and inefficient input units by their ranking position, rather than relying only on optimization procedures.

The fourth and final insight is that input units share affinity-synergy among themselves, which is the main engine that boosts combinations. Hence, this definition implies that the greater the synergy, the greater the outputs given by the efficient frontier; therefore, synergy is the parameter the model seeks to equate as a proxy for risk under scarce resources (critical input). As the merger procedure is the key to the main question, the reasoning is based on the output generated by possible mergers among input units (microconstituents) with affinity-synergy. These input units can be assets of a company, companies inside a sector, or another physical variable that is merged over time. These merged inputs generate an output flow, whose value can be a quarterly income, an energy stream, or another similar time-related parameter.

The output function, the root of the current approach, has been discussed by economists since the late 1800s (Wicksteed (1894)). The most discussed approach is the one that relates output to capital and labor inputs, which is widely covered by literature (Cobb and Douglas (1928), Arrow et al. (1961)). After the seminal article from Arrow et al. (1961), which defined the Constant Elasticity of Substitution (CES), the research on the production function in economics began to decline after the seventies. Part of this fall was probably due to the article from Charnes et al. (1978), which presented DEA's approach founded on Farrel's (1957) proposals. Rather than only presenting an output function for labor and capital, DEA is a non-parametric technique that delivers an efficient frontier comprised of multiple inputs and outputs. The efficient frontier given by DEA is based on benchmarking peers from the same cluster, whose data for the model is straightforward. Thus, due to the broad scope and ease of use, several solutions were presented for many distinct cases in literature: high-tech industry in Feng et al. (2022); banks and finance systems in Silva et al. (2017), Fukuyama (2020) and Zhou et al. (2018); cities in Chen (2017); turbofan aero-engines in Kottas et al. (2011); airlines in Ngo and Tsui (2021); construction firms in Christopoulos et al. (2015); iron and steel industry in Wu et al. (2017); museums in Basso et al. (2017) and several other applications discussed in recent literature. Most of the applications above use an enhancement of DEA developed by Färe and Grosskopf (1996), known as the network model. In this approach, which was conceived primarily for studying the internal functions, the outputs from one stage are the inputs of a further stage. Likewise, the synergic entropy adopts the concept of reflux, which means that part of the output flow returns to the system by an input flow (see Theory's Section).

A typical criticism of the mainstream approaches has been the lack of time embedded in the models, though DEA addresses this issue through dynamic models, as seen in Zýková (2022) and Tone and Tsutsui (2014 and 2010). Moreover, another essential constraint not commonly covered by literature is the finite number of resources, which is the novel critical input parameter. Conversely, both time and critical input are vital economic concepts that are the core of the current approach. Furthermore, the typical economic output functions do not consider the efficiency among companies in a single model, which is the core of DEA and synergic entropy.

As briefly mentioned, the particular case of the current theory leads to the entropies given by Shannon (1948) and Gibbs-Boltzmann[12], which is explained in detail in Appendix B (Shannon's Approach). In such theories, the probability of achieving higher values of entropy is lower, which is the same in the current approach: each occurrence of ε in Figure 1 is given by a possible range of inputs in x_{crs} axis, which interval decreases until $\text{critical}/2$. As critical input is a constraint derived from scarce resources, it can also be acknowledged as the limited number of constituents available to gather in a cluster. Furthermore, at critical input, there is only one possible combination of microconstituents, which is the combination comprising all input units merged simultaneously. Hence, there is no randomness (zero entropy) at critical input because just one combination is possible (all inputs joined). Along with critical input, another novelty relies on the creation of the fundamental unit of the system, where efficiency is achieved for the first time – the first efficient point (fe_{crs}) in Figure 1. This fundamental point can be acknowledged as the minimum number of constituents required to ignite the cluster.

The practical question that arises is how to equate the phenomenon described in Figure 1 using these key concepts. This puzzle is solved in Theory's Section using combined output acceleration as inputs are merged. The solution of the differential equation leads to the following formula:

$$\varepsilon_{max}{}^2 = \frac{3}{ln\ c} . \phi^2 \sigma^2{}_{weighted(x_{crs},t)} ln\left(\frac{c}{x_{crs}}\right) \mid x_{crs} \ \geq \ fe_{crs} \quad (1)$$

$$if\ CRS\ (tan\ \theta \ = \mu_{weighted(x,t)}) \ \approx 0 : x_{crs} = x;\ \varepsilon_{max} \ = \ Y_{max(x,t)} - \mu_{weighted(x,t)} x$$

[12] Gibbs-Boltzmann and Shannon approaches for entropy can be regarded as mathematically equivalent.

In the above equation, *t*he input given by x is chosen as the quarterly total assets detained by each company, although other linear combination using multiple variables is suitable (Appendix C). The output flow given by $Y_{max(x,t)}$ (d_output/dt) *is chosen in this article as the summed combined net incomes ranked by return, though any other flow measure can be applied as well; hence,* $Y_{max(x,t)}$ *is the sum of the combined net incomes ordered by returns on assets (ROA), being ROA the same as the derivative* $dY_{(x,t)}/dx$ *provided by each input unit. The weighted mean and the weighted variance of returns (ROA) is given by* $\mu_{weighted(x,t)}$ and ${\sigma^2}_{weighted(x,t)}$, respectively.

Theory is based on the maximum fluctuations given by ${\varepsilon_{max}}^2$ over the x_{crs} axis, which is the constant returns to scale axis ($CRS = \mu_{weighted(x,t)}$). Hence, the whole data must be rotated counterclockwise until CRS by an angle equal to $tan\,\theta = \mu_{weighted(x,t)}$ (Figure 1) as described in Appendix A. If CRS is small enough ($tan\,\theta \approx sin\,\theta$), ε_{max} can be evaluated as $Y_{max(x,t)} - \mu_{w(x,t)}x$. The squared maximum entropy given by ${\varepsilon_{max}}^2$ depends on a longitudinal effect given by (${\sigma^2}_{weighted(x_{crs},t)}$) along with a cross-sectional effect given by $ln\binom{c}{x_{crs}}$. Hence, the cross-sectional effect is given by the natural logarithm of combinations among constituents up to critical input (c) (Subsection 2.3). As opposed to critical input that measures the maximum number of constituents in a cluster, the first efficient constituent (fe_{crs}) is the minimum required number of units to ignite the cluster in the x_{crs} axis. Finally, synergy (ϕ) can be defined as a constant derived from the effect given by merged input units with affinity over the absence of combinations, which higher values may amplify entropy providing more volatility; conversely, there are sectors whose synergy decreases volatility instead of providing more entropy. This shrinking phenomenon seems unreasonable, but synergy actually measures the summed effect derived from interactions among microconstituents, whose result can amplify or decrease volatility.

As shown above, synergy can also be regarded as the strength of interactions among constituents, which is neglected in most of the entropy's theories until the early 2000s. On the other hand, Beck and Cohen (2003) and Sattin (2006) presented a new interpretation of Tsallis's (1988) entropy, where the coefficient q could be acknowledged as a measure for these interactions among constituents. This new branch is commonly recognized as *superstatistics*. Besides *superstatistics*, there are other applications in literature based on the seminal approach from Tsallis, such as evaluating the stock market using structure entropy (Zhu and Wei (2021)). However, rather than relying primarily on statistics, the current approach is founded on operations and physics concepts developed a priori.

As predicted by theory, a practical application for synergy must be a novel proxy for risk - values above one and below one for synergy expands and shrinks volatility, respectively. The tests regarding this hypothesis are done over the financial statements derived from the 11 sectors of the American Economy (2064 companies), though most of the tests are done over the Consumer Cyclical Sector (324 companies). The results are significant for every sector except for the Healthcare cluster (as explained in Discussion, Healthcare should be broken up by its subsectors, setting apart drug producers from health plan providers). Beta CAPM (Sharpe (1964), Lintner (1965), and Mossin (1966)) tests for financial statements, conversely, do not present significance in any sector whatsoever. Hence, typical parametric approaches do not usually give the required significance for risk management – DEA and other machine learning tools have been widely used to address this issue, as seen in Ben Lahouel et al. (2022), whose research applies DEA in the banking industry aiming at liquidity risk and financial stability. Besides the new proxy for risk given by synergy, the current approach provides a new tool for valuation, which is a novel discount rate. This discount rate is conceived by multiplying synergy by the weighted standard deviation of returns, which proves to be stable and more suitable than typical betas.

The remainder of the article is organized as follows. The Second Section contains the reasoning for the Synergic Entropy Theory. The Third Section contains the Methodology, explaining the application of the synergic entropy in Financial Data in several steps. The Fourth Section presents the Results of the financial statements. The Fifth Section contains the Discussion of the Results. The Sixth Section presents the Conclusion, and the Seventh Section is about the References. Finally, there are four Appendices, as follows: Appendix A is the Rotation Method over the CRS; Appendix B is Shannon's Approach for the Synergic Theory; Appendix C is the Linear Combinations of Inputs; and Appendix D is the Ellipse Method, which is an estimate for critical input when the variable is unknown.

Financial Markets, Institutions and Risks, Volume 7, Issue 1, 2023
ISSN (online) – 2521-1242 ISSN (print) – 2521-1250

1. Theory

1.1. Design of the Empirical Envelopment

The Synergic Entropy relies on combinations among input units that lead to maximum and minimum entropies, which equal absolute value was previously defined as the maximum entropy. These maximum and minimum entropies can be regarded as the efficient and inefficient frontier from DEA theory, respectively. Before solving the maximum entropy equation, the first step is to build the empirical envelopment given by the boundaries of the gray ellipse in Figure 2 – the empirical data used to create Figure 2 is derived from financial statements retrieved from American companies in Consumer Cyclical Sector (better detailed in next section – Methodology). The "empirical envelopment" task is required because this numerical procedure is built by ordering and merging the available empirical data.

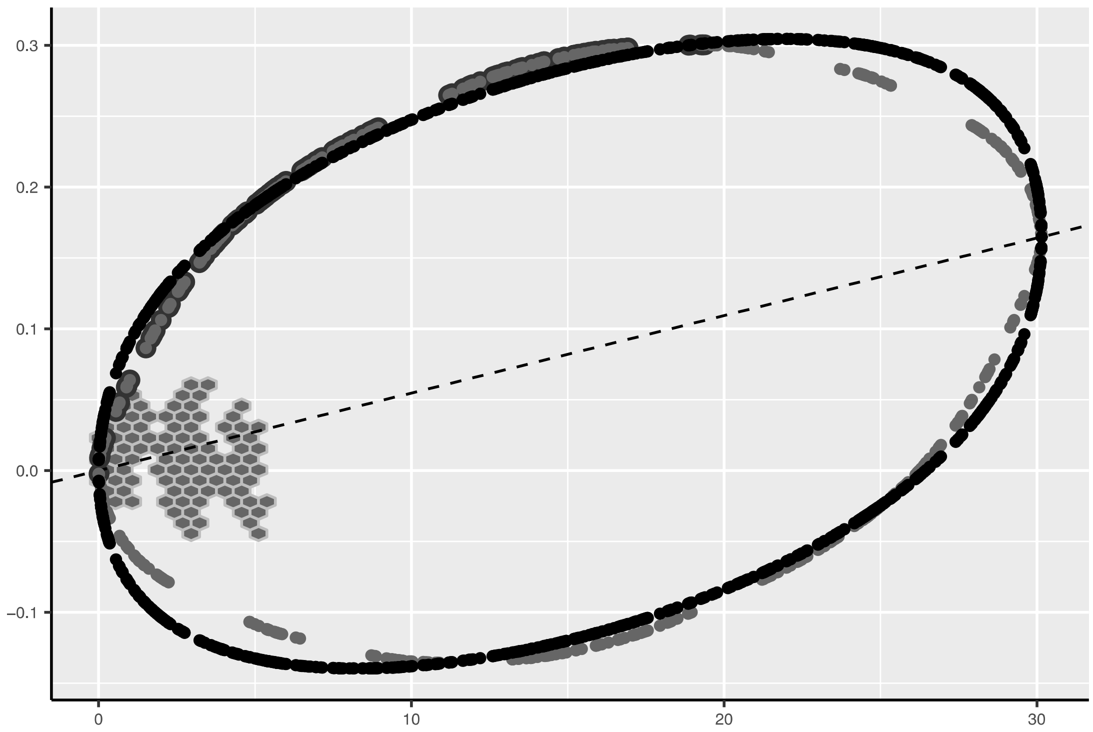

Figure 2. Output Combinations related to Merged Inputs in the Consumer Cyclical Sector

Source: compiled by the author.

Figure 2 is derived from the 4Q2019 balance and income sheets from 324 companies in the Consumer Cyclical Sector (as seen in Methodology Section). The small gray dots and hexagons are single outputs or possible mergers among companies in the Consumer Cyclical Sector. X-axis (Input) is the Total Asset value from individuals or the sum of combined assets when companies are joined. Y axis (Output) is the Income derived from individuals and combined incomes from companies in pairs and trios (hardware limitations limit other combinations), which are shown as gray hexagons; and Y is also a set formed by outputs sorted by minimum and maximum returns, which is the envelopment of the maximum entropy – a gray ellipse with small gray dots. Bigger gray dots above the dashed line up to the maximum derivative are the result of the DEA VRS model (efficiency >97.5%), computed using the empirical frontier. The dashed slope is the weighted average return of the cluster (CRS). The black ellipse is the fitted model from the Synergic Entropy given by (1).

