This is a repository copy of *An information-content based measure of proliferation as a proxi for structural complexity*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/143203/

Version: Accepted Version

Article:

Ruiz-Hernández, D., Menezes, M.B.C. and Amrani, A. (2019) An information-content based measure of proliferation as a proxi for structural complexity. International Journal of Production Economics, 212. pp. 78-91. ISSN 0925-5273

https://doi.org/10.1016/j.ijpe.2019.02.008

Article available under the terms of the CC-BY-NC-ND licence (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

An information-content based measure of proliferation as a proxi for structural complexity

Diego Ruiz-Hernández[a,*], Mozart B.C. Menezes[b], Aicha Amrani[c]

[a] *Sheffield University Management School, Conduit Road, S10 1FL, Sheffield, UK*
[b] *NEOMA Business School, 1 Rue du Maréchal Juin, 76130, France*
[c] *University of Bordeaux, IMS, 351 Cours de la Libération, Talence, 33400, France*

Abstract

In this article we propose the use of an information-content based measure as a proxy for supply chain complexity. The focus of our research is the problem of structural complexity in the supply chain, i.e. the complexity emanating from the proliferation of products, channels and markets. Notwithstanding it is widely agreed among practitioners that this proliferation damages supply chains, rendering them less efficient, there is still need for a mechanism for measuring structural complexity and evaluating its impact on the firm's performance. In an attempt to filling this void, we propose a definition that originates from the firm's business strategy and, based on it, suggest the direct use of entropy as a more austere measure for structural complexity than other available alternatives, which rely heavily in the use of typically hard to acquire data. We show that the suggested measure has some interesting mathematical properties (to which we refer to as internal consistency) together with the capability of reproducing certain empirical regularities observed in supply chain management (external consistency). Moreover, the proposed measure has attributes that are not present in other measures: it requires a limited and easily accessible amount of data, it allows direct comparison between firms or business units, and it is a useful tool for assessing the impact on structural complexity of alternative managerial decisions (the look-ahead property). Numerical examples are provided.
Keywords: Supply chain complexity; structural complexity; pars-complexity

*Corresponding author
Email address: `d.ruiz-hernandez@sheffield.ac.uk` (Diego Ruiz-Hernández)

Please cite as Ruiz-Hernandez, D., Menezes, M.B.C., and Amrani, A. (2019). An information-content based measure of proliferation as a proxi for structural complexity. *International Journal of Production Economics, 212: 78-91.*

measure; information content; entropy; proliferation

1. Introduction

The aim of this article is providing a framework which, on one hand, offers a conceptualisation of the term supply chain structural complexity, which is suitable for academic work and, on the other, provides a justification for the use of entropy as an information-content based measure of complexity that can be applied at the strategic-level analysis. This is achieved by associating the amount of information dealt with by a manager –in the form of messages originating from stock keeping units (SKUs), markets and channels- with what practitioners refer to as "complexity". In this process, we detach our discussion from administrative and operational-level issues in supply chain. Moreover, instead of identifying structural complexity with factors such as the complicatedness of transformation processes, intricateness of managerial procedures, or the uncertainty inherent to supply and demand, we consider it a phenomenon emanating from the proliferation of products, channels and markets (Heywood et al., 2007), which can potentially aggravate the negative effects of those factors on the firm's performance.

In this work we adopt a notion of complexity voiced by practitioners in supply chain management, who associate complexity with numerousness or proliferation[2]. Three examples must suffice to illustrate this. By 2006, researchers from George Group (2006) affirmed: "as businesses increase their product and service portfolios in response to evolving customer demands ..., they run the risk of adding too much complexity, which can tax existing resources and ultimately harm returns". In turn, John Mariotti (2008) wrote: "In the quest for high growth in low/no-growth markets, companies have proliferated nearly every-thing: products, customers, markets, suppliers, facilities, locations, etc. Some of the time, the proliferation actually does lead to top-line revenue growth, but as the top line goes up, the bottom line actually goes down dramatically". Finally, Fisher et al. (2017) affirm that in the quest for increasing revenue.

[2]We hereby stress that our work distinguishes between complexity, associated with numerousness, and risk, mostly associated with uncertainty.

businesses “had continued to chase growth by opening new stores far past the point of diminishing returns... [and] keep expanding until their chains begin to collapse under their own weight” .

An important message emerges from this discussion: supply chain complexity stems from strategic choices and other organisational and operational decisions made without considering their systemic effects (Saeed and Young, 1998). The costs are huge: it erodes profit, increases inventory, hinders the agility of the supply chain (Adams et al., 2016) and may even increase capital investment costs (Saeed and Young, 1998). In 2007, a study by A.T. Kearney suggested that complexity management could lead to an upturn in EBIT of 3 to 5 percentage points (Scheiter et al., 2007). More recently, Adams et al. (2016) mention the case of a large food manufacturer facing about 10% margin loss due to increasing complexity; likewise, in the complexity analysis of a large manufacturer Menezes and Ruiz-Hernández (2019) find that for every unit of increase in complexity, based on the measure proposed in this article, there may be a reduction of between 1.4 and 3.4 percentage points in gross margins[3].

The discussion above stresses the need for a measure of supply chain complexity that can be used for evaluating its impact on costs and financial performance. This article constitutes an attempt for providing such framework. We present a theoretical framework that satisfies the requirement, established by Wittgenstein (1922), of double logical consistency for any model of the physical world to be valid (namely, self-logical or internal consistency and consistency with the world it is describing). An in depth discussion on the empirical and practical implications of the proposed measure is presented in Menezes and Ruiz-Hernández (2019).

In section 2 a working definition of supply chain structural complexity is provided. The theoretical framework for the measurement of supply chain complexity and its mathematical properties are presented in Section 3. In that section we also present a collection of examples illustrating the application of the measure to questions that frequently rise in supply chain management and design. In Section 4 the theoretical results and conclusions are contrasted against

[3]95% confidence interval.

a collection of so-called *stylised facts*, i.e. managerial decisions about whose consequences there is general consensus supported by empirical evidence. The rationale behind is that a measure which is externally consistent should be able to reproduce well established facts or empirical regularities. Section 5 concludes this article.

2. A conceptual framework for supply chain complexity

As Weber (2005) points out, there seems to be some general agreement that something is complex if it is "made of (usually several) closely connected parts". Referring to Heylighen (1999), Weber suggests that complexity increases when the variety (distinction) and dependency (connection) increases; a notion that was already suggested, among others, by Simon (1962) and Waldrop (1993). This highlights a common trait in most definitions of supply chain complexity: proliferation. Indeed, proliferation of activities, channels, customers, processes, products, markets, and so on, is at the root of the definition of complexity for both, practitioners (George and Wilson, 2004; Heywood et al., 2007; Mariotti, 2008; Golfmann and Lammers, 2015; Adams et al., 2016; Hirose et al., 2017) and academics (Rutenberg and Shaftel, 1971; Wilding, 1998; Fisher and Ittner, 1999; Choi et al., 2001; Novak and Eppinger, 2001; Blecker and Kersten, 2006; Choi and Krause, 2006; Abdelkafi, 2008; Schaffer and Schleich, 2008; Bozarth et al., 2009; Subramanian and Rahman, 2014; Aitken et al., 2016). We base our work on the definition provided by Saeed and Young (1998):

> *Complexity is the systemic effect that numerous products, customers, markets, processes, parts, and organizational entities have on activities, overhead structures, and information flows.*

Once it has been agreed that complexity is a consequence of proliferation, we still need to bind the sphere of complexity that we encompass in our definition of supply chain complexity. Several categories have been proposed in literature for classifying complexity, see Table 1. Serdarasan (2013) and de Leeuw et al. (2013) provide excellent reviews of these and other classifications found in literature. Parallel to this, an important amount of work has been devoted to the analysis and measure of product portfolio complexity and its relation to operational performance. Jacobs and Swink (2011) provide a detailed account of the available work in this area.

Focus	**Categories**	**Reference**
Value Source Dynamics Coordination	Good and bad Internal and external Static and operational Objectives, customers and variety	Abdelkafi (2008)
Aggregation	Systems or business unit level	Aitken et al. (2016)
Scope	Structural (market, product, organisational) and process complexity	Lindemann et al. (2008)
Source	External and internal (organisational, products and processes)	Kaivola (2017)
Source	Internal or external (originated by customers or suppliers)	Isik (2010)
System related Supply chain related	Detail or dynamic complexity Up-stream, internal, or downstream complexity	Bozarth et al. (2009)
Source	Horizontal, vertical, or spatial complexity	Bode and Wagner (2015)
Scope	Depth and breadth complexity	Wang et al. 2000

Table 1: Business complexity categories found in literature

The notion of supply chain complexity that we use in this article, the one associated to the proliferation of products, markets and customers, fits within many of these categories but, unfortunately, does not have a direct correspondence with any of them. It can be good or bad depending on its magnitude; can be considered static, but it affects the overall system dynamics; it is closely related to the notion of detail complexity, but at the same time puts focus at both system and business unit levels, etc.

In order to fill this gap, we propose an alternative category: structural complexity. It originates from the firm's business strategy and is linked to the answer given to questions regarding *what do customers want*, *where they are*, and *how can they be reached* in the form of products, markets and channels, respectively. This category fits within the class of breath complexity (Wang and von Tunzelmann, 2000) and is closely related to the notion of structural complexity proposed by Lindemann et al. (2008).

The three dimensional nature of our definition comes from the observation that the physical movement of a product, and therefore the origins of the information that a manager has to deal with, can be summarised by identifying the SKU where it comes from, the market where it is sold and the distribution channel by means of which it is delivered to the market. Each unique combination of these three elements within the supply chain constitutes what we refer to as a *pars* (plural *partes*). In what follows, we consider that a *pars* is fully characterised by the triple $\{SKU, market, channel\}$.

As discussed earlier, there is substantial agreement by academics and practitioners that it is precisely the proliferation of those three elements (products, channels and markets) or –in our terminology- the proliferation of *partes*, which constitutes the main source of complexity in the supply chain. As this proliferation results in an increased amount of information delivered to the manager, measuring the amount of information created by the physical flow of products and services stands for measuring complexity in the supply chain. This approach is supported by some practitioners, see for example Hirose et al. (2017), who suggest that growth in total revenue is driven by the contribution of a limited number of "cells" (out of the thousands typically managed in a firm), which they define as specific combinations of products and geographies. Consequently, they argue, complexity can often be mitigated by reducing the number of low revenue-growth cells.

It is convenient to make here a pause and remember that our notion of structural complexity does not account for complicatedness or uncertainty. Conversely, we consider structural complexity as a factor that may potentiate –or mitigate- the negative effects brought, in one hand, by complicated industrial processes or administrative procedures; and, in the other, by the random nature of the supply, transformation or demand. It can be argued, for example, that proliferation of products or markets is a mechanism for diversifying risk. In such case, the manager of a single line of a seasonal product may be facing a low level structural complexity but still dealing with high uncertainty; whereas a person in charge of a larger number of products may face higher levels of structural complexity but lower total demand variability. Likewise, a complex supply chain serving many differentiated products in a national market may face few administrative challenges; whilst a producer exporting a couple of perishable products will need to deal with numerous and complicated administrative procedures.

In the following section we provide a measure for structural complexity that is simple and independent of the supply chain's scope considered. Moreover, it is shown that this measure satisfies the double logical consistency requirement mentioned in the introduction to this article.