In order to make the envelopment approach clearer, let the article provides a simple example derived from economics. In this theoretical example, a typical industrial plant, called Industry Co, has a classical production function depending on the labor for the short run. In this plant, the law of diminishing marginal returns holds when the building gets overcrowded by the rise of employees. Each employee has different productivity levels, which can be measured, and the number of employees (resources) available to hire in the industry market is limited (critical input). If one uses the typical approaches from literature, the output function might be regressed under labor. In standard practice, the first stage of the curve is a constant slope (constant returns to scale), but the second stage presents a diminishing rate slope where the limit given by capital (physical space) is a constraint.

However, the current approach aims at the most efficient employees from the whole sector (Industry Cluster) where Industry Co is classified. Hence, the current approach requires that the data for the entire sector concerning labor productivity be available. Therefore, the current approach is a benchmarking method in the cluster, rather than an individual assessment. The best company of the Industry's Cluster, which only input variable is labor, is the one that has the best selection tools to pick the best-ranked professionals. Thus, if one makes a ranking for the best productivity professionals, the most efficient virtual company is the one that has all these star professionals until the labor amount required for its own needs. The same reasoning is applied to the most inefficient virtual company. Therefore, the output production function has two boundaries: an upper limit formed by the best employees and a lower limit, created by the most unskilled employees. It seems reasonable that the upper limit is built by ordering the productivities from maximum to minimum, which is the same as ranking the professionals or ordering the derivatives (productivity). The same reasoning applies to the inefficient frontier (lower limit). Thus, after ordering the derivatives (labor productivity), the envelopment is the area formed by the upper and lower limit, which correspond to the efficient and inefficient frontier, respectively. This boundary of probable results is the frontier of the ellipse shown in Figure 1 - Figure 2 and corresponds to the maximum entropy.

Therefore, the curve of the output could be written as follows:

$$\Delta Y_c = r_j \Delta x_c \,, \qquad (2)$$

where Y_c is the cumulative output flow, x_c is the cumulative stock input and r_j is the return from each input unit j used in the cumulative process.

For the given example, x_c is the number of summed professionals (labor), Y_c is the income/time given by x_c and r_j is the productivity from each professional. The order that the cumulative input units are gathered will derive different paths, which the best one is the efficient frontier (maximum entropy) given by the ordered ranked selection.

In Figure 2, the empirical hexagons plotted inside the ellipse are the output combinations from individuals, pairs, trios, and so forth up to the 324 companies, which figure still needs to be fulfilled due to hardware limitations. Notwithstanding, the current approach solves this issue using the ordered derivatives to build the envelopment, as seen in (3). The fitted model from (1) is seen in Figure 2 as the black ellipse line around the gray envelopment (the empirical frontier). However, like DEA, the focus remains on the boundaries of the curve.

If the evaluating cluster has enough number of individuals, (2) can be rewritten as:

$$\frac{dY_c}{dx_c} = r_j \qquad (3)$$

Therefore, the derivative of the cumulative output with respect to a given input is the return of each infinitesimal individual joined in the process. As aforementioned, the path of the cumulative function depends on the order of the combined input units. If one aims to design the boundaries of the curve in Figure 2 (envelopment), the derivatives must be ordered following the rules below:

- Efficient Frontier (maximum entropy) = Derivatives (Returns) in descending order
- Inefficient Frontier (minimum entropy) = Derivatives (Returns) in ascending order

As discussed above and presented in Subsection 2.12, multiple inputs instead of a single one can be used for x_c when the input is comprised of a linear combination of variables.

2. Affinity and Anti-Cluster

As the novel theory's foundation relies on combinations, a central question arises: what is the main engine that drives input units to combine with each other? The first insight is that input units with affinity are merged by an irresistible desire to reach greater combined outputs. In other words, input units are merged because the combined effect is greater than the sum of the individual effects from each input unit, which is the synergy's definition.

Unfortunately, this affinity among input units decreases as more input units are joined, because the probability of finding similar assets decays when cumulative input grows. In the current approach, the "affinity" force generates higher variable returns to scale and higher entropy (higher variance).

A second effect happens when input units with affinity combine. The decaying probability of finding similar assets does not decay at the infinite, but rather is constrained by the available similar inputs in the cluster. This maximum number of available input units with affinity in a cluster is named critical input. It means, for instance, that one cluster has limited resources, which could be available matter, raw materials, customers and skilled labor. In the Industry Co. example, critical input is the sum of every professional available in the industry's cluster. Suppose Industry Co. has become big enough to be the cluster itself. In that case, there is only one possible output: the sum of outputs from every professional of the cluster, without any other possible combinations in the outcome. Thus, at critical input, the deviation from the expected mean is zero, and the ranked selection is useless; thus, at critical input, there is no other choice for Industry Co. rather than hire all the available professionals from the market, without any efficiency advantage.

In order to avoid the critical point, an anti-cluster force pushes the cluster to become another one, no longer limited by the critical input. In this new configuration, critical input is a value in a higher dimension that postpones the zero-deviation phenomenon. In this case, a potential and profitable merger exists with other input units in a new cluster configuration. Therefore, as the cluster follows the path to critical input, the chance of a cluster becoming another one is more significant due to the absence of higher returns above the mean. This new cluster configuration could merge from outer input units into the current one or collapse. If the cluster collapses, its input units can be combined in another greater system or scattered through multiple systems.

3. Critical Input and First Efficient Point

The First Efficient Point $(\boldsymbol{f}e_{crs})$, synergy (ϕ) and critical input (c) are the key parameters for a cluster, which define its creation, behavior, and termination, respectively. Although the first efficient point is not present in the core of the equation given by (1), its value is the first point of the domain: the beginning of the lifecycle.

4. Critical Input

A critical input in the current approach claims that input available to combine in a cluster is made from finite resources or limited individuals. These finite resources or limited individuals are a threshold given by a constant that can be known or unknown. Regardless of understanding the value from critical input (c) previously, the essential feature is that there is only one possible combination at this critical point. The only one combination rule occurs because every input unit must be combined to achieve the value of critical input, not leaving a chance for other possible combinations. Therefore, every known and unknown input unit is combined at critical input, leading to zero entropy or uncertainty (zero variance over the mean).

There are systems in which the critical input is a trivial constant to retrieve, being a known value. The known critical input value in a sector can be the number of available employees to hire in a sector (as given by the example above), the amount of raw land in a country, the supply of a commodity, the limit of machinery in a cluster, and further on. On the opposite, there are systems in which critical input is an unknown value subject to investigation (by the behavior of the other parameters).

In systems where critical input is a non-trivial value, its magnitude can be steady or non-steady. For systems with steady unknown critical inputs, the article presents the Ellipse Method described in Appendix D as a solution for estimating the parameter; furthermore, the same Ellipse Method can be applied to non-steady systems whether the variance of critical input is constant through time.

For non-steady systems, the variance of critical input may occur due to random errors, which can lead to a constant variance through time. In this case, critical input can be evaluated by the Ellipse Method (Appendix D) whether having a constant variance derived from these nonsystematic errors. Conversely, whether the variance is not constant for critical input, the definition of the cluster itself must be changed. This constant variance of critical input in non-steady systems should not be misunderstood with the zero output's variance as aforementioned. An example of a non-steady system with unknown critical input is the one derived from financial statements or other accounting figures, which values may contain random errors - critical input variance can be constant or not.

Based on the previous assumptions, critical input can be summarized as follows:

- Known Critical Input
- Unknown Critical Input
  - Steady value through time. Ellipse Method in Appendix D can be used to evaluate critical input.
  - Non-steady value having constant variance through time. Critical input can also be evaluated through Ellipse Method.
  - Non-steady value with non-constant variance through time. If the critical input's variance is not constant, the definition of the cluster itself must be changed.

Evaluating critical input in empirical cases is a challenge. Even in cases where critical input is a known value, it is unlikely that every input unit of the cluster is available to join in a specific cross-sectional. Therefore, if one uses the known critical input value, the boundary condition (shown in the following sections) will not be met because the sum of the input units at an empirical cross-sectional will not equal the known critical input value.

The solution for this puzzle is an approach where two instances of critical input are built: the cross-sectional critical input and the longitudinal critical input. The cross-sectional critical input is the sum of every available input unit at a cross-sectional (retrieved empirical data). In contrast, the longitudinal critical input is the sum of the seen (retrieved) and unseen input units in a cluster. This approach is used only for solving this empirical puzzle and does not change the shape of equation (1), in which critical input (c) - depending on the assessment - can be the cross-sectional or longitudinal.

The cross-sectional critical input must be chosen whenever one wants to evaluate the efficient frontier comprising the available input units at a certain moment. Another possible use of cross-sectional critical input is to calculate synergy computed from cross-sectionals regressions. On the opposite, longitudinal critical input can be desirable whether one seeks the unobserved critical input of the system. This unobserved value is the edge of a system, a critical point where the system collapses or turns into another one. In this article, the symbol c also refers to cross-sectional critical input, otherwise noted. Thus, the main features of the cross-sectional and the longitudinal critical input are:

I. Cross Sectional Critical Input

a. Sum of the available input units in a cross-sectional

b. Compute entropy (efficient frontier) from available input units in a cross-sectional.

c. Calculate synergy from cross-sectionals regressions.

II. Longitudinal Critical Input

a. Sum of the available and unseen input units in a system

b. Compute the critical value of the system, and where the system changes (Appendix D).

Finally, the input can be comprised of a linear combination of variables, as seen in Subsection 2.12 and Appendix C (Linear Combination of Inputs). This set can be formed by constrained or unconstrained variables. The constrained variables have critical values, whereas the unconstrained variables have infinite input values. Thus, if any of the variables has unconstrained values, critical input will never be reached because there is an endless input source derived from this variable; therefore, there is no competition among input units because there is plenty of inputs. At last, the unconstrained inputs must be kept out of the function removing their values from the output.

5. First Efficient Constituent (Fundamental Unit)

In the early days of a cluster, the constituents are eager to gather to create the first unit of the system (fundamental unit), which must be efficient enough to ignite the cluster - this is the beginning of the lifecycle of a cluster. The fundamental unit for some systems could be trivial because it is simply the number one (1): the first unit of the cluster. In other clusters, due to scale, discover the fundamental unit ($\boldsymbol{f}e_{crs}$) could be the objective of the research itself. In these cases, $\boldsymbol{f}e_{crs}$ can be retrieved by computational methods and further statistical developments.

6. Output Acceleration

Retrieving the example for Industry Co. one more time, one can establish a simple output production given by the following equation:

$$Y_f = rx\,, \tag{4}$$

where Y_f is the output flow (income/time), x is a stock variable that measures the level of inputs (number of professionals), and r is the constant returns to scale (CRS), which is the productivity's mean. Therefore, if the level of the input (the number of skilled workers) is increased by a certain amount, the increased output (incomes/time) is increased by the same amount in a CRS approach.

As Industry Co wants to benefit from the economies of scale, the company reinvests part of the monetary output flow in the production, resulting in an increased input flow. Rather than only labor, part of the input flow could be due to other input variables like materials, machinery, and others. Therefore, this relation between input flow and output flow is established by the following equation:

$$X_f = zY_f \tag{5}$$

where X_f is the input flow and z is the reinvestment coefficient. The variables presented in capital letters in further equations are flow variables, whereas lower-case letters correspond to stock variables.

The output flow (Y_f) and the input flow (X_f) for the infinitesimal period can be written as following, where y is the stock value from the output.

$$Y_f = \frac{dy}{dt} \tag{6}$$

$$X_f = \frac{dx}{dt} \tag{7}$$

If one plugs (6) and (7) in (4), one gets:

$$\frac{dy}{dt} = rx \tag{8}$$

$$\frac{dx}{dt} = rzx \tag{9}$$

Equation (9) can be better acknowledged if one simple finance example is provided. If one takes as input its invested capital in Treasury Notes, the output function could be the TNX yield for that period. If the yield r is totally reinvested ($z = 1$: total capital reinvested), solving (9) with respect to x leads to: $x(t) = e^{rt}$. Thus, the reinvestment concept in the equations above applied in Finance is a well-known result for compounded returns.

The reinvestment in input flow detaches from most typical approaches in the output function. Though it seems simple, all the following equations derive from this concept. Retrieving the Treasury's example, it is evident that the output flow is not constant when the input flow rate is directly related to the output. It is worth emphasizing this point: the output flow rate does not remain steady if part of the output flow is reintroduced in the system by an input flow. This reinvestment in input units is not a hypothetical premise: reinvestment is required if a company wants to move from its current production level (x_c) to another. In closed physical systems, a significant part of the output flow (e.g.: energy) is kept inside the system, which is added to the input units themselves in a chain reaction.

Furthermore, one can define the rate of the output function as the first derivative from (6), which turns out to be defined as the output acceleration (g_y):

$$g_y = \frac{d(Y_f)}{dt} = \frac{d^2y}{dt^2} \tag{10}$$

If one uses (10) aiming at computing the output acceleration for (8): CRS , one gets:

$$g_y = r.\frac{dx}{dt} = r.rzx = zr^2x \quad (11)$$

The simple result given by (11) plays an important role in the novel theory, as this is the start point for modeling the forces. Based on (11), it is intuitive to conceive that infinitesimal output acceleration variation (da_y) may depend on input infinitesimal variation.

Thus, (11) could be rewritten as:

$$dg_{combined} = z.\alpha.dx \quad (12)$$

Where α is the affinity which intends to capture the cross-sectional and longitudinal phenomena. One can notice that acceleration now has a subscript presenting the combination. It means that output's acceleration is an effect derived from combinations derived from the input x.

A simple solution for (12) gives:

$$g_{combined} = z.\alpha.x + C \quad (13)$$

Equation (13) states that the acceleration of production (output) increases as more input units are combined because part of the output flow is reintroduced in the system by reinvestment. Besides the financial example, other cases for this flow's acceleration could be the growth of a bacterial colony, a city growth with limited resources, a chain reaction, and many other possible practical examples.