3. A theoretical framework for measuring complexity

In the following lines we present, and analyse, a measure for *pars*-Complexity based on the *Theory of Communication* developed by Shannon (1948) and Shannon and Weaver (1949). Shannon's measure of information quantifies the expected amount of information required for describing the state of a system. In our context, it provides a measure of the information generated by the physical flow of goods in the supply chain and, therefore, can be used as a measure for a supply chain's structural complexity.

Several measures for supply chain complexity have been proposed in literature. In a classification proposed by de Leeuw et al. (2013), the authors distinguish two main directions of research: "exploratory studies" and "entropy-based studies". The first group, exploratory studies, includes work aimed at evaluating the cost of complexity –or its impact on performance- by indirectly measuring it through a number of indicators or "drivers" without providing a closed form expression for its measurement. The second group includes efforts made for quantifying complexity by means of a entropy-based measure. Table 2 shows some of the available work on each of these categories. Entropy-based measures have also been deployed for measuring decision-making efficiency, e.g. Gong et al. (2014); Fan et al. (2017b); Wang et al. (2017). An alternative line of research, leaded by Blecker and Abdelkafi (2006b), proposes a measure of complexity based on the work by Suh (1999).

As de Leeuw et al. point out, the available measures of complexity require a large amount of data of different nature. For example, Deshmukh et al. (1998)'s measure is defined in terms of parts, routing, and resources from the workshop; Sivadasan et al. (1999) use information on orders, sales, deliveries, production schedule, purchases and so on. Likewise, Isik (2010) proposes a modified version of Shannon's measure that requires information on demand (actual and sched-

Exploratory Studies	Entropy Based Studies
Novak and Eppinger (2001)	Palepu (1985)
Vachon and Klassen (2002)	Karp and Ronen (1992)
Perona and Miragliotta (2004)	Frizelle and Woodcock (1995)
Kaluza et al. (2006)	Deshmukh et al. (1998)
Schaffer and Schleich (2008)	Sivadasan et al. (1999)
Bozarth et al. (2009)	Sivadasan et al. (2002)
Garbie and Shikdar (2011)	Sivadasan et al. (2006)
Jacobs (2013)	Wu et al. (2007)
de Leeuw et al. (2013)	Frizelle and Suhov (2008)
Subramanian and Rahman (2014)	Jacobs (2008)
Bode and Wagner (2015)	Isik (2010)
Hendricks and Singhal (2016)	Sivadasan et al. (2010)
Fan et al. (2017a)	Isik (2011)

Table 2: Alternative measures of complexity in literature

uled) and its variations. Our work departs from this line of by proposing a more parsimonious measure of information, based on the contribution of each *pars* to the total monetary value of sales of the firm or business unit (notwithstanding the fact that other sales-related measures can also be used, monetisation is an important aid for comparability between different SKU's). This approach has already been used, from an economics point of view, by Jacquemin and Berry (1979) and Palepu (1985), who use sales-based entropy as a measure of an industry's or a firm's total diversification, respectively. The main difference lies on the scope: whilst earlier works focused on product diversification, our analysis considers not only products but also channels and markets and encompasses the whole breadth of the supply chain. Moreover, the aim of those studies was establishing a link between diversification and performance, whereas the intention of this work is providing a complete theoretical framework for the measurement of structural complexity and the analysis of the complexity implications of different managerial decisions and supply chain strategies.

Before introducing our measure, we propose a toy model of a supply chain aimed at helping the reader to put the ideas developed in this section into focus. Consider a factory –illustrated in Figure 1- producing four different products or SKUs. The markets are represented by two different stores, namely, A and B. The arrows in the figure represent the direction of the product flow. Following the terminology introduced above, this supply chain can be characterised as a set $\mathcal{P}$ of *partes*, each represented by the triplet $\{SKU, market, channel\}$. Notice that, because in this example SKU_1 is sold in two different shops, $|\mathcal{P}| = 5$.

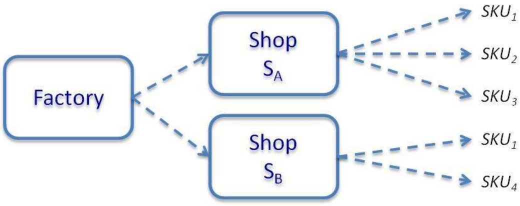


Figure 1: Factory producing four SKUs distributed through two different shops, $|\mathcal{P}| = 5$.

We can now focus on the discussion of the complexity measure. Take a supply chain with a set $\mathcal{P}$ of *partes* and consider a fixed period for discussion. Each *pars* $i \in \mathcal{P}$ contributes a fraction $\mathbf{p}_{(i)}$ of the total sales, with $\mathbf{p} = \left(\mathbf{p}_{(1)}, \ldots, \mathbf{p}_{(|\mathcal{P}|)}\right)$ representing the vector of weights of each of the *partes* in $\mathcal{P}$. Clearly, $\sum_{i=1}^{|\mathcal{P}|} \mathbf{p}_{(i)} = 1$. It will help our following discussion if we notice that these proportions can also be understood as the probability that one monetary unit of revenue has been generated by certain *pars*.

Let, $X \in \{1, \ldots, |\mathcal{P}|\}$ be a random variable representing the *pars* associated with one money unit taken at random from the sales pool within the chosen period. The probability of $X = i$ is given by $\mathbf{p}_{(i)}$ and, therefore, X is fully characterised by $\mathbf{p}$.

Definition 1. *We define* pars*-Complexity as the expected amount of information (measured in bits) necessary to express all possible values of* X*, i.e.*

$$C_p(X) \triangleq C_p(\mathbf{p}) = \sum_i \mathbf{p}_{(i)} \log_2 \left(\frac{1}{\mathbf{p}_{(i)}}\right). \tag{1}$$

Notice that our *pars*-Complexity measure, C_p, is Shannon's information measure (Shannon, 1948) directly applied to a random variable representing the sales distribution among *partes*.

It is important to underline the fact that our approach is in sharp contrast with the views of other authors (e.g. Calinescu et al., 1998; de Leeuw et al., 2013), who consider that entropy-based measures demand a large amount of data, making it a "costly and time consuming exercise". Moreover, by respecting Shannon's formulation, we keep the Information Theory framework intact, therefore inheriting all its properties. This allows us to bring forward a measure that, unlike other exploratory approaches, can be applied and used to making comparisons between business units with different complexity levels (Allesina et al., 2010; de Leeuw et al., 2013); and to establishing a relationship between structural complexity and the firm's financial performance (Menezes and Ruiz-Hernández, 2019).

It is important at this point to highlight that, whereas other proposed measures for complexity rely in actual performance measures and other values which are assumed to be known, but in practice hard to obtain (revenue, sales, processing and set-up times, number of SKU's, markets, facilities, employers or suppliers, and so on), the measure proposed in this article only requires information about the market share that each *pars* in the system represents. This confers our measure a computational simplicy and ease of interpretation that cannot be found in other existing measures of complexity. Please see the Appendix to this work and Table 4 therein for a more in depth discussion of this issue.

In the following section we review the theoretical properties of our *pars*-Complexity measure, and provide some examples which illustrate certain features of the measure which are relevant within the supply chain complexity framework.

3.1. Theoretical Properties and Internal Consistency

The internal consistency of the measure, the fact that it does not lead to contradiction or paradoxes, stems from its definition and its mathematical properties. These properties have been well established by Boltzmann's work (see for example Tolman, 1938) and -regarding its use as a measure for information- by Shannon (1948) and by Shannon and Weaver (1949). In this section we recall some of those properties, together with other useful features that contribute simplify the analysis of a supply chain's structural complexity.

Property 1. ***Shannon (1948)*** *Let $\mathcal{P}$ be the set of partes constituting a supply chain and X be an associated random variable characterised by the probabilities vector $\mathbf{p}$. The complexity measure $C_p(X)$ defined in* (1) *satisfies:*

1.1. $C_p(X)$ is continuous and concave in X.

1.2. If $|\mathcal{P}| = 1$, then the system's complexity is zero, i.e. $C_p(X) = 0$.

1.3. C_p attains a maximum when $\mathbf{p}_{(i)} = \frac{1}{|\mathcal{P}|}$ for all $i = 1, \ldots, |\mathcal{P}|$. Such maximum is equal to $\log_2 |\mathcal{P}|$.

Proof *Properties 1.1 to 1.3 are immediate consequences of the definition of* C_p*.* ■

The following properties show certain attributes of C_p which are desirable in a good measure for structural complexity. Moreover, they guarantee the absence of paradoxes or contradictions in the computation of *pars*-complexity values for different supply chain configurations. For the sake of clarity, and withou loss of generality, these properties will be motivated within the context of supply chain management and structural complexity.

Property 2. *Consider a collection* S° *of* N *independent systems* $\mathcal{P}_1^\circ, \ldots, \mathcal{P}_N^\circ$*. Each of these systems distributes* $K_j = |\mathcal{P}_j|$ *different* partes *and has an associated* K_j*-vector of weights* $\mathbf{p}_j^\circ$*, for* $j = 1, \ldots, N$*. Let* $S^\bullet$ *be a system with cardinality* $|S^\bullet| = N$*, whose elements have a one to one correspondence with the members of* S°*, and with weights represented by the* N*-vector* $\mathbf{p}^\bullet$*, then it holds that*

$$C_p\Big(\oplus_{j=1}^{N} \left(\mathbf{p}_{(j)}^\bullet \, \mathbf{p}_j^\circ\right)\Big) = \sum_{j=1}^{N} \mathbf{p}_{(j)}^\bullet \, C_p\left(\mathbf{p}_j^\circ\right) + C_p\left(\mathbf{p}^\bullet\right) \geq \sum_{j=1}^{N} \mathbf{p}_{(j)}^\bullet \, C_p\left(\mathbf{p}_j^\circ\right) \qquad (2)$$

where $\oplus_{i=1}^{n} q_j = q_1 || q_2 || \cdots || q_n$*, is a notational shortcut for the concatenation of vectors* q_1 *to* q_n*; and* $\mathbf{p}_{(j)}$ *represents the* j*-th element of vector* $\mathbf{p}$*.*
Proof *The first part of equation* (2) *can be obtained, by simple algebraic manipulation, applying the definition of* C_p *in* (1) *on the l.h.s. of equation* (2)*; the second one is a direct consequence of Property 1.1.* ■