7. Flattened Output Acceleration

The foundations regarding output acceleration were previously presented, but the output acceleration that governs the phenomenon has yet to be shown. As discussed before, the primary law for synergic entropy is that the forces of the phenomenon are flattened by the decreasing probability of finding similar input units with affinity. Therefore, the first approach is the flattened output acceleration $(g_{flat(x,t)})$ given by the following based on (12) :

$$dg_{flat(x,t)} = z.\alpha_{(x,t)}\frac{dx}{x} \quad (14)$$

A simple solution for (14) gives:

$$g_{flat(x,t)} = z.\alpha_{(x,t)}ln(x) + C , \quad (15)$$

where $\alpha_{(x,t)}$ is the affinity coefficient that merges the cross-sectional and longitudinal effect.

The first step to acknowledge (14) is to realize that infinitesimal variation of flattened output acceleration depends on successive increments of input units, just like (12). However, rather than stated by (12), the incremental is flattened by (dx/x) due to the decreasing probability of finding similar input units with affinity. The flattening process is the main foundation of the novel reasoning: as the input grows, affinity decreases, inversely proportional to cumulative input. Therefore, equation (15) is the first step to design the forces above and below the CRS, which is the reference axis for the logic developed in this article.

8. Affinity and Anti-Cluster Forces

The previous subsection discussed the flattening process, turning the acceleration weaker as the input grows. As mentioned before, the acting forces act above or below the CRS, which is the reference axis. However, the accelerations are based on natural logarithmic, so their rate would be below the CRS' rate if the CRS were not the reference axis. Therefore, forces are modeled above and below the dashed line (CRS) presented in Figure 1 and Figure 2.

The main difference between the anti-cluster and the affinity force is that the latter depends on critical input, presenting decaying values as the input grows near to critical point. When input grows, affinity is affected by the proximity of the critical point. In contrast, the anti-cluster force acts to find a better cluster configuration (avoiding the lack of input units).

Given the theoretical reasoning, critical input does not constrain the anti-cluster force. Thus, it can be defined by the basic formulation equated in (16), where g_{ac} is the anti-cluster output acceleration:

$$g_{ac\,(x,t)} = z.\alpha_{(x,t)}\, ln(x) + C_{ac} \tag{16}$$

Though not affected by critical input, the anti-cluster force has the exact decaying probabilities of finding similar input units to achieve a new better cluster configuration – a flattened output acceleration.

By contrast, the affinity force depends on critical input, which is the number of inputs available to combine in the cluster. Thus, affinity force must be solved considering critical input and its formulation is presented below.

It is natural to infer that affinity gradually diminishes as the input grows, because the probability of finding similar input units decreases. Thus, as input increases, the bond among units loses its strength, reaching the maximum loss at critical input. In other words, as the input gets closer to critical input (c), the minor is the value of $(c - x)$ and the greater is the loss of affinity.

Therefore, before reaching the affinity formulation, one can make a summary of the premises already discussed:

- As input grows, affinity decreases; on the opposite, the anti-cluster force increases.
- Affinity depends on critical input; conversely, anti-cluster does not depend on a critical point, acting to achieve a better cluster configuration. However, both forces have the common law of diminishing available input units to gather.
- As input grows, the incremental from input units generates an affinity loss. This is a negative flattening process.
- As $(\boldsymbol{c - x})$ becomes lower, loss of affinity becomes greater until reaching its maximum point at critical input.

Based on premises from iii. to iv., equation (13) can be rewritten to formulate the affinity force ($g_{aff(x,t)}$):

$$dg_{aff(x,t)} = z.\alpha_{(x,t)} \frac{-dx}{(c-x)}$$

$$g_{aff(x,t)} = z.\alpha_{(x,t)}\, ln(c-x) + C_{\alpha}\ , \tag{17}$$

where c is the critical input.

The equations (16) and (17), present, respectively, the anti-cluster acceleration and the affinity acceleration. The following subsection will mount the differential equation based on these forces.

9. Mounting the Differential Equation

As the forces are established, it is time to build the differential equation, determining the behavior of the cluster.

Even declining as input evolves, the affinity force is expected to be greater than the anti-cluster one at the first stages. If the opposite were true, the cluster would not begin any combination. Therefore, the differential equation is written based on the positive sign of the affinity acceleration, which is the opposite of the anti-cluster force. So, the reference for the system's output acceleration is based on the positive affinity force, or, in other words, the positive deviation of outputs.

Another important point regarding the reference system is that the forces are designed using the CRS (constant returns to scale) as the reference axis for inputs. Therefore, the x presented in the last equations will be replaced by x_{crs}. Likewise, deviations for outputs above and below the mean are the system's entropy given by ε (entropy). As the design of the model is based on the maximum entropies, the reference axis Y will be replaced by ε_{max}.

Finally, the terms are mounted to discover the final output acceleration of the cluster, already knowing the acting forces:

$$a_{cluster(x,t)} = a_{affinity(x,t)} - a_{anti-cluster(x,t)} \tag{18}$$

Plugging in (10), (16) and (17) in (18):

$$\frac{d^2Y_{eff}}{dt^2} = z\,(\alpha_{(x,t)}\,ln(c-x) + \frac{C_\alpha}{z} - \alpha_{(x,t)}\,ln(x) - \frac{C_{ac}}{z})$$

Translating the equation in the correct reference system given by the reasoning (CRS):

$$\frac{d^2\varepsilon_{max}}{dt^2} = z\,(\alpha_{(\,x_{crs},t)}\,ln(c-\ x_{crs}) + \frac{C_\alpha}{z} - \alpha_{(\,x_{crs},t)}\,ln(\,x_{crs}) - \frac{C_{ac}}{z})$$

where c is critical input measured in CRS axis.

From (5), having constants merged:

$$\frac{d^2x_{crs}}{dt^2} = z^2\left(\alpha_{(x_{crs},t)}\,ln(c-x_{crs}) - \alpha_{(x_{crs},t)}ln(x_{crs}) + C_0\right) \tag{19}$$

If one multiplies every term of (19) by dx and integrate every part, one gets:

$$\int \frac{d^2x_{crs}}{dt^2}\,dx_{crs} = z^2\left(\int \alpha_{(x_{crs},t)}\,ln(c-x_{crs})\,dx_{crs} - \int \alpha_{(x_{crs},t)}\,ln(x_{crs})\,dx_{crs} + \int C_0\,dx_{crs}\right)$$

Solving this simple integral leads to the following:

$$\frac{1}{2}\left(\frac{dx_{crs}}{dt}\right)^2 = z^2\left(-\alpha_{(x_{crs},t)}x_{crs}\,ln(x_{crs}) - \alpha_{(x_{crs},t)}(c-x_{crs})\,ln(c-x_{crs}) + C_0\,x_{crs} + C_1\right)$$

As we are interested in the Output, we can redefine the above equation using (5) :

$$\frac{1}{2}\,(\varepsilon_{max})^2 = -\alpha_{(x_{crs},t)}x_{crs}\,ln(x_{crs}) - \alpha_{(x_{crs},t)}(c-x_{crs})\,ln(c-x_{crs}) + C_0\,x_{crs} + C_1 \tag{20}$$

where ε_{max} is the maximum entropy.

One important point from the integration is that neither the affinity coefficient $\alpha_{(x_{crs},t)}$, nor the critical input, which may depend on time (longitudinal effect), are affected by the integration over x_{crs}.

The equation (20) could be daunting at first. However, the mathematical approximation given by the combinatorics approach (Subsection 2.10) gives the simplified version seen in (1). As the output flow given by (20) is the squared output flow, ε_{max} is the maximum entropy (efficient frontier); and $-\varepsilon_{max}$ is the minimum entropy (inefficient frontier). Moreover, the equation is built over the CRS line: ${\varepsilon_{max}}^2$ is a squared value above or below the cluster's mean, which is the concept of variance-entropy. Finally, it is worth remembering that (20) is designed for the CRS System (dashed line from Figure 1 and 2 with its orthogonal axis ε).

10. The Boundaries Conditions

The boundaries defined by C_0 *and* C_1 must be investigated to complete the final equation. As mentioned in the previous section, the output flow above or below the CRS line will be regarded as the output deviation, becoming a new measure for maximum entropy.

At first, the boundary around zero input is calculated. Then, if the input is zero or t*ends to zero*, it is expected that the variance output will also approach zero – the reason is that there are no available input units to combine, nor exists a real cluster.

As input approaches zero, the term $-\alpha x_{crs} * \ln(x_{crs})$ tends to zero as well. Thus, the equation (19) can be rewritten to determine boundaries when input approaches zero.

Thus, being $x_{crs} = 0$, follows:

$$\frac{1}{2}(\varepsilon_{max})^2 = 0$$

Plugging in (20) the term $(-\alpha_{(x_{crs},t)}x_{crs} * ln(x_{crs})\ =\ 0)$ leads to:

$$\frac{1}{2}\,(\varepsilon_{max})^2 = -\alpha_{(xx_{crs},t)}x_{crs}\,ln(x_{crs}) - \alpha_{(x_{crs},t)}(c - x_{crs})\,ln(c - x_{crs}) + C_0\,x_{crs} + C_1$$

$$\frac{1}{2}(0)^2 = 0 - \alpha_{(x_{crs},t)}(c) * ln(c) + C_0 * 0 + C_1$$

$$C_1 = \alpha_{(x_{crs},t)}c\,ln\,c \tag{21}$$

Plugging in (21) in (20):

$$\frac{1}{2}(\varepsilon_{max})^2 = \alpha_{(x_{crs},t)}c\,ln\,c - \alpha_{(x_{crs},t)}x_{crs}\,ln(x_{crs}) - \alpha_{(x_{crs},t)}(c - x_{crs})\,ln(c - x_{crs}) + C_0\,x_{crs} \tag{22}$$

The other extreme boundary must be solved now. Looking at the extreme side of the curve, as the input approaches critical input, the advantages of the combination tend to disappear. As discussed, the number of possible combinations at critical input is just one, leaving no other possible combination (zero variance).

Thus, when cumulative input tends to critical input, we get:

$$x_{crs} \to c : (\varepsilon_{max})^2 = 0 \tag{23}$$

As input approaches critical input, it leads to $c - x_{crs} \to 0$. Therefore, the following term $-\alpha(c - x_{crs}) * ln(c - x_{crs})$ tends to zero. Thus, plugging in (23) in (22) along with limit conditions at critical input, one gets:

$$\frac{1}{2}(0)^2 = \alpha_{(x_{crs},t)}c\,ln\,c - \alpha_{(x_{crs},t)}clnc + 0 + C_0c$$

$$C_0 = 0$$

Thus, the cross-sectional equation, after the boundaries are defined, is presented below, making the 0.5 coefficient in the left as part of the affinity coefficient's $\alpha_{(x_{crs},t)}$ expectancy:

$${\varepsilon_{max}}^2 = \alpha_{(x_{crs},t)}(c\,ln\,c - x_{crs}\,ln(x_{crs}) - (c - x_{crs})\,ln(c - x_{crs})) \tag{24}$$

The longitudinal effect has not been discussed so far. This topic is explained at Subsection 2.11, which establishes the affinity's ($\alpha_{(x_{crs},t)}$) expectancy depending on time.

Another important topic to address in equation (24) is that it can be rewritten as a probability function of x with respect to c, given by $p_x = x_{crs}/c$ - a measure of uncertainty. In this approach, as seen in Appendix B, equation (24) is a Shannon's entropy binary function.

11. Rotating the Axes

As discussed before, the correct system of reference at which the equation (24) is established is the line of CRS. Therefore, all data must be rotated in order to comply with equation (24). Thus, if one rotates x and Y making x aligned with the vector which slope is the dashed line in Figure 1 and 2 (CRS), one gets the correct reference system now. After rotation given by $tan\,\theta = \mu_{w(x,t)}$, the $y\ axis$ which is orthogonal to x_{crs} is the ε axis, which maximum value is the maximum entropy (${\varepsilon_{max}}^2$). The rotation matrix and the explanation for $tan\,\theta = \mu_{w(x,t)}$ is described in Appendix A (Rotation Method).

For systems with low CRS ($\mu_{w(x,t)} = tan\,\theta \approx sin\,\theta$), the equation (24) becomes (deducted in Appendix A):

$$\left(Y_{max\,(x,t)} - x\mu_{w(x,t)}\right)^2 = \alpha_{(x,t)}(c\,ln\,c - x\,ln(x) - (c - x)\,ln(c - x)) \tag{25}$$

where $\varepsilon_{max} = Y_{max(x,t)} - x\mu_{w(x,t)}$; $x = x_{crs}$

12. A Combinatorics Approach

The mathematical link between the novel theory and combinatorics is discussed in this subsection. As theory relies on combinations of inputs to develop its reasoning, one must expect that (24) *may have some connection with combinatorics. Trying to establish the connection between* (24) *and combinatorics, the Sterling Approximation* (26) *is applied to the natural logarithmic of combinations* ($ln\binom{c}{x}$)*:*

$$Sterling: ln(x!) = xln(x) - x \tag{26}$$

Plugging in Sterling Approximation (26) *to* $ln\binom{c}{x}$, *one gets the following expression after algebraic simplification:*

$$ln\binom{c}{x} = c\,ln(c) - x\,ln(x) - (c-x)ln(c-x) \tag{27}$$

Equation (27) *is exactly the right part of equation* (24) *multiplied by the affinity coefficient* $(\alpha_{(x_{crs},t)})$, *which leads to a Sterling approximation of the exact form:*

$$\varepsilon_{max}{}^2 = \alpha_{(x_{crs},t)}\, ln\binom{c}{x_{crs}} \tag{28}$$

The last part of equation (28) given by $ln\binom{c}{x_{crs}}$ does not depend on time, as proved by the solution of the differential equation over dx_{crs} : one could claim that the cross-sectional effect is given by the combinations among x input units, which threshold is c.

13. Evaluating Time in the Affinity Coefficient

14. Foundations

The affinity coefficient has been named so far as $\alpha_{(x_{crs},t)}$ to keep in mind that time needs to be evaluated, which is addressed now. However, it is worth remembering that time was not assessed so far because the solution of the integral given by (19) does not depend on time, but rather holds only for the cross-sectional. Therefore, the longitudinal effect must be equated.

If one retrieves the solution from (19), it is natural to infer that the affinity coefficient must be formed by two effects: one constant effect in x_{crs}, due to the cross-sectional solution from the integral; and another component derived from the variable longitudinal effect given by time. Moreover, their effect is multiplicative instead of an additive one because a sum inside $\alpha_{(x_{crs},t)}$ would change the solution from the integral. Based on this premise, the affinity coefficient $\alpha_{(x_{crs},t)}$ could be written as:

$$\alpha_{(x_{crs},t)} = \alpha_{(x_{crs})}{}^2 . \alpha_{(t)}{}^2 \tag{29}$$

where $\alpha_{(t)}{}^2$ is a time-dependent variable and $\alpha_{(x_{crs})}$ is the affinity constant.

In order to assess $\alpha_{(t)}{}^2$, the cross-sectional effect is frozen by letting x_{crs} constant in a slice of the Figure 1, which is taken at $x_{crs} = 1$ for theoretical purposes (ease of computing the efficient frontier). It is worth mentioning that even at the point $x_{crs} = 1$ the cluster is being merged with affinity and synergy among input units.

It seems unreasonable to assess the efficient frontier at $x_{crs} = 1$ when there are so many individuals with so different input values. However, each input unit, before the envelopment, is formed by several input units which sum is its own value for x_{crs}. Thus, these statements are described below:

$$\varepsilon_{i\,(x_{crs}=i)} = f(x_{crs_i})$$

$$r_{i(x_{crs}=1)} = \frac{f(x_{crs_i})}{x_{crs_i}} \tag{30}$$

$$\varepsilon_{i\,(x_{crs}=i)} = r_{i(x_{crs}=1)} . x_{crs_i} \tag{31}$$

where r_i is the return for each input unit.

Equation (31) seems a simple approach, though it explains that for a given input x_{crs_i} there are x_{crs} input units with the same output at $x_{crs} = 1$, which is r_i . A simple numerical example can explain this approach: a single input unit that has 100 inputs generating five outputs can be regarded as a sum of 100 input units with the same return (5%) at $x_{crs} = 1$.

Therefore, the probability p_i of occurrence for each input unit at $x_{crs} = 1$ leads to:

$$p_i\left(\varepsilon_{i\,(x_{crs}=1)}\right) = \frac{x_{crs_i}}{\sum_{i=1}^{n} x_{crs_i}} \tag{32}$$

where n is the number from individuals in the cluster.

15. Maximum Entropy given by Combinatorics

After main foundations, it is time to retrieve (28), making $x_{crs} = 1$:

$$\varepsilon_{max}{}^2 = \alpha_{(x_{crs},t)}\, ln\binom{c}{1}$$

$$\varepsilon_{max}{}^2 = ln(c)\, \alpha_{(x_{crs})}{}^2 . \alpha_{(t)}{}^2 \tag{33}$$

Thus, the maximum entropy at $x = 1$ depends only in a variable term given by (33), which is the time-effect ($\alpha_{(t)}{}^2$) being equated. One could argue again why the time effect is assessed at $x = 1$, rather than other sections. However, the main reason was given above, explaining that data provided by individuals is already available, without requiring computing the envelopment at the very first moment; however, the envelopment at $x = 1$ exists, being computed in the following subsection.

16. A uniform distribution for the Maximum Entropy

The size of the maximum entropy given by ε_{max} in (28) is based on the natural log of the combinations for each slice - each slice can be regarded as $x_{crs} = constant$. Hence, each slice is a gathering from multiple possibilities ranging from $-\varepsilon_{max}$ to $+\varepsilon_{max}$. At this slice, each occurrence has the same probability of happening as the others; therefore, it can be regarded as a uniform distribution from $-\varepsilon_{max}$ to $+\varepsilon_{max}$. This insight allows estimating the maximum entropy from another approach rather than the one given by (33).

Thus, in this subsection, the occurrences within the envelopment (the small gray hexagons in Figure 2) are compared against its extreme boundaries given by the maximum entropy ($\pm\varepsilon_{max}$). Finally, given a specific slice of $x_{crs} = constant,$ the occurrences inside the envelopment are named as ε, which are equally probable values in the ε axis. Therefore, any deviation inside the envelopment is $\pm\varepsilon$, which boundaries are the maximum entropies ($\pm\varepsilon_{max}$).

Therefore, the variance given by the uniform distribution ($p_{(\varepsilon)} = 0.5\varepsilon_{max}$) is given by:

$$\sigma^2 = \int_{-\varepsilon_{max}}^{\varepsilon_{max}} \frac{(\varepsilon - \mu_\varepsilon)^2}{2\varepsilon_{max}} d\varepsilon$$

As ε is orthogonal to x_{crs}, $\mu_\varepsilon = 0$. Solving the above equation, leads to:

$$\varepsilon_{max}{}^2 = 3\sigma^2 \tag{34}$$

The problem now relies on estimating σ^2, which is solved at first by estimating the variance from discrete individuals at $x_{crs} = 1$. If one uses this variance from individuals at $x_{crs} = 1$, synergy is not computed because there are several possible combinations given by $\binom{c}{1}$ (see (33)). Therefore, there is already affinity-synergy among input units at this location, which effect gives the maximum entropy. Thus, in order to correct this discrepancy, the effect derived from combinations must be computed by applying a factor over the variance from individuals.

In the following, there is a summary of the above statements:

- First, the variance inside the envelopment is estimated using the variance from discrete individuals, a known parameter.
- Second, a factor must amplify (or decrease) the discrete variance from individuals to achieve maximum entropy. This factor is derived from the power of possible combinations given by c at $x_{crs} = 1$.

Therefore, one can conclude from the above statements:

$$\sigma^2 = \phi^2 . \sigma_{ind}{}^2 \tag{35}$$

where ${\sigma_{ind}}^2$ is the variance given by discrete individuals without combinations and ϕ is a multiplicative constant – the synergy effect provided by the power of combinations, which could amplify or decrease the weighted variance of individuals.

Thus, derived from (35) and previous assumptions, the synergy effect derived from combinations among input units is a multiplicative effect over the variance given by discrete individuals. From (34) and (35):

$$ {\varepsilon_{max}}^2 = 3\phi^2 {\sigma_{ind}}^2 \quad (36) $$

Thus, for $x_{crs} = 1$, the estimated discrete variance from ${\sigma_{ind}}^2$ is:

$$ {\sigma_{ind}}^2 = \sum_{i=1}^{n} \left(r_{i\,(x_{crs},t)} - \mu_{w(x_{crs},t)}\right)^2 . p_{(i)} \quad (37) $$

where n is the number of individuals and $p_{(i)}$ is the probability for each individual.

Plugging (32) in (37) leads to:

$$ {\sigma_{ind}}^2 = \sum_{i=1}^{n} \frac{x_{crs_i}\left(r_{i\,(x_{crs},t)} - \mu_{w(x_{crs},t)}\right)^2}{\sum_{i=1}^{n} x_{crs_i}} = \sigma^2{}_{w(x_{crs},t)} $$

$$ {\sigma_{ind}}^2 = \sigma^2{}_{w(x_{crs},t)} $$

where $\sigma^2{}_{w(x_{crs},t)}$ is the weighted variance, x_{crs_i} is the weight given by the input value of each individual and $r_{i\,(x_{crs},t)}$ is the return $(\varepsilon / x_{crs_i})$ of each individual.

As ε is orthogonal to x_{crs}, $\mu_{w(x_{crs},t)} = 0$.

Replacing ${\sigma_{ind}}^2$ by $\sigma^2{}_{w(x_{crs},t)}$ in (36) leads to:

$$ {\varepsilon_{max}}^2 = 3\phi^2 \sigma^2{}_{w(x_{crs},t)} \quad (38) $$

17. Matching the Frontiers

Making equivalent both frontiers given by (33) and (38) leads to:

$$ 3\phi^2\sigma^2{}_w = ln(c)\, {\alpha_{(x_{crs})}}^2 . {\alpha_{(t)}}^2 $$

$$ {\alpha_{(x_{crs})}}^2 . {\alpha_{(t)}}^2 = \frac{3}{ln\ c} . \phi^2 . \sigma^2{}_{w(x_{crs},t)} \quad (39) $$

As every term from (39) is constant with exception from $\sigma^2{}_{w(x_{crs},t)}$, the conclusion is that the time-effect component is derived from the weighted variance of returns. Therefore:

$$ \alpha_{(t)} = \sigma^2{}_{w(x_{crs},t)} $$

$$ {\alpha_{(x_{crs})}}^2 = \frac{3}{ln\ c} . \phi^2 \quad (40) $$

One can realize that the affinity constant given by ${\alpha_{(x_{crs})}}^2$ is replaced by another constant term directly proportional to squared synergy (ϕ^2) and inversely proportional to critical input's logarithm. Therefore, the affinity is boosted by synergy and constrained by critical input. Therefore, from (35) and (40), squared synergy (ϕ^2) is an effect over the weighted variance of individuals derived from the affinity-synergy among their combinations, compared to the absence of mergers. Thus, higher values given by synergy amplify entropy (the efficient frontier). On the opposite, whether synergy is below 1, the combined effect from mergers shrinks the weighted variance of individuals – this is a defensive or nuclear system.

Another important conclusion follows. As stated by (29), the affinity coefficient $\alpha_{(x_{crs},t)}$ is the squared affinity constant ${\alpha_{(x_{crs})}}^2$ multiplied by the weighted variance of returns $\sigma^2{}_{w(x_{crs},t)}$. If the weighted variance of returns is zero, all the returns are the same. It means, in other words, that there is just one individual in the cluster and there is no envelopment. Without at least two different individuals, there is no affinity-synergy in combinations. Therefore, the affinity constant $\alpha_{(x_{crs})}$ must be zero because its numerator given by synergy is null whether there is only one individual; on the other hand, critical input is not zero because its value corresponds to the value of the single input itself. Hence, the equation (29), whose independent parameter is ${\sigma_{w(t)}}^2$, is a regression without intercept.

Therefore, the final complete equation is given by (41)*:*

$${\varepsilon_{max}}^2 = \frac{3}{ln\ c} . \phi^2 {\sigma^2}_{w(x_{crs},t)} (c\ ln\ c\ - x_{crs}\ ln(x_{crs}) - (c - x_{crs})\ ln(c - x_{crs})) \tag{41}$$

Equation (41) can be rewritten using the combinatorics approach, leading to the shorter version given by (1):

$${\varepsilon_{max}}^2 = \frac{3}{ln\ c} . \phi^2 {\sigma^2}_{w(x_{crs},t)} ln \binom{c}{x_{crs}} \tag{1}$$

18. Multiple Inputs with Constant Returns to Scale

As given by foundations from theory, the solution of the integral given by (24) is over the constant returns to scale (CRS). Hence, the model holds whether the mean of returns is kept constant regardless of the value from combinations. This is a straightforward application for the linear combination of inputs where the coefficients (partial derivatives) are constant (hyperplanes). Besides the regular linear combination of inputs, the current approach still holds when using constant elasticities transformed by a log operation (linear parameters).

In multiple inputs, an important topic to address concerns critical input. As discussed before, if any of the variables has infinite resources (unconstrained) compared to other inputs with scarce resources, these variables are regarded as not owing a critical input themselves - they must be kept out of the model. Finally, critical input for multiple variables is a linear combination of the constrained variables belonging to the set. Further explanations regarding multiple inputs are better detailed in Appendix C (linear combination of inputs).

19. Synergy's Discount Rate

One of the essential variables in valuation methods is the discount rate for cash flows. The mainstream methods based on cash flows rely on Beta CAPM for obtaining the cost of capital (ROE). However, the beta market risk (SP&500) is unstable and often provides negative values; furthermore, the beta in emerging markets does not present the significance obtained in developed markets. Hence, in some emerging countries, the discount rate that is effectively used is imported from the sectors derived from the American Market. Conversely, as discussed in this subsection, the Synergic Entropy provides a more stable and suitable discount rate.

At first, if one intends to measure the cost of capital and debt simultaneously, the inputs could be a linear combination of Equity and Debt, respectively. This approach is described below:

$$x = b_d x_d + b_e x_e \tag{42}$$

where d is debt and e is equity and the coefficients b_d/b_e are estimated under a OLS regression. However, this article keeps computing the Return on Assets (ROA) instead of splitting equity and debt into two separate variables. Appendix C (Multiple Inputs) better explains using multiple variables instead of one.

After choosing the proper input variables (or a single one), the second step is to compute the synergy constant for the whole sector. As the synergy constant is the effect over the weighted variance from individuals, it is possible to use (35) in order to define a new estimate for risk:

$$r_\phi = r_c + \ \phi . \sigma_{w(x,t)} \tag{43}$$

where r_ϕ is the discount rate for the cluster based on synergy (ϕ); $\sigma_{w(x,t)}$ is the weighted standard deviation of returns among companies in the same sector; and r_c is the minimum required return for each cluster (r_c), depending on stakeholders' decisions. In this article, the algorithm sets r_c to zero because the minimum required return depends on each investor's risk appetite.