Property 3. *Let* $\mathbf{p}^\circ$ *be a* N*-vector of weights representing the distribution of sales in a particular system,* S°*. Let* $\mathcal{P}_j$ *be the set of* K_j *inputs used for the production of* pars j *in* S° *and* $\mathbf{p}_j$ *their contributions (in percentage) to* j*. Define now* $\mathbf{p}_j^I = \mathbf{p}_{(j)}^\circ \cdot \mathbf{p}_j$*, for all* $j = 1, \ldots, N$*, representing the decomposition of the contribution to sales of* pars j *in terms of its inputs. With these elements, the concatenation* $\mathbf{p}^I = \oplus_{j=1}^{N} \mathbf{p}_j^I$ *is a vector consisting of the fractions of total sales attributed to each input-*pars *combination in system* S°*. It holds that*

$$C_p\Big(\mathbf{p}^I\Big) \geq C_p\Big(\mathbf{p}^\circ\Big) \, .$$

Proof *This result follows directly from Property 2.* ■

Property 4. *Consider a system* S° *consisting of* N *independent subsystems* $\mathcal{P}_1^\circ, \ldots, \mathcal{P}_N^\circ$*. The contributions of the subsystems to the total sales of* S° *are given by the weights vector* $\mathbf{p}^\circ$*. Additionally, each subsystem* $\mathcal{P}_j$ *distributes a number* K_j *of* partes*, with an associated vector of weights* $\mathbf{p}_j^\circ$*. The complexity of system* S° *can be readily computed by means of equation* (2)*, and is represented by* C_p°*. Consider now an alternative design where the supply chain is characterised by a system* $S^\bullet$ *consisting of* M *independent subsystems* ($M \neq N$) *with weights vector represented by* $\mathbf{p}^\bullet$*. Each subsystem* $\mathcal{P}_h^\bullet$*, delivers* G_h partes *with weights* $\mathbf{p}_h^\bullet$*. The complexity of this alternative system is given by* $C_p^\bullet$*. If* $\sum_{j=1}^{N} K_j = \sum_{h=1}^{M} G_h$*, i.e. the number of* partes *in both systems is the same;*

and $\oplus_{j=1}^{N}\left(\mathbf{p}^{\circ}_{(j)}\mathbf{p}^{\circ}_{j}\right) = \oplus_{h=1}^{M}\left(\mathbf{p}^{\bullet}_{(h)}\mathbf{p}^{\bullet}_{j}\right)$*, i.e. the contributions to the total sales of each pars is the same in both systems, then*

$$C_p^{\circ} = C_p^{\bullet} .$$

Proof *This property is a direct consequence of Property 2.* ■

Property 5. *Consider a pars* r *that contributes a fraction* p_r *of total sales in system* $\mathcal{P}$*, and let* $C_p\left(\mathbf{p}_{[r]}\right)$ *be the system's complexity, where* $\mathbf{p}_{[r]}$ *is used to emphasize the dependency of* C_p *on* pars r*. Assume now that* pars r *is substituted by two alternative* partes*, each contributing to total sales by* λp_r *and* $(1-\lambda)\,p_r$*, with* $\lambda \in [0,1]$*, and let* $\mathbf{p}_{[r,\lambda]}$ *be the updated vector of percentage contributions to sales. Then,*

$$C_p\left(\mathbf{p}_{[r]}\right) \leq C_p\left(\mathbf{p}_{[r,\lambda]}\right) .$$

Proof *This property is a consequence of the concavity of* C_p *and Property 2.* ■

Comment

Property 1.2 states that a single SKU generates minimal information and, thus, shows no complexity. Property 1.3 indicates that maximal complexity is attained when the market is shared evenly by all SKU's. Property 2 establishes that whenever a supply chain consists of two subsystems at different levels, its total complexity is computed as the weighted sum of the downstream subsystem's complexity plus the complexity derived from the composition of the upper level subsystem. As it will be shown in Section 3.2.6, and it is stated in Proposition 2, this property can be easily generalised to any number of levels or subsystems. Property 3 states that up-stream complexity cannot be smaller than downstream complexity when the system is considered as a whole. Property 4 establishes that, as long as the final product mix remains the same, the total structural complexity is independent of the design of the supply chain. Finally, Property 5 indicates that whenever a *pars* is substituted for two alternative ones whose contributions to the system add up to that of the one they substitute, the overall system's complexity increases.

3.2. Illustrative Examples

In this section we present a set of examples aimed at illustrating the different attributes and properties of the measure. They will also allow us to derive some general results which are stated as propositions.

3.2.1. Basic configuration

Consider two independent systems, S^A and S^B, illustrated in Figure 2. System S^A produces and distributes two *partes* with weights vector $p^A = (0.6, 0.4)$ñ whereas system S^B produces and distributes three *partes* with weights $p^B = (0.85, 0.1, 0.05)$. After applying (1) we obtain

$$C_p(\mathbf{p}^A) = 0.6 \log_2 \frac{1}{0.6} + 0.4 \log_2 \frac{1}{0.4} = 0.971; \text{ and}$$
$$C_p(\mathbf{p}^B) = 0.85 \log_2 \frac{1}{0.85} + 0.1 \log_2 \frac{1}{0.1} + 0.05 \log_2 \frac{1}{0.05} = 0.748,$$

showing that although there are more SKUs on system S^B, its *pars*-complexity value is lower than that of system S^A. This is so because in S^B there is more concentration of sales on one particular item, while in system S^A market shares are more uniformly spread. This result is closely related to Property 2, which states that the more evenly distributed are the weights in a system, the more complex it is.

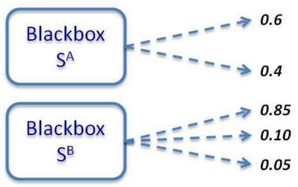


Figure 2: Supply chain with two distribution centres serving two markets.

Let us now assume that systems S^A and S^B are distribution centres of the same company. Assume also that system S^A contributes with 75% of the total sales of the company, system S^B with the remaining 25%, and that there is a central facility, S° that supplies both centres. This situation is illustrated in Figure 3.

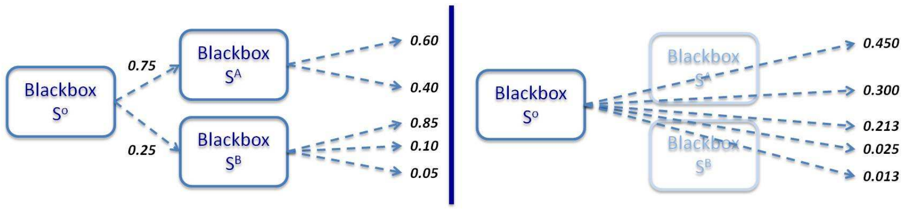


Figure 3: Distribution centre serving two markets. On the left panel, S° manages two subordinated systems with limited visibility; on the right one, S° has full visibility of the market.

.

The *pars*-complexity of this system can be computed using the l.h.s. of expression (2), i.e.

$$\begin{aligned} C_p\Big((0.75\mathbf{p}^A) \oplus (0.25\mathbf{p}^B)\Big) &= (0.75 \times 0.6)\log_2 \frac{1}{(0.75 \times 0.6)} + (0.75 \times 0.4)\log_2 \frac{1}{(0.75 \times 0.4)} + \\ &\quad (0.25 \times 0.85)\log_2 \frac{1}{(0.25 \times 0.85)} + (0.25 \times 0.1)\log_2 \frac{1}{(0.25 \times 0.1)} + \\ &\quad (0.25 \times 0.05)\log_2 \frac{1}{(0.25 \times 0.05)} \\ &= 1.726, \end{aligned}$$

Notice that the complexity of the whole system integrated under the management of S° (i.e. $C_p\left((0.75\mathbf{p}^A) \oplus (0.25\mathbf{p}^B)\right) = 1.726$) is larger than the weighted sum of the individual complexities of the two integrating subsystems (i.e. $\left(0.75C_p(\mathbf{p}^A)\right) + \left(0.25C_p(\mathbf{p}^B)\right) = 0.9153$). The difference between these two values corresponds to the extra complexity introduced in the system when joining S^A and S^B together under the management of S°, i.e. $C_p\left(\mathbf{p}^\circ\right) = 0.811$. Indeed, from the right hand side of (2) we have that $C_p\left((0.75\mathbf{p}^A) \oplus (0.25\mathbf{p}^B)\right) = C_p(\mathbf{p}^\circ) + \left(0.75C_p(\mathbf{p}^A)\right) + \left(0.25C_p(\mathbf{p}^B)\right)$. This discussion is summarised in Figure 4.

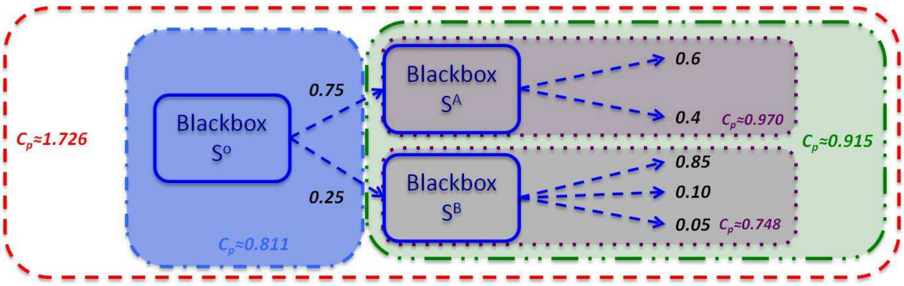


Figure 4: Measures of supply chain complexity at different levels of aggregation

3.2.2. Supply chain expansion

Consider now the supply chain described in the previous example, with total *pars*-complexity is $C_p = 1.726$. Assume now that the management decides to expand its network by 20% opening a new facility C that serves four new markets with weights $\mathbf{p}^C = (3/20, 1/4, 12/25, 3/25)$. With the new addition, the system's weights become $\mathbf{p}^\bullet = (5/8, 5/24, 1/6)$. It is straightforward to verify that the expanded system's *pars*-complexity increases to $C_p = 2.386$.

This suggests that horizontally increasing the scope of a supply chain (i.e. adding same-level supply chain partners) may also increase the *pars*-complexity value of the chain. However, the reader must be aware that, although an increase in the size of the supply chain will typically result in an increase in complexity, this may not always be the case. Take for example the (unrealistic) case where an expansion rebounds in a nine times increase in market size, with the new market weights given by $\mathbf{p}^C = (3/4, 1/8, 1/16, 1/16)$. In such case, the upper-level weights will be given by $\mathbf{p}^\bullet = (5/72, 1/43, 49/54)$ and the system's *pars*-complexity becomes $C_p = 1.681$. This is explained by the following result:

Proposition 1. *The larger the market share of a facility is, the closer the system's complexity becomes to that particular facility's complexity value.*

Proof *The proof goes straightforwardly by considering that* $\sum_{k=1}^{N} \mathbf{p}^\bullet_{(k)} = 1$ *and taking the limit in the left hand side of* (2) *when* $\mathbf{p}^\bullet_{(j)} \to 1$ *for a given* j *in* $j = 1, \ldots, N$. ∎

3.2.3. Supply chain design

Property 4 can be illustrated by introducing an alternative design for the supply chain described in section 3.2.2. Consider that, instead, the supply chain consists of three distribution centres with demand shares given by $\mathbf{p}^\bullet = (13/35, 1/18, 5/96)$. The distribution centres serve the same nine retailers but now the weights are distributed in the following way: $\mathbf{p}^A = (37/55, 17/79, 1/9)$, $\mathbf{p}^B = (23/51, 35/97, 3/16)$, and $\mathbf{p}^C = (17/55, 36/55, 2/55)$. As the final products' distribution remains the same, it is easy to verify that the *pars*-complexity value of the new design remains $C'_p = 2.386$.

Going back to the original configuration of the supply chain, consider now that the management decides to split market A_1 into two separate markets, one of them representing 75% of the original demand and the other the remaining 25%. The weights vector for centre A becomes $\mathbf{p}^{A'} = (9/20, 3/20, 2/5)$ and its complexity value $C_p\left(\mathbf{p}^{A'}\right) = 1.458$. Consequently, the total system's complexity is now $C'_p = 2.691$. This illustrates Property 5.