As synergy is a longitudinal and cross-sectional parameter, its value should not change over time. On the other hand, $\sigma_{w(x,t)}$ could be subject to evolve through time; hence, the discount rate given by (43) presents a relative deviation (RSD) through time. It is important to notice that the weighted mean of returns (CRS) is not included in (43) for each cross-sectional, because it should not be misunderstood as a parameter for risk - its values can be either positive or negative and risk is evaluated around CRS axis. Therefore, rather than relying on a specific level of weighted returns (as CAPM), the risk is based on the synergic fluctuations around the several levels (weighted returns) provided by each cross-sectional.

The variance's fluctuation over time given by $\sigma_{w(x,t)}$ in (43) could undermine the synergy's discount rate. Fortunately, for financial statements, this weighted variance of returns when accrued for large windows (>=3 years) presents more stability. The reason probably relies on accounting performance, which makes some companies postpone or anticipate results to achieve best figures in financial statements; however, these deviations are commonly melted when the accounting figures for more than 2 years are joined. Conversely, it is worth remembering that the synergy parameter, rather than the synergy's discount rate, does not have an issue regarding the weighted variance of returns (the fluctuating values from $\sigma_{w(x,t)}$ are filtered out from the estimation).

Finally, if the subject is the company itself rather than the cluster it belongs, the cluster must be the company itself without its peers on the market. Therefore, all the input variables and output variables within a company must be evaluated, which could be the collection of its assets and their respective returns.

20. Methods and Data

The data used to evaluate the model are the financial statements (balance and income sheets) from the 11 sectors in America (2064 companies). However, the figures and tables are from the Consumer Cyclical Sector (324 companies). The 11 sectors used in this article, which classification criteria by company are provided in the dataset, are the following: Basic Materials; Business Services; Consumer Cyclical; Consumer Defensive; Energy; Financial Services; Healthcare; Industrials; Real State; Technology; and Utilities. The dataset containing all financial statements was retrieved from the following site belonging to the company SimFin: https://simfin.com/data/bulk. The SimFin's dataset and the files in R containing the algorithm are available for download at https://osf.io/ax4ny/?view_only=476431405b 9f4f9894200cd5d9ebef97.

The financial statements used are the quarterly incomes and balance sheets from January 2015 up to December 2019 (5 years). Due to the COVID-19 pandemic, there was significant volatility in results regarding 2020 and 2021; hence, these years are removed from the data to filter this effect. The variables used in the article are net incomes and total assets, which are the flow output and input, respectively. As the total assets' variable is used as an employment from equity and debt, there is no need for unlever net incomes. Therefore, the article uses returns on assets (ROA) instead of returns on equity (ROE) as the proxy for returns. Total assets are brought to future value (last quarter from data) by the Treasury Yield for ten years, using the code TNX from Yahoo Finance; and returns for each quarter are subtracted from the treasure yield accrued for three months - decreasing net income by the risk-free rate.

After retrieving net income, total assets and return for every company classified by sector, the following steps are required:

a) Establish the reference axes.

The reference axes must be built as explained in Subsection 2.9 and Appendix A, using $\boldsymbol{\varepsilon_{max}}$ and $\boldsymbol{x_{crs}}$ as the orthogonal axes for any operation that follows. The $\boldsymbol{x_{crs}}$ axis is rotated counterclockwise by the angle $\boldsymbol{tan\,\theta = \mu_{w(x_{crs},t)}}$ over $\boldsymbol{x}$ (cumulative inputs). As $\boldsymbol{\varepsilon_{max}}$ is orthogonal to $\boldsymbol{x_{crs}}$, $\boldsymbol{\varepsilon_{max}}$ is also rotated counterclockwise by the same angle over $\boldsymbol{Y_{(x,t)}}$ (cumulative outputs). The $\boldsymbol{\mu_{w(x_{crs},t)}}$ is the weighted average return, using $\boldsymbol{x_{crs\,i}}$ as the weight from individuals.

As data from the article is comprised of small returns on assets (ROA), one can use the approximation given in Appendix A:

$$if\ CRS\ \left(tan\theta = \mu_{weighted(x,t)}\right) \approx 0$$
$$x_{crs} = x\ ;\ \varepsilon_{max} = Y_{(x,t)} - x\mu_{w(x,t)}$$

b) Build an efficient envelopment.

The cumulative vectors made from joined outputs ($\boldsymbol{\varepsilon_{max}}$) and merged inputs ($\boldsymbol{x_{crs}}$) are derived from two laws of accrual: the first one accrues the outputs sorted from maximum to minimum returns, which is the maximum entropy: $+\boldsymbol{\varepsilon_{max}}$; and the second one accrues the outputs sorted from minimum to maximum returns, which is the minimum entropy: $-\boldsymbol{\varepsilon_{max}}$.

c) For each cross-sectional, estimate and store $\boldsymbol{\alpha}_{(\boldsymbol{x_{crs}},\boldsymbol{t})}$ given by $(\mathbf{29})$:

$$\boldsymbol{\varepsilon_{max}}^{2} = \boldsymbol{\alpha}_{(\boldsymbol{x_{crs}},\boldsymbol{t})}\, \boldsymbol{ln}\left(\frac{\boldsymbol{c}}{\boldsymbol{x_{crs}}}\right)$$

where:

$x_{crs} = x$ for small returns (data used in this article); and

$\boldsymbol{c = critical\ input}$: total of inputs at each cross-sectional, as described in Subsection Cross-sectional critical input.

d) For each cross-sectional, store the weighted variance of returns $\boldsymbol{\sigma^{2}}_{\boldsymbol{w}(\boldsymbol{x_{crs}},\boldsymbol{t})}$ from individuals and the cross-sectional critical input computed for that cross-sectional.

e) Estimate Synergy ($\boldsymbol{\phi}$) by regressing the affinity coefficient $\boldsymbol{\alpha}_{(\boldsymbol{x_{crs}},\boldsymbol{t})}$ for every cross-sectional depending on the weighted variance and critical input stored from last step. From $(\mathbf{39})$:

$$\alpha_{(x_{crs},t)} = \phi^{2}.\frac{3}{ln\ c}\sigma^{2}{}_{w(x_{crs},t)}$$

$$\sqrt{\alpha_{(x_{crs},t)}} = \phi\sqrt{\frac{3}{ln\ c}}\sigma_{w(x_{crs},t)} \qquad (44)$$

Finally, Synergy ($\boldsymbol{\phi}$) is the angular coefficient derived from $(\mathbf{44})$ after joining $\sqrt{\frac{\mathbf{3}}{\boldsymbol{ln\ c}}}$ and $\boldsymbol{\sigma}_{\boldsymbol{w}(\boldsymbol{x_{crs}},\boldsymbol{t})}$ in one independent term. As discussed above, cross-sectional critical input ($\boldsymbol{c}$) is a constant.

f) After estimating Synergy ($\boldsymbol{\phi}$), the Synergy's discount rate ($\boldsymbol{r_{\phi}}$) given by $(\mathbf{43})$ is computed for olling windows of 3 years (12 quarters of compounded returns).

21. Results

Table 1. Regressed Model for the Consumer Cyclical Sector

The $\mu_{(x,t)}$ and $\sigma_{w\ (x,t)}$ *are the weighted mean and weighted standard deviation from the entire cluster, respectively (weights are the values from inputs of individuals). The fitted model for the affinity coefficient* ($\sqrt{\alpha}_{(x,t)}$) *is from* (28)*, using Total Assets from each cross-sectional as an estimate for critical input. *** =* p-*value < 3.09E-287.* [1]*CAPM Beta* p-value *is computed over the financial statements using the whole Consumer Cyclical Sector against all market (2064 companies) with 1 year of rolling windows*

Section	U$ Billions Total Assets	%/quarter $\mu_{(t)}$	%/quarter $\sigma_{w\ (t)}$	Synergic Model			CAPM Beta[1] *p-value*
				$\sqrt{\alpha}_{(x,t)}$	*t*-stat	R^2	
2015 Q1	2152.04	0.90%	4.62%	4.17%***	69.5	89.1%	NA
2015 Q2	2286.76	0.99%	2.25%	3.84%***	86.6	92.6%	NA
2015 Q3	2271.68	0.82%	2.03%	3.36%***	181.4	98.2%	NA
2015 Q4	2351.23	1.01%	2.67%	4.30%***	153.0	97.4%	0.56
2016 Q1	2393.12	0.79%	1.52%	3.05%***	174.4	98.0%	0.47
2016 Q2	2432.34	0.91%	1.98%	3.10%***	283.7	99.2%	0.43
2016 Q3	2479.04	1.07%	1.74%	3.45%***	167.4	97.8%	0.87
2016 Q4	2509.09	0.96%	3.42%	5.19%***	99.0	93.8%	0.88
2017 Q1	2535.67	0.62%	1.59%	3.22%***	165.2	97.7%	0.78
2017 Q2	2627.52	0.55%	2.01%	3.77%***	217.0	98.6%	0.79
2017 Q3	2674.66	0.62%	2.64%	4.81%***	245.7	98.9%	0.85
2017 Q4	2696.81	0.72%	3.26%	6.71%***	252.1	99.0%	0.90

2018 Q1	2710.14	0.70%	2.20%	4.52%***	140.2	96.8%	0.91
2018 Q2	2620.27	0.80%	2.00%	4.04%***	146.2	97.0%	0.93
2018 Q3	2654.40	0.81%	2.43%	4.27%***	254.0	99.0%	0.93
2018 Q4	2673.39	0.55%	2.51%	4.46%***	238.5	98.9%	0.43
2019 Q1	2869.13	0.50%	1.68%	3.64%***	164.4	97.7%	0.52
2019 Q2	3022.82	0.90%	3.45%	5.55%***	154.4	97.3%	0.46
2019 Q3	2970.58	0.75%	2.13%	4.14%***	189.3	98.2%	0.51
2019 Q4	3015.25	0.55%	2.33%	4.64%***	221.4	98.7%	0.55

Source: compiled by the author.

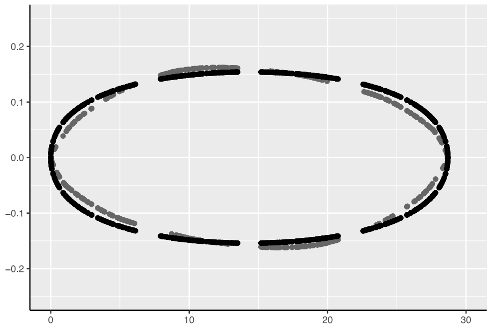

Figure 3a: 1Q 2019. $\sigma_{w=1.68\%}$; $c = 2869$

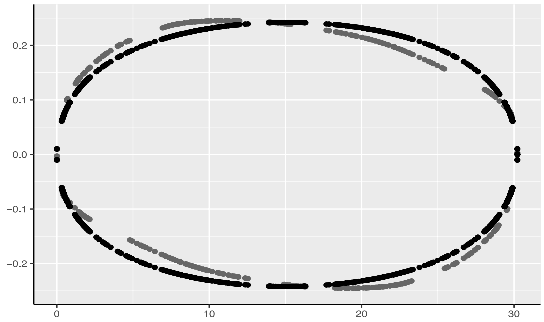

Figure 3b: 2Q 2019. $\sigma_{w=3.45\%}$; $c = 3023$

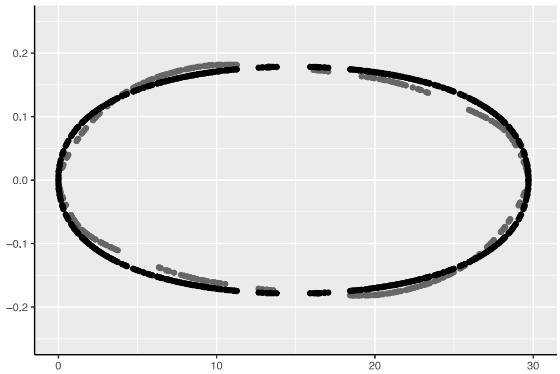

Figure 3c: 3Q 2019. $\sigma_{w=2.13\%}$; $c = 2971$

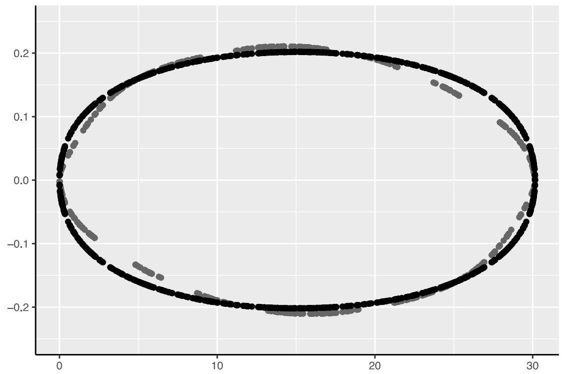

Figure 3d: 4Q 2019. $\sigma_{w=2.33\%}$; $c = 3015$

Figure 3. Longitudinal and Cross-Sectional Effects

Source: compiled by the author.

X axis (Input) is the Total Asset value from individuals or the sum from combined assets of joined companies. Y axis (Output) is the Income Generated by the maximum and minimum ordered returns (gray ellipse). The black ellipse is the fitted model from (28). $\sigma_{w(t)}$ is the weighted standard deviation from input units at each cross-sectional, given by Table 1. c is the critical input computed as the sum of every input unit at the cross-sectional (step iii in Method).

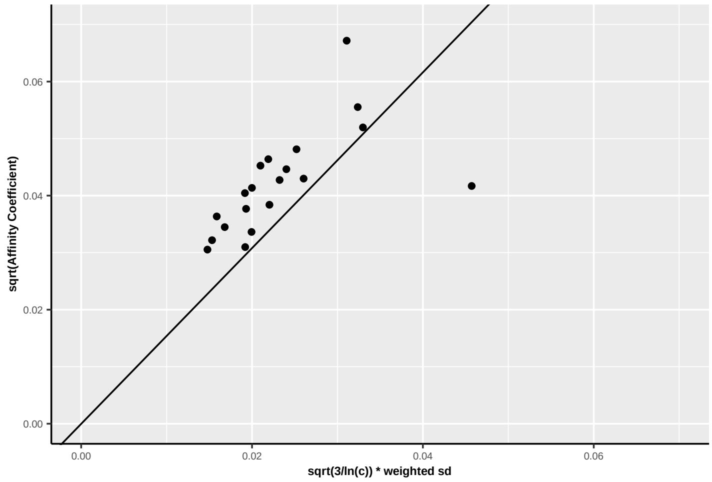

Figure 4. Synergy Model for Consumer Cyclical Sector

Source: compiled by the author.

Y Axis is the square root of the Affinity Coefficient ($\sqrt{\alpha_{(x,t)}}$) from (28) shown in Table 1; X Axis is the weighted standard deviation of returns ($\sigma_{w(x,t)}$) multiplied by the square root of the constant ($3/ln\, c$). Synergy (ϕ) is the angular coefficient given by (44).

Table 2. Synergy Discount Rate (r_ϕ) for Consumer Cyclical in rolling windows

Synergy (ϕ) is the angular coefficient computed by (44). Synergy's discount rate (r_ϕ) is computed for rolling windows of 3 years multiplying Synergy (ϕ) by the weighted standard deviation of returns using *(43)* ; μ (r_ϕ) is the mean *of* r_ϕ through time and *RSD is the relative standard deviation given by the variance of* r_ϕ *through time given by* $\sigma_{r_\phi} / \mu_{r_\phi}$. The weighted return in the second column is the CRS axis for the rolling window, but it is not part of the *Synergy's* discount rate (r_ϕ).

Section		Synergy ϕ = 1.54	
	Weighted Return %/year	Weighted Standard Deviation %/year	Synergy's rate r_ϕ %/year
2015 Q1 - 2017 Q4	3.22%	5.24%	8.07%
2015 Q2 - 2018 Q1	3.20%	5.69%	8.77%
2015 Q3 - 2018 Q2	3.11%	5.91%	9.10%
2015 Q4 - 2018 Q3	3.13%	5.95%	9.17%
2016 Q1 - 2018 Q4	3.05%	5.92%	9.13%
2016 Q2 - 2019 Q1	3.13%	6.05%	9.32%
2016 Q3 - 2019 Q2	3.29%	6.14%	9.45%
2016 Q4 - 2019 Q3	3.06%	6.23%	9.60%
2017 Q1 - 2019 Q4	2.92%	6.14%	9.46%
		$\mu\ r_\phi$	9.12%
		RSD r_ϕ	5.09%

Source: compiled by the author.

Table 3. Synergy (ϕ) and r_ϕ for All Sectors

Synergy is the angular coefficient estimated by (44). *** = p-*value* < 0.001 ; ** = p-*value* < 0.01. Synergy's discount rate (%/year) is computed using *(43) in rolling windows of 3 years. RSD is the relative standard deviation given by the variance of* r_ϕ *through time given by* $\sigma_{r_\phi} / \mu_{r_\phi}$.[1]Cheniere Energy Partners (CQH) is removed from Energy Cluster due to abnormal results over 1000% in 2017Q4, 2018Q1 and 2018Q2.

Sector	t-stat	R^2	Synergy ϕ	Synergy's rate r_ϕ	RSD r_ϕ
Basic Materials	8.02	75.98%	0.78 ***	3.57%	5.39%
Business Services	7.42	72.97%	1.10 ***	4.41%	7.48%

Table 3 (cont.). Synergy (ϕ) and r_{ϕ} for All Sectors

Consumer Cyclical	7.23	71.96%	1.54 ***	9.12%	5.09%
Consumer Defensive	9.99	83.16%	1.21 ***	6.57%	6.83%
Energy[1]	8.48	78.01%	1.09 ***	5.42%	15.39%
Financial Services	30.19	97.85%	1.20 ***	6.83%	2.93%
Healthcare	0.464	1.12%	0.01	NA	NA
Industrials	2.77	25.04%	0.98 **	4.49%	4.35%
Real State	8.63	78.60%	0.52 ***	1.54%	2.03%
Technology	15.84	92.59%	2.09 ***	15.51%	7.22%
Utilities	6.18	65.03%	0.84 ***	1.56%	10.99%

Source: compiled by the author.

Discussion

The empirical results show clear frontiers for maximum entropy, signaling a driven phenomenon, rather than a pure stochastic combination among companies. Furthermore, some envelopments are not entirely symmetrical due to imperfections related to cluster definition and other issues derived from accounting and publishing procedures (as seen in Figure 3).

As expected by the current approach, the cross-sectionals regressions are significant for every time slice (Table 1). In contrast, the typical beta for financial statements did not present any significance (tough beta from the stock market is significant as widely covered by literature). The longitudinal regressions prove that the affinity constant (α_x) , which derives synergy (ϕ), is significant and stable for every sector (Table 2). The longitudinal result in Table 3 is plotted for Consumer Cyclical Sector in Figure 4. The outlier in Figure 4 is due to an abnormal income loss from Amazon Inc. and Hertz Co. in 2015Q1, two of the major companies in the sector. All the regressions are made without the intercept, for which reasons are presented as follows (topic briefly discussed in theory). At first, the cross-sectionals regressions cannot have an intercept because output deviation must be zero whether the input is zero. A similar approach is valid for longitudinal regressions: whether the weighted variance of returns is zero, there is no envelopment because all outputs are located at CRS; hence, the output deviations are zero. Therefore, whether squared deviations and variance are zero and inputs are not zero (there is at least one input unit to compute the weighted variance), the null term in (28) can only be derived from the affinity coefficient ($\alpha_{(x,t)}$).

Comparing the sections in Figure 3, it is possible to understand better the longitudinal and cross-sectional effects of the deducted phenomenon. For example, in section 3Q2019 and section 4Q2019, which have almost the same critical input (U$ 3 trillion), the difference regarding the shape between them is given by the longitudinal volatility of each section ($\sigma_{w(t)}$). Now, suppose one compares the shapes from 1Q2019 and 2Q2019. In that case, there are two size effects in entropy: the first effect is the cross-sectional one generated from the greater inputs (x) and critical input (c), deriving a proportional increase in squared entropy given by linearity (homogeneity in Appendix D); and the second effect is the longitudinal one provided by the greater ($\sigma_{w(t)}$), which squared value also derives a proportional increase in squared entropy.

The synergy, as discussed before, is the effect derived from combined inputs with affinity compared to the effect without mergers. Therefore, it is expected that occurs greater values for riskier sectors and lower values for defensive sectors, which common belief is confirmed by Technology (2.09) and Real State (0.52), respectively. One important threshold for synergy is the unitary value: clusters with synergy beyond one generate deviations above the weighted variance from individuals when there is combination; conversely, clusters with synergy below 1 means that their affinity is a defensive-nuclear one, decreasing the weighted variance of returns when inputs are merged. If synergy is below one, one could claim that it is not synergy, because the combined effect is lower than the sum of the individual effects; however, this anti-synergy behavior is a significant phenomenon that shrinks volatility and should not be misunderstood with clusters that do not present any affinity among its members.

Regarding clusters without affinity or cohesion, the synergy among their input units should have insignificant levels. This fact truly happened for the Healthcare Sector, which does not present a significant value for synergy. The reasons for not achieving significance are related to the proper classification criteria of the cluster. In Healthcare, for example, there are heterogeneous companies without strong commonalities: some companies manufacture drugs (e.g.: Gilead, Pfizer and Johnson & Johnson); others provide medical devices and equipment (e.g.: Medtronic, Becton Dickinson and Bard); some companies offer medical diagnostics and research (e.g: Danaher and Thermo); and finally, some companies are health care providers (e.g: UnitedHealth, CVS and HCA Holdings). As the Healthcare sector is highly heterogeneous, the best approach is to compute synergy for each subsector. On the other hand, one could argue that there are other heterogeneous sectors in the economy, like Consumer Cyclical; however, the effect derived from discretionary consumption in this sector is more dominant than the opposite one given by the heterogeneity between its companies.

On the other hand, the Finance Sector is the one that presents the highest significance for synergy. Due to the liquidity and correlation of financial assets compared to real assets, the mergers among financial assets are more efficient and synergic than other sectors with real assets. At last, the Industrial Sector is the cluster that presents the lowest significance for synergy (p-*value* < 0.01) compared to the values achieved by other sectors (p-*value* < 0.001); as discussed above, whether one changes the classification criteria in proper subsectors, the significance must get higher.

The rationale that computes the synergy's discount rate (r_ϕ) is shown in Table 2, which is the one built for Consumer Cyclical Sector. The rate r_ϕ is not constant through time due to the fluctuating weighted variance of returns, as explained in the 2.13 subsection. The final rate r_ϕ is the mean of the rates computed for each cross-sectional, and RSD is the relative squared deviation of these values. As shown in Table 3, r_ϕ is also computed for all sectors, presenting stable values with low RSD. The magnitude of r_ϕ regarding each sector seems suitable whether one picks the rule of thumb that gives market risk a value of 6% - the finance sector, which is the one most alike the stock market's behavior, presents r_ϕ very similar to the stock's market risk: 6.83%. Technology and Consumer Cyclical, as expected, are the ones that present the riskiest behavior, achieving 15.51% and 9.12%, respectively. Conversely, there are risk-averse sectors, such as Utilities and Real State with rates around 1.5% yearly over the Treasury yield for ten years.

Other input variables rather than the one used in this article, which is "total assets", can be eventually evaluated as seen in Subsection 2.12. and Appendix C (linear combination of inputs). For example, other inputs such as debt (leverage), equity, liquidity ratio, working capital, labor and operational profitability can be evaluated together compounding input x.

Conclusion

As discussed, the theory aims to design a semi-deterministic approach for volatility (entropy), whose objective calls into question the usual stochastic behavior found in the literature. The foundation for this objective is based on the causal relationship between joined inputs and merged outputs, which combined output may be bounded inside an envelopment. This envelopment that removes the stochastic behavior can be acknowledged as similar to the efficient frontier given by DEA, shown in the seminal article from Charnes et al. (1978).

The boundaries of the envelopment in the current approach are given by the maximum and minimum ordered derivatives (returns), rather than the optimization procedure (objective function) found in DEA. Furthermore, the model for the envelopment is built under an approach based on time, combinations and novel parameters, such as critical input (scarce resources), synergy and the first efficient constituent. Some of these parameters and variables are well known in Economics, but the current theory establishes the relationship between them in a single model over time; even time itself is neglected in most economic theories, because these approaches are hardly interested in the longitudinal effect.

A theoretical achievement is the Critical Input variable embedded in the model. This is the name theory gives for the concept related to scarce resources. It is widely accepted that most of the inputs in a system are limited by constrained resources, such as the maximum number of skilled engineers, current oil reserves, available capital to invest and further on. Physical systems, which can also lead to further empirical research in their respective fields, comprise a limited number of forests, a limited number of available matter or even the maximum number of people infected by a virus. Critical input can be trivial for some systems but unknown for others. Critical input is the cause that makes the curve from the novel theory become an ellipse rather than a common logarithmic curve. Besides critical input, another point of interest is the first efficient constituent. This is the first real point in the ellipse that entropy is maximum for a given level of combinations – this is the minimum number of constituents required to ignite the cluster.

The strength of the interactions among constituents in a cluster is defined by a constant named synergy, in which higher values amplify entropy (volatility) and lower values below one shrink entropy; in the case that synergy is below one, the sector has a defensive (nuclear) behavior, which could lead synergy to be named as anti-synergy. The puzzle for solving synergy is derived from its definition: a comparison between the effect without mergers against the boosted effect provided by the interactions (combinations) among its constituents. Therefore, the deducted equation has the effect of interactions embedded in its formulation, which is neglected in most mainstream approaches for entropy until Tsallis (1988). However, rather than the statistical approach in Tsallis (1988), the current theory is based on economical-physical reasoning established a priori whose foundations are the variables and parameters discussed above. Furthermore, the cross-sectional special case of Synergic Entropy is the entropy given by Gibbs-Boltzmann and Shannon (Appendix B). The general relationship between the output flow and its input can be applied in other physical systems, such as *hadronization*. As hadronization is conceived by an influx of energy, the input flow ($hadrons/time$) can be established as given by (5). Moreover, the fundamental particle of a system (the first efficient constituent) can be indirectly estimated whether one has already computed the other parameters, which precise procedures are subject to further research.

The results presented in this article for economic clusters have shown that synergy is stable, significant and constant through time. These synergy advantages confirm its adequacy as a good proxy for risk, whose values above one mean risky sectors and values below one corresponds to defensive-nuclear sectors (which anti-synergy shrinks volatility). Therefore, based on its significance and adequacy, synergy is the core of a novel approach for valuation that does not depend on the return of overall market risk, like SP&500. Rather than relying on a market index that can be either positive or negative, the new discount rate is obtained by multiplying synergy by the weighted standard deviation of returns. This practical formula for synergy's discount rate proves to be adherent to the perceived risk regarding each sector. Besides that, the subjectivity that exists in selecting portfolios in CAPM (Sharpe (1964), Lintner (1965) and Mossin (1966)) is removed by synergy, because the significance of the selection is tested against the proper definition of the cluster ($\phi\ p-value$) .