3.2.4. Pars consolidation

An important consequence of Property 5 is illustrated by the following example. Consider the supply chain configuration resulting from the expansion described in Section 3.2.2. Assume that the firm decides to eliminate distribution centre B, and to deliver its sales from facility A. A will now be in charge of

83.3% of the market, with weights vector $\mathbf{p}^{A'} = (9/20, 3/10, 17/80, 1/40, 1/80)$. It is easy to verify that the total *pars*-complexity of the consolidated system (obtained by means of the expression in the centre of equation (2)) will remain equal to the original, i.e. $C_p = 2.386$.

Imagine now that, before closing centre B, certain market was served by both A and B. Assume that this market's contribution to total sales was delivered by A_2 and B_1. The firm can further reduce complexity by consolidating those sales (we refer to this as *pars* consolidation). Under consolidation, the aggregated weights vector becomes $\mathbf{p}^{A''} = (9/20, 41/80, 1/40, 1/80)$ and the complexity of the consolidated system is $C_p^{\circ} = 1.968$. This result is summarised by the following property:

Property 6. *Consider two partes in a system, originating from the same SKU, i, and delivered to the same market through two different channels that distribute proportions λ and $(1-\lambda)$ of the total sales, respectively. Let $\mathbf{p}_{A,i}$ be the weight of SKU i in channel A and $\mathbf{p}_{B,i}$ its weight in channel B. If the sales of SKU i are consolidated in one single channel, then it holds that*

$$C_p\left(\mathbf{p}_{[\mathbf{p}_{A,i},\mathbf{p}_{B,i},\lambda]}\right) \geq C_p\left(\mathbf{p}^{\circ}_{[i]}\right) \tag{3}$$

where $\mathbf{p}^{\circ}_{[i]} = \left(\lambda \mathbf{p}_{A,i} + (1-\lambda)\,\mathbf{p}_{B,i}\right)$ represents the weight of i in the consolidated system; and $\mathbf{p}_{[\mathbf{p}_{A,i},\mathbf{p}_{B,i},\lambda]}$ is used to stress the dependency of C_p on channels A and B and contributions parameter λ.

This result is a direct consequence of Property 5 and can be easily generalised to any number of *partes*. It states, in short, that consolidation of *partes* will always reduce complexity.

3.2.5. Adding/Removing partes

Suppose now that the firm decides to remove one *pars* from the expanded system described in Section 3.2.2, say B_3. In order to adjust the contributions to total sales of the remaining *partes* we define $\overline{\mathbf{p}}_{(i)} = \frac{\mathbf{p}_{(i)}}{1-\mathbf{p}_{(j)}}$, where j is the position in $\mathbf{p}$ of the withdrawn *pars*, and $i = 1, \ldots, N$. Finally, after removing element j we obtain the updated contributions vector $\overline{\mathbf{p}}$, where $|\overline{\mathbf{p}}| = N-1$. Given that B_3 occupies the 5th position in $\mathbf{p}$, with contribution $\mathbf{p}_{(5)} = 1/96$, the resulting system's complexity after removing pars B_3 becomes $C_p(\overline{\mathbf{p}}) = 2.327$; i.e. removing one pars has reduced the total complexity of system S°.

Alternatively we can consider the inclusion of a new *pars*. In such case, the elements of the contributions vector will be adjusted by $\widetilde{\mathbf{p}}_{(i)} = \frac{\mathbf{p}_{(i)}}{1+\mathbf{p}_{(N+1)}}$, where $\mathbf{p}_{(N+1)}$ represents the expected contribution of the new pars to the total output. Finally, $\mathbf{p}_{(N+1)}$ is appended to the new contributions vector, $\widetilde{\mathbf{p}}$. If, for example, the new *pasr* brings an increase of 15% in total sales, is straightforward to compute vector $\widetilde{\mathbf{p}}$ and to verify that $C_p(\widetilde{\mathbf{p}}) = 2.634$; namely, adding a new *pars* has increased the system's complexity.

Notwithstanding these examples agree with the general notion that withdrawing a *pars* will reduce complexity and, conversely, including a new one will increase it, it will not be hard for the reader to build examples that work in the opposite direction (for example, if $\mathbf{p}_{N+1} = 17/20$, then $C_p(\widetilde{\mathbf{p}}) = 2.285$. This issue is explored in depth in Section 4.5.

3.2.6. Multi-level System

Let us now introduce an additional level to our expanded network. Assume now that our system has three level. It may well be seen as a supply chain consisting in a production facility S° that delivers its products to three regional distribution centres according to the weights vector $\mathbf{p}^\circ = (5/8, 5/24, 1/6)$. These regional distribution centres serve, in turn, a number of local distribution centres according to the weights vectors: $\mathbf{p}^A = (3/5, 2/5)$; $\mathbf{p}^B = (17/20, 1/10, 1/20)$; and $\mathbf{p}^C = (3/20, 1/4, 12/25, 3/25)$, respectively. Finally, each of the local distribution centres delivers its products to different retailers. The markets served by the facilities depending on regional centre A are given by $\mathbf{p}_1^A = (1/8, 3/8, 1/2)$; and $\mathbf{p}_2^A = (1/3, 2/3)$. Facilities served by centre B have the following market distribution: $\mathbf{p}_1^B = (1/4, 1/12, 1/6, 5/12, 1/12)$; $\mathbf{p}_2^B = (1/5, 3/5, 1/5)$; and $\mathbf{p}_3^B = (3/8, 1/4, 3/8)$. Finally, regarding centre C we have $\mathbf{p}_1^C = (4/9, 5/9)$; $\mathbf{p}_2^C = (4/11, 5/11, 2/11)$; $\mathbf{p}_3^C = (7/12, 5/12)$; and $\mathbf{p}_4^C = (1/4, 1/3, 1/4, 1/6)$.

Successive applications of equation (2) reveal that the total *pars*-complexity of this system is $C_p = 3.756$. However, it is important to highlight that such number can be obtained using (2) at different levels of aggregation. First, we can take a more telescopic approach and consider only the relation between the producer and the final markets, by computing the share of the total sales represented by each final retailer we can obtain $C_p(\widetilde{\mathbf{p}}^\circ) = 3.756$, where $\widetilde{\mathbf{p}}^\circ = (3/64, 9/64, 3/16, \ldots, 1/300)$. Secondly, we can aggregate the markets and start by computing the complexity faced by the regional distribution centres (ignoring the existence of the local ones). Taking for example centre A, we see that its weights will be given by $\widetilde{\mathbf{p}}^A = (3/40, 9/40, 3/10, 2/15, 4/15)$ and, therefore, $C_p\left(\widetilde{\mathbf{p}}^A\right) = 2.182$. Following the same procedure for $C_p\left(\widetilde{\mathbf{p}}^B\right)$ and $C_p\left(\widetilde{\mathbf{p}}^C\right)$, taking weights and adding $C_p\left(\mathbf{p}^\circ\right)$ we obtain $C_p = (2.182 \cdot 5/8 + 2.709 \cdot 5/24 + 3.014 \cdot 1/6) + 1.326 = 3.756$. Finally, the same value can be obtained taking into account the complexity observed at each local distribution centre. In this case, it is easy to confirm that the complexity at centre A_1 is given by $C_p\left(\mathbf{p}_1^A\right) = (0.375 + 0.531 + 0.500) = 1.406$; that the complexity at centre A_2 is $C_p\left(\mathbf{p}_2^A\right) = 0.918$; and, therefore $C_p\left(\mathbf{p}^A\right) = (0.6 \cdot 1.406 + 0.4 \cdot 0.918) + 0.971 = 2.182$. The total system's complexity is computed as the weighted mean of the regional centres' complexities and is equal to 3.756 as before. This nesting property can be formalised as follows:

Proposition 2. *Let $S^\bullet$ be a distribution system consisting of N subsystems S_n°, $n = 1, \ldots, N$ consisting, in turn, of a number n_m of subsystems each; then*

$$C_p = \sum_{n=1}^{N} \mathbf{p}_{(n)}^\bullet \left(\sum_{k=1}^{n_m} \mathbf{p}_{n,(k)}^\circ \, C_p\left(\mathbf{p}_k^n\right) + C_p\left(\mathbf{p}_n^\circ\right) \right) + C_p\left(\mathbf{p}^\bullet\right) . \tag{4}$$

Proof *This feature is a consequence of the additive property of the complexity measure* (1)*, summarised in Property 4 and the discussion around it.* ■

This result allows us to compute the overall supply chain's complexity by weighting previously calculated downstream complexity levels and adding the complexity increase due to the inclusion of up-stream levels in the supply chain.

3.2.7. Production facilities

Consider now the case where system S° is a production facility. The inputs necessary for the SKUs produced by S° are given in Table 3. Columns include *pars* identifiers, the distribution centres from which each *pars* is dispatched, the market where it is sold and the SKU to which it belongs. The *Input* columns represent the composition of each *pars* from out of four possible raw materials. We ask the reader to recall that a *pars* is represented by a triple consisting of information about the SKU, distribution channel and market where the product is sold. Therefore, Table 3 fully characterises our supply chain.

Assume that, as presented in the *basic configuration* and depicted in Figure 4, centre A distributes 75% of the production, and centre B is in charge of the remaining 25%. The weights vector for centre A is given by $\mathbf{p}^A = (3/5, 2/5)$;

Table 3: Input table for system S°. Rows are *pars* identifiers. Contribution to sales, content of each input and SKU are shown in the columns.

Pars				Input			
Number	Dist. Centre	Market	SKU	1	2	3	4
1	A	1	I	0.10	0.50	0.40	–
2	A	2	II	0.50	0.50	–	–
3	B	1	I	0.10	0.50	0.40	–
4	B	2	II	0.50	0.50	–	–
5	B	3	III	0.20	–	0.70	0.10

whereas the corresponding vector for centre B is $\mathbf{p}^B = (17/20, 1/10, 1/20)$.

In order to obtain the complexity of this system, we start by computing the complexity intrinsic to the transformation process associated with each *pars*. Let C_p^s represent the complexity brought by the production of the SKU with the identifier s; therefore, using (1) we can easily obtain $C_p^I = 1.361$; $C_p^{II} = 1$; and $C_p^{III} = 1.157$. Moreover, from the sales distribution, we can see that SKU I contributes to $3/4 \cdot 3/5 + 1/4 \cdot 17/20 = 53/80$ of the sales, SKU II with $13/40$ and SKU III with $1/80$. Therefore, the total complexity of the transformation process is given by the weighted sum of the SKUs' complexities, i.e. $1.361 \cdot 53/80 + 1 \cdot 13/40 + 1.157 \cdot 1/80 = 1.2411$.

Now, read loosely, Property 4 can be expressed by the following statement: *the total complexity of a cohort is given by the sum of the weighted complexities of the offspring plus the complexity of the parent.* Consequently, the complexity of the first level of system S° will be given by $C_p^{\circ} = 0.811 + 1.2411 = 2.0521$. Going one level downstream, we observe that the weighted complexity of systems A and B is equal to 0.915, consequently, the overall system's complexity becomes $C_p = 2.967$. See Figure 5 for a visual representation of these results.

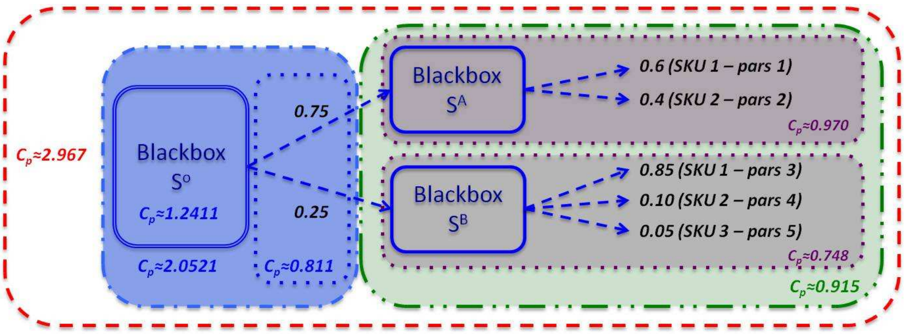


Figure 5: System S°: supply chain with a production facility. Overall system's complexity.