Optimization procedures can classify systems in future research, whose tasks can be done by fine-tuning the significance of synergy. Further research can also investigate the survival of a system, in which ignition is played by the first efficient constituent (fundamental unit), and the termination is given by critical input.

Acknowledgments

This research was supported by the Central Bank of Brazil, which allowed full-time dedication to conducting the research from March 2018 until December 2021. The views expressed in this work are those of the author and not necessarily those of the Central Bank of Brazil or its members.

References

1. Arrow, K.J., Chenery, H.B., Minhas, B.S. & Solow, R.M. (1961). Capital Labour Substitution and Economic Efficiency. *Review of Economics and Statistics*, 63, 225-250. [CrossRef]

2. Basso, A., Casarin, F., & Funari, S. (2017). How well is the museum performing? A joint use of DEA and BSC to measure the performance of museums. *Omega.* [CrossRef]

3. Beck, C. & Cohen, E.G.D (2003). Superstatistics. *Physica*, 322, 267. [CrossRef]

4. Ben Lahouel, B., Taleb, L., Ben Zaied, Y. *et al.* (2022). Financial stability, liquidity risk and income diversification: evidence from European banks using the CAMELS–DEA approach. *Ann Oper Res.* [Crossref]

5. Charnes A., W. W. Cooper & Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research.* 2, 429-444. [CrossRef]

6. Chen, C. C. (2017). Measuring departmental and overall regional performance: applying the multi-activity DEA model to Taiwan's cities/counties. *Omega*, 67, 60–80. [Crossref]

7. Christopoulos, A. G., Dokas, I. G., Katsimardou, S., & Vlachogiannatos, K. (2015). Investigation of the relative efficiency for the Greek listed firms of the construction sector based on two DEA approaches for the period 2006–2012. *Operational Research, 16(3),* 423–444. [CrossRef]

8. Cobb, C.W. & Douglas, P.H. (1928) A Theory of Production. *American Economics Review, 18*(1), 139-165. [Link]

9. Diewert, W.E. (1971). An Application of the Shepherd Duality Theorem: A Generalized Leontief Production Function. *The Journal of Political Economy*, *79*(3), 481-507. [CrossRef]

10. Färe, R., & Grosskopf, S. (1997). Intertemporal Production Frontiers: With Dynamic DEA. *Journal of the Operational Research Society, 48*(6), 656–656. [CrossRef]

11. Farrell M.J (1957). The measurement of productive eficiency. *J R Stat Soc 120:253–281*. [CrossRef]

12. Feng, J., Geng, L., Liu, H. *et al* (2022). Efficiency evaluation of the high-tech industry chain with a two-stage data envelopment analysis approach. *Oper Manag Res*. [CrossRef]

13. Fukuyama, H., Matousek, R., & Tzeremes, N. G. (2020). A Nerlovian cost inefficiency two-stage DEA model for modeling banks' production process: Evidence from the Turkish banking system. *Omega, 102198.* [CrossRef]

14. Kottas, A. T., Bozoudis, M. N., & Madas, M. A. (2019). Turbofan Aero-Engine Efficiency Evaluation: An Integrated Approach Using VSBM Two-Stage Network DEA. *Omega, 102167.* [CrossRef]

15. Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics. 47*(1), 13–37. [CrossRef]

16. Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrics*, *34*(4), 768-783. [CrossRef]

17. Ngo, T., Tsui, K.W.H (2021). Estimating the confidence intervals for DEA efficiency scores of Asia-Pacific airlines. *Operational Research International Journal*. [CrossRef]

18. Silva, T. C.; Tabak, B. M.; Cajueiro, D. O.; Dias & M. V. B. (2017). A comparison of DEA and SFA using micro- and macro-level perspectives: Efficiency of Chinese local banks. *Physica A: Statistical Mechanics and Its Applications*, *469*, 216–223. [CrossRef]

19. Sattin, F (2006). Bayesian Approach to Superstatistics. *The European Physical Journal B. B49 219*. [CrossRef]

20. Shannon, C.E (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27, 379-423, 623-656. [CrossRef]

21. Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance. 19*(3), 425–442. [CrossRef]

22. Sharpe, W. F. (1994). The Sharp Ratio. *The Journal of Portfolio Management 21*(1), 49-58. [CrossRef]

23. Tone, K and Tsutsui, M. (2014). Dynamic DEA with network structure: A slacks-based measure approach. *Omega*, *42*, 124-131. [CrossRef]

24. Tone, K and Tsutsui, M. (2010). Dynamic DEA: A slacks-based measure approach. *Omega*, *38*, 145-156. [CrossRef]

25. Tsallis, C (1988). Possible generalization of Boltzmann-Gibbs statistics. *J Statistical Physics*, *52*, 479–487. [CrossRef]

26. Wicksteed, P.H. (1894). An Essay on the Co-ordination of the Laws of Distribution. *Macmillan & Co., London*. [Link]

27. Wu, H., Lv, K., Liang, L., & Hu, H. (2017). Measuring performance of sustainable manufacturing with recyclable wastes: A case from China's iron and steel industry. *Omega*, *66*, 38-47. [CrossRef]

28. Yang, Guo-liang; Fukuyama, Hirofumi & Chen, Kun (2018). Investigating the Regional Sustainable Performance of the Chinese Real Estate Industry: A Slack-based DEA Approach. *Omega, S0305048317311374*. [CrossRef]

29. Zhou, Xiaoyang; Xu, Zhongwen; Chai, Jian; Yao, Liming; Wang, Shouyang & Lev, Benjamin (2018). Efficiency evaluation for banking systems under uncertainty: A multi-period three-stage DEA model. *Omega*, *S0305048317313063*. [CrossRef]

30. Zhu, J., & Wei, D. (2021). Analysis of stock market based on visibility graph and structure entropy. *Physica A: Statistical Mechanics and Its Applications, 576, 126036*. [CrossRef]

31. Zýková, P (2022). The overall efficiency of the dynamic DEA models. *Central Europenean Journal of Operations Research, 30*(2), 495–506. [CrossRef]

Appendix A: Rotation of Reference Axes

A.1. General

As required by theory, input (x) and output (Y_f) axes must be rotated counterclockwise by an angle θ until the input axis gets aligned with the CRS (Constant Returns to Scale) axis. Therefore, it is necessary to determine the rotation angle θ, which defines the CRS. At first, it is worth remembering that the inputs and outputs are cumulative figures in both axes. The first point of the line that defines the CRS is (0,0) because cumulative outputs do not emerge when cumulative inputs remain without input units. The second point which finally establishes the line of the cumulative CRS is at critical input $(c, Y_{f(c)})$. One could argue why the second point of CRS is defined as critical input. The answer is that theory establishes critical input at the point where all input units are joined with only one possible outcome: the total outputs combined. As there is only one possible output, the cumulative CRS must be located at this point without any other possibility around critical input. It is important to notice that critical input for multiple constrained inputs is shown in Appendix C (linear combination of inputs). Slope between (0,0) and $(c, Y_{f(c)})$:

$$tan\theta = \frac{Y_{f(c)}}{c} \qquad (A.1)$$

As explained before, critical input can be known or unknown. If critical input is unknown, $tan\theta$ can be estimated by retrieving the definition of CRS – the constant mean of returns. In this case, one must compute the total outputs and inputs from known individuals in the cluster to estimate the mean of returns. This is the same as estimating the weighted mean of returns, which must have approximately the same slope that connects (0,0) and $(c, Y_{f(c)})$.

Thus, from the above statements:

$$tan\theta = \frac{\sum_{i=0}^{i=n} Y_{f\ all}}{\sum_{i=0}^{i=n} x_i}$$

$$tan\theta = \frac{\sum_{i=0}^{i=n} r_i . x_i}{\sum_{i=0}^{i=n} x_i} \ ; \ r_i = \frac{Y_{f\ i}}{x_i} \qquad (A.2)$$

$$tan\theta = \mu_w \qquad (A.3)$$

where r_i is the return of each input unit, n is the number of companies in the cluster and μ_w is the weighted return. Rather than using the cumulative input in the x axis, one can use the average output given by CRS in the x axis. In this case, as seen in Multiple Inputs in Appendix C, $tan\theta = 1$.

Being defined the rotation angle, it is possible to present the transformation matrix for a xy plane which will be rotated to $x'y'$. The $x'y'$ plane is formed by orthogonal vectors and x' is a vector rotated $arc\ tan\theta$ counterclockwise from x. The $x'y'$ plane is the solution given by (24) where $x' = x_{crs}$ and $y' = \varepsilon_{max}$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$x_{crs} = x cos\theta + y sin\theta \qquad (A.4)$$

$$\varepsilon_{max} = -x sin\theta + y cos\theta \qquad (A.5)$$

A.2. Rotation for Small Returns

For systems having small returns, the return $(\mu_w = tan\theta)$ is commonly a small number compared to cumulative inputs. Thus, some properties can be used:

$$cos\theta \approx 1 \qquad (A.6)$$

$$sin\theta \approx tan\theta \quad (A.7)$$

It is important to emphasize that this approach is not valid for systems with multiple linear combinations of inputs, nor systems using elasticities as variables.

Therefore, if one applies $(A.6)$ and $(A.7)$ to $(A.4)$, one gets the first equation for transformed axis x_{crs}:

$$x_{crs} = x.1 + y(small\ to\ x').sin\theta(small\ to\ x') \quad (A.8)$$

The second term of $(A.8)$ tends to minimum significance, as both y and $sin\ \theta$ are small numbers compared to x.

Therefore, one gets:

$$x_{crs} = x \quad (A.9)$$

Making the same approach for Y axis in (A.5), one gets:

$$\varepsilon_{max} = y - x\mu_w \quad (A.10)$$

Finally, plugging $(A.9)$ and $(A.10)$ in (24), one gets equation (25) with rotated axes:

$$\left(Y_{max(x,t)} - x\mu_{(x,t)}\right)^2 = \alpha_{(x,t)}(c\ ln\ c\ - x\ ln(x) - (c - x)\ ln(c - x)) \quad (25)$$

It is important to notice that the approximation given by (25) is only valid for small returns. However, when the scale of outputs is similar to the one given by inputs, the rotation matrix must be applied – this is the case for the linear combination of inputs and elasticities.

Appendix B: Shannon's Entropy

In this Appendix the commonalities between the current approach and Shannon's Entropy are presented. Likewise, the same comparison can be applied to Gibbs-Boltzmann's equation due to the commonalities between their mathematical formulation. In (B.1), Shannon's entropy seminal equation is reproduced:

$$H\ =\ -\sum_{i=1}^{n} p_i\ log_b x\ (B.1)$$

Assuming that:

$$\lim_{p \to 0^+} plog(p) = 0\ (B.2)$$

The assumption in $(B.2)$ was the same previously taken to define the boundaries conditions around $\alpha x log(x)$ and $\alpha(c - x)log(c - x)$ (See Subsection 2.8).

Shannon's entropy was conceived for multiple different states of x in $(B.1)$. It means, for example, that x could be any of the 26 letters (states) of an English word in a transmitted message, in which some letters have more probability of occurring than others. However, Shannon's entropy in this article can be acknowledged as having only two states: the first one is the certainty (+cl) about the unit (particle) being part of the cluster; and the second state is the certainty about the particle being part of another cluster (−cl), not belonging to the current one. Thus, equation $(B.1)$, written for two possible states, that are the cluster (+cl) and the anti-cluster one (−cl), becomes:

$$H\ =\ -p_{i=1}\ log_2\ p_{i=+\mathrm{cl}}\ - p_{i=-1}\ log_2\ p_{i=-\mathrm{cl}}$$

If one changes the logarithm's base in the above equation, one gets a new equation rewritten for natural logarithm:

$$H = \frac{1}{log2}(-p_{i=1} log p_{i=+\mathrm{cl}} - p_{i=-1} log p_{i=-\mathrm{cl}}) \quad (B.3)$$

As the two possible events are complementary, the probabilities of each one can be written as follows:

$$p_{i=\mathrm{cl}} = p \quad (B.4)$$

$$p_{i=-\mathrm{cl}} = 1 - p \quad (B.5)$$

Plugging $(B.4)$ and $(B.5)$ in $(B.3)$ one gets:

$$H = \frac{1}{log2}\left(-plogp - (1-p)log(1-p)\right) \quad (\mathrm{B}.6)$$

Equation $(B.6)$ is the entropy binary function for two possible complementary states. As the first state is the cluster force, one can realize that the greater the p the greater the certainty of that particle belongs to the current cluster. The scalar $(\frac{1}{log2})$ was necessary to make a base adjustment, though other adjustments might be necessary to calibrate the correct entropy measure. Thus, equation $(B.6)$ can be rewritten again:

$$H = constant * \left(-plogp - (1-p)log(1-p)\right) \quad (\mathrm{B}.7)$$

Now, it is time to adapt and adjust the novel approach for the cross-sectional given by (24). The first adjustment is to transform input in a probabilistic function, having critical input as the possible threshold for input. When the input is equal to critical input, the chance of that level of the particle (input) belonging to a certain cluster is maximum. The reason for maximum certainty at critical input is that one can only be entirely sure whether a particle actually belongs to a cluster or not whether the level of that particle is the entire cluster itself, which is the value of critical input. Besides critical input, the other constant that defines a cluster is synergy – however, synergy remains constant for the cross-sectional given by (24). Thus, if one takes input and transforms its scale in a unit-based normalization allowed by linearity given by $(D.4)$, one gets:

$$x_{crs}{}^{*} = x_{crs}/c \quad (B.8)$$

Thus,

$$x^{*} = 1, for\ x = c$$

Applying unit-normalization given by $(B.8)$ in (24), one gets:

$$\varepsilon_{max}{}^{2} = \alpha_{(x_{crs},t)} 1\, ln\, 1 - \alpha_{(x_{crs},t)} x_{crs}{}^{*}\, ln(x_{crs}{}^{*}) - \alpha_{(x_{crs},t)}(1 - x_{crs}{}^{*})\, ln(1 - x_{crs}{}^{*})$$

As $x_{crs}{}^{*}$ is the chance of achieving critical point as input increases, ranging from zero to 1, one can define $x_{crs}{}^{*}$as an input's (particle) density probability. Therefore, one gets:

$$x_{crs}{}^{*} = p_x$$

$$\varepsilon_{max}{}^{2} = \alpha_{(x_{crs},t)}\left(-p_x\, ln(p_x) - (1-p_x) ln(1-p_x)\right) \quad (B.9)$$

When equations $(B.9)$ and $(B.7)$ are compared, one realizes that both approaches are equivalent: $\varepsilon_{eff}{}^{2}$ can be acknowledged as a measure for entropy; the constant in $(B.7)$ is the affinity coefficient given by $\alpha_{(x_{crs},t)}$; and the other terms of the equation are equivalent.