Alternatively, this system's complexity can be computed at the lowest level of aggregation by defining a vector whose elements are the contributions of each *pars* to the total sales, i.e. $\mathbf{p} = (0.75 \cdot 0.6 \cdot 0.1,\ 0.75 \cdot 0.6 \cdot 0.5, \ldots, 0.25 \cdot 0.05 \cdot 0.1)$, with *pars*-complexity $C_p(\mathbf{p}) = 2.967$. The last two examples suggest that the more we increase the vertical scope of our analysis, the higher will be the system's *pars*-complexity value.

3.2.8. Real -life example

For the sake of completeness we finish this section with a real-life illustrating example. Figure 6 depicts a simplified map of the supply chain of a manufacturing firm (actual values have been masked to protect confidentiality). The firm has one plant and central distribution centre (CDC) which serves six main distribution centres (MDC) serving the markets in Europe (EUR), China, Middle-East (MDE), the Far-East, NAFTA area, and Latin-America (LATAM). These centres, in turn, serve directly a number or regional stores; with the exception of China and the Far-East, which also serve two regional distribution centres (RDC) for South-East Asia and Japan. In total the firm has 1637 shops around the globe.

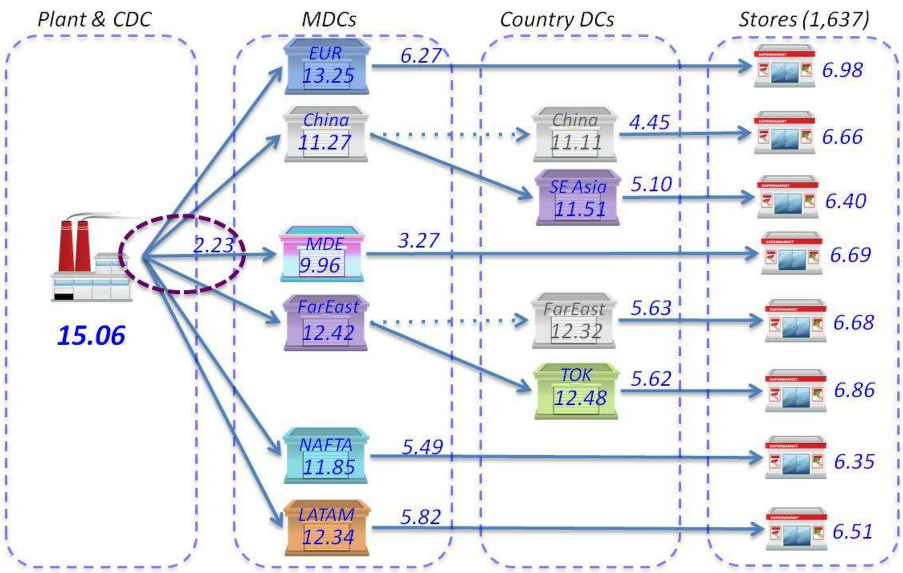


Figure 6: Supply chain complexity of a global firm.

The rightmost panel in Figure 6 shows the downstream *pars*-complexity of the stores served by each MDC. For example, the weighted *pars*-complexity of the shops served by the EUR centre is 6.98. If we add to this number the *pars*-complexity faced by the European MDC (6.27) we obtain the total complexity of the European market, i.e. 13.25. This rational can also be applied to the Middle-East, North-American and Latin-American markets.

Let's now focus on the China MDC. This distribution centre serves the stores located in China itself, plus a regional distribution centre, serving South-East Asia. The partial *pars*-complexity faced by the China MDC by serving the Chinese market is 11.11; whereas the *pars*-complexity faced by the South-East Asia RDC is 11.51. The weighted average of these two values gives the total *pars*-complexity of the market served by the China MDC, i.e. 11.27. The same logic is applied to the Far-East market.

Finally, the total *pars*-complexity faced by this firm (15.06) is the weighted average of the MDCs' *pars*-complexities ($p^{EUR}13.25 + \cdots + p^{LATAM}12.34$) plus that faced by the CDC (2.23).

3.3. Profiling pars-Complexity

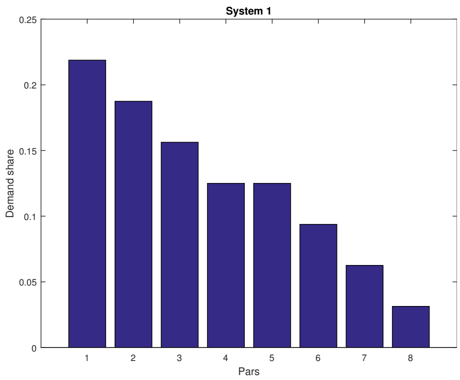
So far, we have computed *pars*-complexity in a number of situations that may arise when managing real-life supply chains. However, nothing has been said about its magnitude, i.e. when the observed *pars*-complexity is too high, neither how does the *pars*-complexity observed in certain system compares to the one observed in a different one (this becomes an issue when the systems being compared have different sizes). To be able to answer those questions we introduce the following measure for the length of the tails in a system's market profile:

Definition 2. *Let $C_p(S)$ be the pars-complexity of certain system S consisting of $|\mathcal{P}|$ partes. Then, the tails of the demand distribution of system S are characterised by*

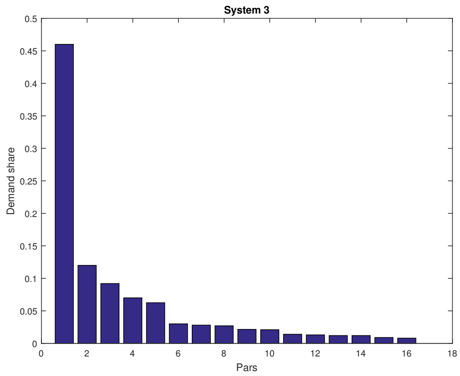
$$\widetilde{C}_p(S) = 1 - \frac{C_p(S)}{\log_2(|P|)} \tag{5}$$

From property 1.3 we have that the maximum complexity of system S is given by $\log_2|\mathcal{P}|$, therefore, the normalised version of C_p given by $\widetilde{C}_p$ is defined in the interval $[0, 1)$.

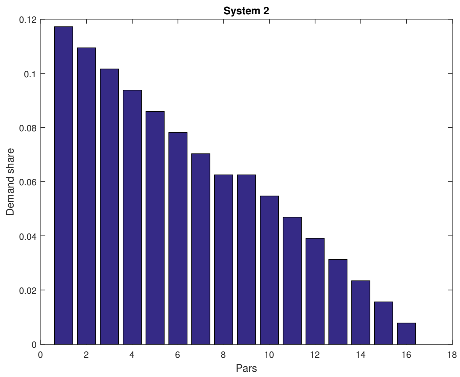
Notice that systems with shorter tails, i.e. a more uniform distribution of demand, will show values of $\widetilde{C}_p$ closer to zero, whereas systems with larger tails,



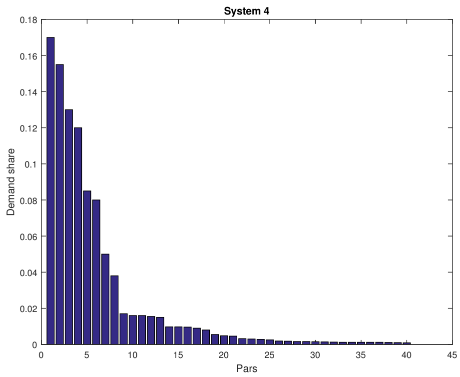
$C_p = 2.8273$, $\widetilde{C}_p = 0.0576$



$C_p = 2.8284$, $\widetilde{C}_p = 0.2929$



$C_p = 3.7784$, $\widetilde{C}_p = 0.0554$



$C_p = 3.7785$, $\widetilde{C}_p = 0.2900$

Figure 7: Four systems with different complexity profile

i.e. a larger number of *partes* with small contribution to demand, will return larger values of $\widetilde{C}_p$.

Figure 7 illustrates four different cases. Systems in the same row show the same value of *pars*-complexity but different tail lengths. This suggests that when high complexity is perceived as a problem, more uniform distributions may indicate a more acute problem than those with longer tails. Likewise, systems in the same column show the same value of $\widetilde{C}_p$ but different complexity, suggesting that systems showing different complexity levels may actually have equally heavy tails. Clearly, extracting conclusions about potential courses of action for reducing complexity can only be done with the aid of a thorough analysis of the relationship between complexity and costs or profits, an issue that goes out of the scope of this paper and is addressed in Menezes and Ruiz-Hernández (2019).

3.4. The Look-ahead property

Before finishing this section, it is convenient to make a remark about an important feature of *pars*-complexity: the look ahead property. As it has been mentioned earlier, one of the main attributes of the *pars*-complexity measure is the austerity in the use of information, requiring only information about the market share that each *pars* in the system represents. This means that a simple estimation of the changes in demand triggered by a managerial decision (*pars*-consolidation, expansion to a new market, release of a new product or withdrawal of a discontinued one, etc.) is enough for obtaining a quick estimation of the decision's impact on the firm's structural complexity.

This attribute, which is a fundamental attribute of our measure, cannot be found in many other measures of complexity and, when available, it relies on painstaking prediction of production, orders, suppliers, employees, customers, revenue, and other variables.

4. External Consistency

In order to test our complexity measure for external consistency we recall a series of empirical regularities or, what economists refer to as, stylised facts. The reasoning behind is that a consistent model of reality should be able to reproduce these facts. In our case, the consistency test consists of showing that the proposed measure for complexity does reproduce the expected impact on supply chain's structural complexity of certain strategic decisions. These decisions are: a) mergers and acquisitions; b) component commonality; c) modularity; d) postponement; and e) changes in the firm's product mix.

4.1. Mergers and acquisitions

Stylised Fact: Notwithstanding the consolidation effort, mergers do not reduce the system's average complexity (for recent discussion on this fact please see Herd and McClelland (2017) or Aries and Hu (2009), among others).

Consider two merging firms A and B. Let π_A and π_B be r-dimensional vectors representing the absolute value of each firm' sales in r different markets. Notice that the number of markets where each firm is present is not necessarily equal to r. Indeed, if after the merger no market consolidation is conducted,

$r = r_A + r_B$ and $\pi_A = (s_A \oplus \mathbf{0}_{r_B})$; where s_A is an r_A-vector representing the absolute value of firm A's sales and $\mathbf{0}_{r_B}$ an r_B-vector of zeros; correspondingly, $\pi_B = (\mathbf{0}_{r_A} \oplus s_B)$. Using $|\mathbf{x}|_1$ as a notational shortcut representing the sum of all elements in $\mathbf{x}$, we can see that $\mathbf{p}_F = \frac{\pi_F}{|\pi_F|_1}$, for $F \in \{A, B\}$. Assume also that firm A's participation in the extended market is represented by the shares vector λ; correspondingly, firm B's shares are given by $(1-\lambda)$.

Proposition 3. *The pars-complexity of the merged system C_p^M is larger or equal than the weighted sum of the pars-complexities of the two individual systems before the merger C_p°; i.e. $C_p^M \geq C_p^\circ$.*

Proof *For given values of $\mathbf{p}^A = \frac{\pi_A}{|\pi_A|_1}$, $\mathbf{p}^B = \frac{\pi_B}{|\pi_B|_1}$, and the corresponding λ, we have that*

$$\lambda \frac{\pi_A}{|\pi_A|_1} + (1-\lambda)\frac{\pi_B}{|\pi_B|_1} = \frac{(\pi_A + \pi_B)}{|(\pi_A + \pi_B)|_1}$$

therefore, after computing the complexity measure on both sides of the expression above, we get

$$C_p\left(\frac{(\pi_A + \pi_B)}{|(\pi_A + \pi_B)|_1}\right) = C_p\left(\lambda \frac{\pi_A}{|\pi_A|_1} + (1-\lambda)\frac{\pi_B}{|\pi_B|_1}\right) \geq \lambda C_p\left(\frac{\pi_A}{|\pi_A|_1}\right) + (1-\lambda)\, C_p\left(\frac{\pi_B}{|\pi_B|_1}\right) \tag{6}$$

which follows from Property 5 and is a direct consequence of Property 1.1 . ■

This property establishes, in concordance with the stylised fact, that the *pars*-complexity of a merged firm is always larger than the weighted *pars*-complexity of its predecessors. If, otherwise, the merging firms take advantage of existing synergies and their supply chains are consolidated, Proposition 4 shows that there is still certain margin for reducing the merged system's *pars*-complexity.

Before presenting our result, it is worth noticing that if both firms were present in exactly the same markets, then $r = r_A = r_B$, $\pi_A = (s_A)$, and $\pi_B = (s_B)$; in any other case, r will represent the total number of markets where either A, B, or both were selling their products. Moreover, π_A (equivalently π_B) will be an r-vector with zeros in the positions representing markets where A (respectively B) had no presence before the merger. With these elements we can now introduce the following result:

Proposition 4. *Let $\mathcal{N}_A$ and $\mathcal{N}_B$ represent the sets of* partes *in systems A and B, respectively, before the merger. Let vector $\mathbf{p}_{AB}$ be a weights s-vector, where $s = |\mathcal{N}_A \cup \mathcal{N}_B|$, whose elements are given by the expression $\mathbf{p}_{AB,i} = \lambda \mathbf{p}_{A,i} + (1-\lambda)\, \mathbf{p}_{B,i}$, for $i = 1, \ldots, s$. Moreover, assume –without any loss of generality- that for any market $i = 1, \ldots, r$ and firm $F \in \{A, B\}$ such that*

$p_{F,i} = 0$, $\log_2\left(\frac{1}{p_{F,i}}\right) = 0$. *Then*

$$C_p\left(\left(\lambda \mathbf{p}_A + (1-\lambda)\,\mathbf{p}_B\right)\right) \geq C_p\left(\mathbf{p}_{AB}\right) \geq \lambda C_p\left(\mathbf{p}_A\right) + (1-\lambda)\,C_p\left(\mathbf{p}_B\right) \quad (7)$$

Proof *The first inequality in expression* (7) *is a direct consequence of Proposition 6. For the second one, it is enough to develop the last two terms in* (7)*, i.e.*

$$\begin{aligned} C_p\left(\mathbf{p}_{AB}\right) &= \sum_{i=1}^{s} p_{AB,i}\log_2\left(\frac{1}{p_{AB,i}}\right) = \sum_{i=1}^{s}\left(\lambda p_{A,i} + (1-\lambda)\,p_{B,i}\right)\log_2\left(\frac{1}{\lambda p_{A,i} + (1-\lambda)\,p_{B,i}}\right) \\ &\geq \lambda\sum_{i=1}^{s} p_{A,i}\log_2\left(\frac{1}{p_{A,i}}\right) + (1-\lambda)\sum_{i=1}^{s} p_{B,i}\log_2\left(\frac{1}{p_{B,i}}\right) = \lambda C_p\left(\mathbf{p}_A\right) + (1-\lambda)\,C_p\left(\mathbf{p}_B\right) \end{aligned}$$

for noticing that, due to the concavity of C_p (Property 1.1), the inequality will hold for any admissible combination of the parameters, i.e. for $p_{A,i}, p_{B,i} \in [0,1]$, $i = 1,\ldots,s$ and $\lambda \in (0,1)$. ■

Propositions 3 and 4 confirm that the *pars*-complexity measure C_p does indeed reflect the stylised fact: it does not matter how big the consolidation effort for taking advantage of synergies is, the *pars*-complexity of the merged enterprise will never be smaller than the weighted-average *pars*-complexity of the original firms.

4.2. Component Commonality

Stylised fact: Notwithstanding component commonality decreases inventory requirements; it does not affect downstream structural complexity. However, if up-stream stages are taken into account, then the overall system's complexity is indeed reduced (for discussion on this fact see Cook (2001); Blecker and Abdelkafi (2007); Abdelkafi (2008); Wazed et al. (2009); Bernstein et al. (2011); Weiser et al. (2016) and Kaivola (2017), among others).

Before starting our discussion, we introduce a generic system $S^{\bullet}$ that will be of use in the analysis of this and subsequent stylised facts. Consider a firm that produces N different SKUs using M different inputs or components. We refer to this firm's supply chain as system $S^{\bullet}$. For the sake of simplicity we assume that each SKU is sold in only one market, and therefore the number of *partes* in the system equals the number of SKUs [4] Moreover, and without any loss

[4]Results in this section can easily be extended to cases where the same SKU can be sold in different markets.

of generality, we assume that the SKUs are distributed from the factory to H different markets through a number F of distribution facilities. System $S^\bullet$ can, therefore, be characterised by the tuple $\{\mathbf{p}^\bullet, \mathbf{p}^\circ, \mathbf{p}^+, \mathbf{q}\}$, where $\mathbf{p}^\bullet$ represents the F-vector of weights of the production facility; $\mathbf{p}^\circ$ is an $F \times H$ matrix whose rows are weights vectors corresponding to the distribution facilities; $\mathbf{p}^+$ is a $F \times H \times N$ matrix with rows representing the weights vector for each market; finally, $\mathbf{q}$ is an $N \times M$ matrix representing the composition of each SKU in terms of its inputs and components. Each row $\mathbf{q}^i$ in $\mathbf{q}$, is an $M-$vector representing the composition of *pars* i, for $i = 1, \ldots, N$.

In order to establish our main result recall system $S^\bullet$ and notice that whenever input k, for $k = 1, \ldots, M$, is not used for the production of *pars* i, $\mathbf{q}^i_{(k)} = 0$. Therefore, as shown in illustrative example 3.2.7, the complexity of $S^\bullet$ can be readily computed as

$$C_p\left(S^\bullet\right) = \sum_{i=1}^{N} \overline{\mathbf{p}}_i^\bullet \sum_{k=1}^{M} \mathbf{q}^i_{(k)} \log\left(\frac{1}{\mathbf{q}^i_{(k)}}\right) + \sum_{j=1}^{F} \mathbf{p}^\bullet_{(j)} C_p\left(s^\circ_j\right) + C_p\left(\mathbf{p}^\bullet\right) \qquad (8)$$

where $\overline{\mathbf{p}}_i^\bullet = \sum_{j=1}^{F} \sum_{h=1}^{H} \mathbf{p}^\bullet_j \mathbf{p}^\circ_{j,(h)} \mathbf{p}^+_{j,h,(i)}$ for all $i = 1, \ldots, N$ and s°_j represents the subsystem starting from the second level of the supply chain downwards.

Suppose now that a subset $\mathcal{C}$ of mutually exclusive inputs (i.e. only one of them can be used as input for *pars* i) with $|\mathcal{C}| < M$, is replaced by a common component. This defines a new system $S^\star$. For the sake of simplicity, let us assume that the inputs to be substituted are located in the first $|\mathcal{C}|$ positions of vectors $\mathbf{q}^i$, $i = 1 \ldots, N$. Notice that this defines a new family of composition vectors, each of them with $(M - |\mathcal{C}| + 1)$ elements, represented by $\overline{\mathbf{q}}^i = \left(\sum_{j=1}^{|\mathcal{C}|} \mathbf{q}^i_{(j)}, \mathbf{q}^i_{(|\mathcal{C}|+1)}, \ldots, \mathbf{q}^i_{(M)}\right)$. The *pars*-complexity of the new system, can be readily obtained by substituting the first term in the right hand side of equation (8) for:

$$\sum_{j=1}^{N} \overline{\mathbf{p}}_i^\bullet \sum_{k=1}^{M-|\mathcal{C}|+1} \overline{\mathbf{q}}^i_{(k)} \log\left(\frac{1}{\overline{\mathbf{q}}^i_{(k)}}\right). \qquad (9)$$

The following proposition establishes that the proposed measure reproduces the stylised fact that component commonality does not affect downstream *pars*-complexity:

Proposition 5. *The pars-complexities of systems $S^{\bullet}$ and $S^{\star}$, which only differ in their technology vectors $\mathbf{q}^i$, $i = 1, \ldots, N$ and $\overline{\mathbf{q}}^i$, $i = 1, \ldots, N$, are identical.*

Proof *Notice that the pars-complexity equation of systems $S^{\bullet}$ and $S^{\star}$ only differs in the first term on the r.h.s. of equation* (8)*. Moreover, given the fact that $\overline{\mathbf{q}}^i_{(1)} = \sum_{j=1}^{|\mathcal{C}|} \mathbf{q}^i_{(j)}$ for all $i = 1, \ldots, N$, the result follows immediately from the fact that*

$$\sum_{k=1}^{M} \mathbf{q}^i_{(k)} \log \left(\frac{1}{\mathbf{q}^i_{(k)}} \right) = \sum_{k=1}^{M-|\mathcal{C}|+1} \overline{\mathbf{q}}^i_{(k)} \log \left(\frac{1}{\overline{\mathbf{q}}^i_{(k)}} \right), \quad \forall\, i = 1, \ldots, N$$

■

Notwithstanding the previous result, it is important to notice that if the scope of the analysis is expanded to including one more up-stream level in the supply chain, i.e. the supply of raw materials, then the reduction in *pars*-complexity induced by component commonality is reflected by the proposed measure. In order to see this, we introduce the expanded systems $\widetilde{S}^{\bullet}$ and $\widetilde{S}^{\star}$, which extend the original ones by including the input's procurement stage. With these elements, the stylised fact can be expressed as

$$C_p\left(\widetilde{S}^{\bullet}\right) \geq C_p\left(\widetilde{S}^{\star}\right) . \tag{10}$$

Introducing the input vectors $\widetilde{\mathbf{q}}^{\bullet}$ and $\widetilde{\mathbf{q}}^{\star}$ (where $\widetilde{\mathbf{q}}^{\bullet}_{(k)} = \sum_{i=1}^{N} \mathbf{q}^i_{(k)}$, for $k = 1, \ldots, M$; and $\widetilde{\mathbf{q}}^{\star}_{(k)} = \sum_{i=1}^{N} \overline{\mathbf{q}}^i_{(k)}$, with $k = 1, \ldots, M - |\mathcal{C}| + 1$), and given that the *pars*-complexity of systems $\widetilde{S}^{\bullet}$ and $\widetilde{S}^{\star}$ is equal -in each case- to the *pars*-complexity of the lower level system plus the *pars*-complexity of the supply, we can express inequality (10) as:

$$C_p\left(S^{\bullet}\right) + C_p\left(\widetilde{\mathbf{q}}^{\bullet}\right) \geq C_p\left(S^{\star}\right) + C_p\left(\widetilde{\mathbf{q}}^{\star}\right) . \tag{11}$$

Using the result in Proposition 5 together with equation (1), developing the sums and after eliminating common terms we get that inequality (11), and therefore expression (10), reduces to

$$\sum_{k=1}^{|\mathcal{C}|} \widetilde{\mathbf{q}}^{\bullet}_{(k)} \log_2 \left(\frac{1}{\widetilde{\mathbf{q}}^{\bullet}_{(k)}} \right) \geq \widetilde{\mathbf{q}}^{\star}_{(1)} \log_2 \left(\frac{1}{\widetilde{\mathbf{q}}^{\star}_{(1)}} \right) \tag{12}$$

which is a direct consequence of Property 5.