Appendix C: Linear Combination of Inputs

C.1. Simple Linear Combination

As discussed in Subsection 2.12, the theory holds if the input defined by x is a linear combination of several inputs, as seen below:

$$x = b_1 x_1 + b_2 x_2 + \ldots b_n x_n \quad (C.1)$$

When $x = c$, it comes:

$$c = b_1 c_1 + b_2 c_2 + \ldots b_n c_n \quad (C.2)$$

where n is the number of constrained parameters in input.

The unconstrained variables are out of the set because they do not provide causality relations between their magnitude and the variance given by the current approach. For example, suppose one uses an unconstrained variable as a single parameter. In that case, this approach gives the same result as the one provided by a pure CRS approach with random errors, an OLS with homoscedasticity. As the companies or input units do not need to compete themselves for inputs, because there is plenty of them, there is no difference regarding efficiency among them – the envelopment does not exist for this single unconstrained variable. In this case, there is no theoretical variance driven by a phenomenon – however, due to random errors, a constant variance derived from OLS might happen. In order to avoid this constant variance at critical input, whose variance must be zero, the unconstrained variables must be kept out from the variance's prediction model. Hence, the unconstrained variables must be subtracted from the output flow, keeping only the constrained ones.

As the single input x is compounded by a linear combination of several inputs $x_1 ... x_n$, its coefficients $(b_1, b_{2,} ... b_n)$ can be estimated by an OLS multiple linear regression, as given by $(C.3)$ below. As defined before, this is a regression without intercept because there is no output when zero input units are available to combine.

Therefore, the coefficients $b_1, b_{2,} ... b_n$ are estimated by:

$$\hat{Y}_f = b_1 x_1 + b_2 x_2 + \cdots b_n x_n \quad (C.3)$$

$$x = \hat{Y}_f \quad (C.4)$$

One can argue that replacing x by $\hat{Y}_f$ in the input axis destroys the relationship between output and input. However, instead of the CRS itself, the novel theory provides the variance (entropy) over the CRS. Thus, the reasoning that predicts the fluctuations above and below CRS is not changed whether one replaces x by $\hat{Y}_f$.

As the axes are formed by cumulative values, x is replaced by the cumulative values from the individuals $(\hat{Y}_f)$ in $(C.4)$ and Y_f remains the cumulative values from outputs. Thus, equation (3) for the ordered returns that build the envelopment becomes:

$$\frac{dY_{f(i)}}{d\hat{Y}_{f(i)}} = r_i \quad (C.5)$$

Therefore, the derivatives built in ascending and descending order can be calculated using the same definition provided by (3) mirrored for more dimensions as stated by $(C.5)$.

Therefore, the axis x is formed by cumulative fitted values from output $(\hat{Y}_f)$ and the y axis is the cumulative actual data from output (Y_f). The envelopment for this data will show an ellipse around a CRS which slope is 45^0 $(y = x)$. This 45^0 CRS is derived from the line that connects the points $(0{,}0)$ and $(c, Y_{f(c)})$, which was explained in Appendix A. Therefore, $tan\theta$ can be derived directly from $(A.1)$ and $(C.2)$:

$$tan\theta = \frac{Y_{f(c)}}{c}$$

$$tan\theta = \frac{\hat{Y}_{f(c)}}{\sum b_1 c_1 + \sum b_2 c_2 + ... \sum b_n c_n}$$

$$tan\theta = \frac{\hat{Y}_{f(c)}}{\hat{Y}_{f(c)}} = 1 \quad (C.6)$$

where $Y_{f(c)} = \hat{Y}_{f(c)}$ is derived from the definition of critical input as detailed in Appendix A. At critical input, there are no deviations between the envelopment $(Y_{f(c)})$ and the CRS $(\hat{Y}_{f(c)})$.

C.2 CES Production Function and Other Output Functions

The CES production function is a first-degree homogeneous equation. Therefore, the CES production function from Arrow et al. (1961) can be used as a CRS baseline for the current approach. Furthermore, the current approach holds for other functions whether a log transformation gives a linear combination of elasticities. Therefore, output functions like Cobb-Douglas (1927) can be used as CRS baselines in (1) when applying the proper log transformation.

Appendix D: Ellipse Method

This section presents a method for estimating longitudinal critical input when its value is unknown, and its variance is steady throughout time (better detailed in 2.3.1). Before describing the ellipse method, some important properties and definitions must be addressed: the maximum point of deviation and linearity.

D.1 Maximum Point of Deviation

The maximum deviation is another practical assessment given by the deducted equation, allowing the maximum output estimation. This value is deducted below to compute the y axis required from the Ellipse Method. In order to get the maximum point, the first derivative of (24) is shown below:

$$\varepsilon_{max} = \pm\sqrt{\alpha_{(x_{crs},t)}(c\,ln\,c\; - x_{crs}\,ln(x_{crs}) - (c - x_{crs})\,ln(c - x_{crs}))}$$

$$\frac{d\varepsilon_{max}}{dt} = \pm\frac{\alpha_{(x_{crs},t)}\,(ln(c - x_{crs})\; - ln(x_{crs}))}{\sqrt{\alpha_{(x_{crs},t)}(c\,ln\,c\; - x_{crs}\,ln(x_{crs}) - (c - x_{crs})\,ln(c - x_{crs}))}} \quad (\text{D}.1)$$

The root for this equation might be at half of critical input, because the curve is symmetric. At $\frac{c}{2}$, the numerator of equation $(S.1)$ turns out to be zero, proving that this is actually the root. Therefore,

$$Roots(\text{D}.1) = \frac{c}{2} \quad (\text{D}.2)$$

Considering the root as $\frac{c}{2}$, *the maximum deviation* $max\ \varepsilon_{max}{}^2$, *becomes*:

$max\ \varepsilon_{max}{}^2 = \alpha_{(x_{crs},t)}\left(c\,ln(c) - \frac{c}{2}ln\left(\frac{c}{2}\right) - \left(c - \frac{c}{2}\right)ln\left(c - \frac{c}{2}\right)\right)$, which leads to:

$$max\ \varepsilon_{max}{}^2 = \alpha_{(x_{crs},t)}c\,ln(2) \quad (D.3)$$

D.2 Linearity

An important property of the cross-sectional equation given by (28) is that the multiplication of c and x by the same constant k leads to multiplication of $\varepsilon_{max}{}^2$ by the same constant. Thus, multiplying every term from the right side from (28) by k:

$$\varepsilon_{max}{}^2 = \alpha_{(x_{crs},t)}(kc\,ln\,kc\; - kx_{crs}\,ln(kx_{crs}) - (kc - kx_{crs})\,ln(kc - kx_{crs}))$$

If one uses logarithmic properties and some simple algebraic manipulation, the above equation is equal to:

$$\varepsilon_{max}{}^2 = k\alpha_{(x_{crs},t)}\,[c\,ln\,c\; - x_{crs}\,ln(x_{crs}) - (c - x_{crs})\,ln(c - x_{crs})] \quad (\text{D}.4)$$

Another important parameter that gives a linear effect for the squared deviations is inside the affinity coefficient given by $\alpha_{(x_{crs},t)}$, which is $\sigma^2{}_{w(x_{crs},t)}$:

$$\alpha_{(x_{crs},t)} = \frac{3}{ln\,c}.\phi^2\sigma^2{}_{w(x_{crs},t)} \quad (44)$$

Finally, linearity is derived from a proportional increase in the squared deviations given by:

- a proportional increase in input units in the cross-sectional, which also leads to a proportional increment in cross-sectional critical input; and
- a proportional increase in the weighted variance of returns.

An issue regarding linearity could arise from $\frac{3}{ln\,c}$ in equation (1), because $ln\,c$ should be multiplied by k leading to $\frac{3}{ln\,k + ln\,c}$. However, as seen in the equalization procedure, the k parameter is given by $k_j = \frac{c_{last}}{c_j}$, leading to:

$equalization\ (\frac{3}{ln\,c_j}) = \frac{3}{ln\,(k.c_j)} = \frac{3}{ln\,(\frac{c_{last}}{c_j}.c_j)} = \frac{3}{ln\,c_{last}} = constant$

Thus, the term $\frac{3}{ln\,c}$ does not present any effect over the pooled regression, remaining constant.

D.3 The Ellipse Method

The geometric model given by (24) resembles an ellipse. An ellipse has two focus of convergence, which has a parallel with the conflicting affinity and anti-cluster forces. Retrieving the equation of the ellipse, one gets:

$$\frac{(y'-c_2)^2}{b^2}+\frac{(x'-c_1)^2}{a^2}=1 \qquad (\text{D}.5)$$

If one replaces the ellipse variables given by $(\text{D}.5)$ with the following variables, an ellipse can be drawn as an approximation for (24).

y' : ε_{max}

x' : x_{crs}

c_1 : $c/2$ (center of the ellipse shifted from x')

c_2 : 0 (center of the ellipse shifted from y')

a : $c/2$ (major axis)

b : $\sqrt{ln2\alpha c}$, from $(\text{D}.3)$ (minor axis)

Thus, using the above variables, (25) is rewritten using $(\text{D}.5)$:

$$\frac{{\varepsilon_{max}}^2}{ln2\alpha c}+\frac{(x-\frac{c}{2})^2}{(\frac{c}{2})^2} = 1$$

This ellipse routine must be used after the rotation of the system (Appendix A).

Solving algebraically, leads to:

$${\varepsilon_{max}}^2 = 4\alpha ln2x\left(1-\frac{x}{c}\right)$$

Dividing both sides by x, leads to:

$$\frac{{\varepsilon_{max}}^2}{x} = 4\alpha ln2\left(1-\frac{x}{c}\right) \qquad (D.6)$$

$$Intercept: 4\alpha_{(x_{crs},t)}ln2$$

$$Slope: -\frac{4\alpha_{(x_{crs},t)}ln2}{c}$$

Finally, using a regression model based on the Ellipse Approach given by $(D.6)$, the result leads to an estimate of longitudinal critical input:

$$critical\ input(c) = -\frac{Intercept}{Slope} \qquad (\text{D}.7)$$

D.4 Data Preparation

The ellipse method requires a pooled regression using all cross-sectionals together. In order to prepare data for use in the Ellipse Method, every cross-sectional must be prior equalized using the following properties: i. linearity given by $(D.4)$; and ii. the linear effect given by the weighted variance of returns; Therefore, the following variables and parameters must be scaled to provide proper equalization:

- Inputs (x);
- Squared deviations $({\varepsilon_{max}}^2)$.

The equalization is required because every cross-sectional has different levels of inputs and weighted variance of returns, which leads to different levels of deviations $({\varepsilon_{max}}^2)$. Therefore, in the equalization procedure, each cross-sectional must be scaled up to the level of the last cross-sectional $(last)$ using linear effects. These linear effects in empirical data are shown in Figure 3 and explained in Discussion. Finally, after the proper equalization, the pooled regression with all cross-sectionals together can be computed.

From linearity given by $(D.4)$, the first step in equalization concerns the input values (x). This approach is described in $(D.8)$, where c_j is the sum of every input unit (x) at a given cross-sectional (j) , which is the cross-sectional critical input; c_{last} is the c_j value for the last cross-sectional; and k_j is the scalar value for equalization under linearity mentioned in $(D.4)$:

$$k_j = \frac{c_{last}}{c_j} \tag{D.8}$$

The same procedure can be done for the weighted variance of returns, replacing k_j by w_j:

$$w_j = \frac{{\sigma^2}_{w(last)}}{{\sigma^2}_{w(j)}} \tag{D.9}$$

Finally, every input x for each cross-sectional j must be multiplied by k_j to compare every x at the same level when all cross-sectionals are joined in a pooled regression. Simultaneously, as given by $(D.4)$, the squared entropy from section j must also be multiplied by k_j. Furthermore, the weighted variance of returns also gives a linear effect on squared entropies (adjusted by w_j). Thus, the proper equalization for both ${\varepsilon_{max}}^2$ and x_{crs} for each cross-sectional is:

$${\varepsilon_{max_j EQUALIZED}}^2 = {\varepsilon_{max_j}}^2 . k_j . w_j \tag{D.10}$$

$$x_{crs_j\,EQUALIZED} = x_{crs_j} . k_j \tag{D.11}$$