4.3. Modularity

Stylised fact: Modularity, seen an extension of the notion of component commonality where not only one but several –not necessarily exclusive- components are substituted for one single module, as well as providing flexibility and reducing inventory, decreases the supply chain's structural complexity (for discussion in this subject please refer to Baldwin and Clark (1997); Da Silveira et al. (2001); Blecker and Abdelkafi (2006b); Marti (2007); Golfmann and Lammers (2015) and Kaivola (2017)).

In order to assess the capability of the complexity measure for reflecting this stylised fact, we introduce a minor modification in system $S^{\bullet}$ product design (where $S^{\bullet} = \{\mathbf{p}^{\bullet}, \mathbf{p}^{\circ}, \mathbf{p}^{+}, \mathbf{q}\}$). Assume that each *pars* uses one out of $|\mathcal{C}|$ exclusive inputs and all of them use a number $|\mathcal{D}|$ of common inputs. Assume that $|\mathcal{C}| + |\mathcal{D}| \leq M$. As discussed before, the complexity of this system can be computed as the sum of the transformation and distribution complexities as given by equation (8). Suppose now that a module is introduced which integrates the attributes of the $|\mathcal{C}|$ exclusive components together with the ones of the $|\mathcal{D}|$ of common inputs. This module is to be used in the production of all *partes*. If, for the sake of simplicity, we assume that these inputs are located in the first $|\mathcal{C}| + |\mathcal{D}|$ positions of vectors $\mathbf{q}^i$, then the introduction of the module defines a new family of composition vectors $\widehat{\mathbf{q}}^i = \left(\sum_{j=1}^{|\mathcal{C}|+|\mathcal{D}|} \mathbf{q}^i_{(j)}, \mathbf{q}^i_{(|\mathcal{C}|+|\mathcal{D}|+1)}, \ldots, \mathbf{q}^i_{(M)}\right)$, for all $i = 1, \ldots, N$. The *pars*-complexity of this new system, $S^{\star}$, can be computed as

$$C_p\left(S^{\star}\right) = \sum_{j=1}^{N} \overline{\mathbf{p}}_i^{\bullet} \sum_{k=1}^{M-|\mathcal{C}|-|\mathcal{D}|+1} \widehat{\mathbf{q}}^i_{(k)} \log\left(\frac{1}{\widehat{\mathbf{q}}^i_{(k)}}\right) + \sum_{j=1}^{N} \mathbf{p}^{\bullet}_{(j)} C_p\left(s_j^{\circ}\right) + C_p\left(\mathbf{p}^{\bullet}\right) . \tag{13}$$

The following proposition establishes the fact that our measure reflects the stylised fact that modularity reduces structural complexity:

Proposition 6. *Keeping everything else constant, introducing modularity in the production process reduces pars-complexity. Namely, whenever the technology vectors,* $\mathbf{q}^i$ *and* $\widehat{\mathbf{q}}^i$*, of systems* $S^{\bullet}$ *and* $S^{\star}$*, respectively, satisfy* $\widehat{\mathbf{q}}^i_{(1)} = \sum_{j=1}^{|\mathcal{C}|+|\mathcal{D}|} \mathbf{q}^i_{(j)}$ *and* $\widehat{\mathbf{q}}^i_{(h)} = \mathbf{q}^i_{(|\mathcal{C}|+|\mathcal{D}|+h-1)}$ *for* $i = 1, \ldots, N$ *and* $h = 2, \ldots, M - |\mathcal{C}| - |\mathcal{D}| + 1$*, then* $C_p\left(S^{\bullet}\right) \geq C_p\left(S^{\star}\right)$*.*

Proof *Given equations* (8) *and* (13) *we have that for establishing* $C_p\left(S^\bullet\right) \geq C_p\left(S^\star\right)$ *it suffices to proof that*

$$\sum_{k=1}^{M} \mathbf{q}^i_{(k)} \log_2 \left(\frac{1}{\mathbf{q}^i_{(k)}}\right) \geq \sum_{h=1}^{M-|\mathcal{C}|-|\mathcal{D}|+1} \widehat{\mathbf{q}}^i_{(h)} \log_2 \left(\frac{1}{\widehat{\mathbf{q}}^i_{(h)}}\right), \quad \forall\, i = 1, \ldots, N$$

equivalently

$$\sum_{k=1}^{|\mathcal{C}|+|\mathcal{D}|} \mathbf{q}^i_{(k)} \log_2 \left(\frac{1}{\mathbf{q}^i_{(k)}}\right) \geq \widehat{\mathbf{q}}^i_{(1)} \log_2 \left(\frac{1}{\widehat{\mathbf{q}}^i_{(1)}}\right), \quad \forall\, i = 1, \ldots, N$$

which is a direct consequence of Property 5. ■

4.4. Postponement

Stylised fact: Keeping the number of SKUs constant, postponement reduces complexity locally (at certain given stage of the supply chain) but the overall system's complexity remains the same (for discussion on this fact please see Cook (2001); Blecker and Abdelkafi (2006a); Abdelkafi (2008), and Brunoe and Nielsen (2016)).

In order to establish that our measure is indeed consistent with this stylised fact, let us recall our baseline system $S^\bullet = \{\mathbf{p}^\bullet, \mathbf{p}^\circ, \mathbf{p}^+, \mathbf{q}\}$ and assume that transformation is conducted at the top level of the supply chain. System $S^\bullet$ *pars*-complexity is given by equation (8). Consider now an alternative design of the supply chain, $S^\star$, where transformation is performed at a lower level of the supply chain. In order to obtain an expression for the *pars*-complexity of this new system, we first develop a version of equation (8) valid for all intermediate facilities $j = 1, \ldots, F$:

$$C_p\left(s^\star_j\right) = \sum_{i=1}^{N} \overline{\mathbf{p}}^\circ_i \sum_{k=1}^{M} \mathbf{q}^i_{(k)} \log_2 \left(\frac{1}{\mathbf{q}^i_{(k)}}\right) + \sum_{h=1}^{H} \mathbf{p}^\circ_{j,(h)} C_p\left(s^+_{j,h}\right) + C_p\left(\mathbf{p}^\circ_j\right) ; \quad (14)$$

where $\overline{\mathbf{p}}^\circ_i = \sum_{h=1}^{H} \mathbf{p}^\circ_{j,(h)} \mathbf{p}^+_{j,h,(i)}$ for all $i = 1, \ldots, N$. With these elements, the *pars*-complexity of system $S^\star$ can now be expressed as

$$C_p\left(S^\star\right) = \sum_{j=1}^{F} \mathbf{p}^\bullet_{(j)} C_p\left(s^\star_j\right) + C_p\left(\mathbf{p}^\bullet\right) . \quad (15)$$

The following two results establish the main claim of this section, namely, that the *pars*-complexity measure reflects the stylised fact that postponement displaces the transformation complexity to the stage where production is conducted.

Proposition 7. *The stage in the supply chain where transformation is conducted does not affect its overall pars-complexity value, i.e.* $C_p\left(S^{\bullet}\right)=C_p\left(S^{\star}\right)$.

Proof *Equation* (15) *can be developed as*

$$C_p\left(S^{\star}\right)=\sum_{j=1}^{F}\mathbf{p}^{\bullet}_{(j)}\left[\sum_{i=1}^{N}\overline{\mathbf{p}}^{\circ}_{i}\sum_{k=1}^{M}\mathbf{q}^{i}_{(k)}\log_2\left(\frac{1}{\mathbf{q}^{i}_{(k)}}\right)+\sum_{h=1}^{H}\mathbf{p}^{\circ}_{j,(h)}C_p\left(s^{+}_{j,h}\right)+C_p\left(\mathbf{p}^{\circ}_{j}\right)\right]+C_p\left(\mathbf{p}^{\bullet}\right)$$

which simplifies to

$$=\sum_{j=1}^{F}\mathbf{p}^{\bullet}_{(j)}\left[\sum_{i=1}^{N}\overline{\mathbf{p}}^{\circ}_{i}\sum_{k=1}^{M}\mathbf{q}^{i}_{(k)}\log_2\left(\frac{1}{\mathbf{q}^{i}_{(k)}}\right)\right]+\sum_{j=1}^{F}\mathbf{p}^{\bullet}_{(j)}C_p\left(s^{\circ}_{j}\right)+C_p\left(\mathbf{p}^{\bullet}\right)\ ;$$

upon substitution of $\overline{\mathbf{p}}^{\circ}_{i}$ *in the previous expression we obtain*

$$=\sum_{j=1}^{F}\mathbf{p}^{\bullet}_{(j)}\sum_{i=1}^{N}\sum_{h=1}^{H}\mathbf{p}^{\circ}_{j,(h)}\mathbf{p}^{+}_{j,h,(i)}\sum_{k=1}^{M}\mathbf{q}^{i}_{(k)}\log_2\left(\frac{1}{\mathbf{q}^{i}_{(k)}}\right)+\sum_{j=1}^{F}\mathbf{p}^{\bullet}_{(j)}C_p\left(s^{\circ}_{j}\right)+C_p\left(\mathbf{p}^{\bullet}\right)$$

finally, after reordering and collecting terms, we have

$$=\sum_{i=1}^{N}\overline{\mathbf{p}}^{\bullet}_{i}\sum_{k=1}^{M}\mathbf{q}^{i}_{(k)}\log_2\left(\frac{1}{\mathbf{q}^{i}_{(k)}}\right)+\sum_{j=1}^{F}\mathbf{p}^{\bullet}_{(j)}C_p\left(s^{\circ}_{j}\right)+C_p\left(\mathbf{p}^{\bullet}\right)=C_p\left(S^{\bullet}\right)$$

■

A direct consequence of this result is that transformation complexity is accounted for at the transformation stage:

Corollary 1. *Postponement displaces pars-complexity to the supply chain's stage or level where transformation is conducted.*

Proof *It has already been established that the pars-complexity at any stage of the supply chain can be expressed as the sum of the pars-complexity of the parent system,* $C_p(\mathbf{p})$*, plus the weighted average of the offspring's pars-complexities. If, additionally, the production is performed at that precise level, the total pars-complexity of the stage will be that of the parent system plus the transformation complexity,* $C_p(\mathbf{q})$*. Therefore, it follows from Proposition 7 that if transformation is postponed from central production to some other stage, the parent's pars-complexity will reduce from* $C_p(\mathbf{p})+C_p(\mathbf{q})$ *to* $C_p(\mathbf{p})$*; whereas the pars-complexity of the new transformation stage will increase from* $C_p(\mathbf{p}')$ *to* $C_p(\mathbf{p}')+C_p(\mathbf{q})$. ■

4.5. Product Mix

Stylised Fact: Eliminating products from the firm's product mix may contribute to reducing the overall system's complexity (see, for example, Blecker et al. (2004); Blecker and Abdelkafi (2006a); Anderson et al. (2007); Marti (2007); Aries and Hu (2009), and Hirose et al. (2017), among others).

Example 3.2.5 illustrates a case that reflects this stylised fact: removing (adding) a *pars* reduces (increases) the value of C_p. However, notwithstanding there is a general agreement among practitioners and academics that the withdrawal of a *pars* from the supply chain will reduce its structural complexity, there is one important caveat: If the product to be withdrawn represents a significantly large proportion of the firm's market, the structural complexity of the supply chain may increase. This stems from the fact that, when removing a big seller, the focal point shifts from the operations of one SKU to those of many different ones, which previously were irrelevant but now play a major role in the new setting. Therefore, the inadvertent withdrawal of a critical *pars* may have the unexpected consequence of increasing the supply chain's structural complexity instead of reducing it as intended.

The following result provides a sufficient condition on the *pars*' market share for *pars*-complexity to be reduced when the pars is removed from the product mix:

Proposition 8. *Given a system $S^{\bullet}$ with weights vector $\mathbf{p}$, if* pars *j for some $j = 1, \ldots, N$ is such that $\mathbf{p}_{(j)} \leq \frac{1}{N}$, then withdrawing* pars *j from the market will decrease the system's* pars*-complexity with certainty.*

Proof *Consider system $S^{\bullet}$ and assume, without any loss of generality, that the* pars *to be withdrawn is* pars *N. Assume that the remaining system is given by $S^{\star}$ with weights vector $\mathbf{q}$ where $\mathbf{q}_{(j)} = \frac{\mathbf{p}_{(j)}}{1-\mathbf{p}_{(N)}}$, $j = 1, \ldots, N-1$.*

For establishing $C_p(S^{\bullet}) \geq C_p(S^{\star})$ we need to show that

$$\sum_{j=1}^{N-1} \mathbf{p}_{(j)} \log_2 \left(\frac{1}{\mathbf{p}_{(j)}} \right) + \mathbf{p}_{(N)} \log_2 \left(\frac{1}{\mathbf{p}_{(N)}} \right) \geq \sum_{j=1}^{N-1} \frac{\mathbf{p}_{(j)}}{1 - \mathbf{p}_{(N)}} \log_2 \left(\frac{1 - \mathbf{p}_{(N)}}{\mathbf{p}_{(j)}} \right)$$

whenever $\mathbf{p}_{(N)} \leq \frac{1}{N}$. This inequality can be reduced to

$$\left(1 - \frac{1}{1 - \mathbf{p}_{(N)}}\right) \sum_{i=1}^{N-1} \mathbf{p}_{(j)} \log_2 \left(\frac{1}{\mathbf{p}_{(j)}} \right) \geq \log_2 \left(1 - \mathbf{p}_{(N)}\right) - \mathbf{p}_{(N)} \log_2 \left(\frac{1}{\mathbf{p}_{(N)}} \right) \tag{16}$$

Defining $x = 1 - \mathbf{p}_{(N)}$ and given that $\left(1 - \frac{1}{x}\right) \leq 0$ whenever $x \in (0,1]$, the l.h.s. of equation (16) *minimises whenever $\mathbf{p}_{(j)} = \frac{1-\mathbf{p}_{(N)}}{N-1} = \frac{x}{N-1}$. Hence, after substitution and upon simplification, we have that* (16) *will hold whenever*

$$-(1-x) \log_2 \left(\frac{N-1}{x} \right) \geq \log_2 (x) - (1-x) \log_2 \left(\frac{1}{1-x} \right)$$

which is satisfied whenever x is such that

$$-\log_2\left(\frac{N-1}{x}\right) \geq \log_2(1-x)$$

or, equivalently,

$$x \geq \frac{N-1}{N} \quad \Rightarrow \quad \mathbf{p}_{(N)} \leq \frac{1}{N}$$

■

Notwithstanding this condition can be used as a general rule of thumb, in many situations it may be preferable to have an exact threshold for the value at which the withdrawal of a *pars* with weight $\mathbf{p}_{(j)}$ will surely reduce supply chain's structural complexity. Moreover, given that Proposition 8 provides only a sufficient condition, there will (almost certainly) be cases where the withdrawal of a *pars* j such that $\mathbf{p}_{(j)} \gg \frac{1}{N}$ will still decrease the system's *pars*-complexity. Proposition 9 provides such threshold.

Proposition 9. *If a pars $\mathbf{p}_{(j)}$ is withdrawn from system $S^{\bullet}$ such that $\mathbf{p}_{(j)} \leq \mathbf{p}^*$, where $\mathbf{p}^*$ is the unique solution to*

$$C_p(\mathbf{p}) = \log_2\left(\frac{1}{\mathbf{p}^*}\right) - \frac{1-\mathbf{p}^*}{\mathbf{p}^*}\log_2(1-\mathbf{p}^*)\ , \tag{17}$$

the system's complexity decreases.

Proof *Assume, without loss of generality, that the* pars *withdrawn occupies position N in $\mathbf{p}$. Let $S^{\star}$ be the resulting (reduced) system and $\mathbf{q}$ its weights vector; then the condition $C_p(S^{\bullet}) \geq C_p(S^{\star})$ is given by equation (16), i.e.*

$$\left(1-\frac{1}{1-\mathbf{p}_{(N)}}\right)\sum_{i=1}^{N-1}\mathbf{p}_{(j)}\log_2\left(\frac{1}{\mathbf{p}_{(j)}}\right) \geq \log_2\left(1-\mathbf{p}_{(N)}\right) - \mathbf{p}_{(N)}\log_2\left(\frac{1}{\mathbf{p}_{(N)}}\right)$$

reordering terms we have

$$\mathbf{p}_{(N)}\log_2\left(\frac{1}{\mathbf{p}_{(N)}}\right) \geq \log_2\left(1-\mathbf{p}_{(N)}\right) + \left(\frac{\mathbf{p}_{(N)}}{1-\mathbf{p}_{(N)}}\right)\sum_{j=1}^{N-1}\mathbf{p}_j\log_2\left(\frac{1}{\mathbf{p}_{(j)}}\right)$$

using $C_p(\mathbf{p}) = \sum_{j=1}^{N-1}\mathbf{p}_{(j)}\log_2\left(\frac{1}{\mathbf{p}_{(j)}}\right) + \mathbf{p}_{(N)}\log_2\left(\frac{1}{\mathbf{p}_{(N)}}\right)$ we obtain

$$\left(\frac{1}{1-\mathbf{p}_{(N)}}\right)\mathbf{p}_{(N)}\log_2\left(\frac{1}{\mathbf{p}_{(N)}}\right) \leq \log_2\left(1-\mathbf{p}_{(N)}\right) + \left(\frac{\mathbf{p}_{(N)}}{1-\mathbf{p}_{(N)}}\right)C_p(\mathbf{p})$$

which, after reordering terms, gives the inequality

$$C_p(\mathbf{p}) \geq \log_2\left(\frac{1}{\mathbf{p}_{(N)}}\right) - \frac{1-\mathbf{p}_{(N)}}{\mathbf{p}_{(N)}}\log_2\left(1-\mathbf{p}_{(N)}\right) .$$

That $\mathbf{p}^*$ is the unique $\mathbf{p}_{(N)}$ solution to the last inequality follows directly from the fact that the expression in the r.h.s. is strictly decreasing for $\mathbf{p}_{(N)} \in (0,1]$. ■

Mirroring the previous discussion we notice that, even though it is widely accepted that *pars* proliferation increases structural complexity, there are many cases where increasing the firm's product mix may still be profitable. Consequently, it may be convenient to know whether or not the release of a new product will increase the system's *pars*-complexity. The following proposition provides such threshold in the form of the new *pars*'s contribution to total sales.

Proposition 10. *If a pars $\mathbf{p}_{(N+1)}$ is added to system $S^\bullet$ such that $\mathbf{p}_{(N+1)} \geq \overline{\mathbf{p}}$, where $\overline{\mathbf{p}}$ is the unique solution to*

$$C_p(\mathbf{p}) = \frac{1}{\overline{\mathbf{p}}}\log_2(1+\overline{\mathbf{p}}) + \log_2\left(\frac{1+\overline{\mathbf{p}}}{\overline{\mathbf{p}}}\right) , \tag{18}$$

the system's structural complexity decreases.

Proof *By defining $\mathbf{q}_{(j)} = \frac{\mathbf{p}_{(j)}}{1+\widetilde{\mathbf{p}}}$, $\mathbf{q}_{(N+1)} = \frac{\widetilde{\mathbf{p}}}{1+\widetilde{\mathbf{p}}}$, and developing the inequality $C_p(\mathbf{q}) \leq C_p(\mathbf{p})$, i.e.*

$$\sum_{j=1}^{N+1} \mathbf{q}_{(j)}\log_2\left(\frac{1}{\mathbf{q}_{(j)}}\right) \leq \sum_{j=1}^{N} \mathbf{p}_{(j)}\log_2\left(\frac{1}{\mathbf{p}_{(j)}}\right)$$

and working along the same lines as that of Proposition 9, we get

$$C_p(\mathbf{p}) \geq \frac{1}{\overline{\mathbf{p}}}\log_2(1+\overline{\mathbf{p}}) + \frac{\widetilde{\mathbf{p}}}{\overline{\mathbf{p}}}\log_2\left(\frac{1+\overline{\mathbf{p}}}{\widetilde{\mathbf{p}}}\right)$$

which only holds when $\widetilde{\mathbf{p}} \geq \overline{\mathbf{p}}$. This completes the proof. ■

5. Conclusion

The focus of this article is the problem of complexity in supply chain management. The degree in which proliferation of products, markets and channels is permeating in the business environment is dubbed by practitioners as a *complexity crisis*. It is actually affirmed that complexity damages the tissue of organisations' supply chains, rendering them less efficient.

Complexity is... complex, and as such there is a strong need for a proper definition of what is meant when talking about supply chain complexity. Notwithstanding there is a vast amount of literature addressing the problem of complexity, its costs and potential solutions, this is the first time, to our knowledge, that a definition of complexity –which emanates from the business strategy- is provided. This definition, in turn, allows us to bring the discussion on supply chain complexity to a quantitative framework, providing a direct and austere measure that can be used at high-level decision making.

The article starts with a discussion on supply chain complexity and introduces the notion of *structural complexity*, a numerousness related concept associated with the proliferation of products, channels and markets. We then develop the theoretical framework for its measurement based on the so-called *pars*-complexity measure, in information-content based measure of proliferation that stems from the Theory of Communication. We show that this measure satisfies two fundamental requirements for the scientific validity of any representation of the real world: internal and external consistency. We argue that the internal consistency is inherited from Shannon's Theory of Communication (whose measure of information content is maintained intact), and present a collection of theoretical properties that highlight attributes which are desirable in a good measure for structural complexity. The discussion on external consistency is based on certain *stylised facts* on supply chain management that should be familiar for both practitioners and academics. The underlying reasoning is that a good model of the reality should be consistent with the world it is aimed at representing. We show that the results emanating from our representation of structural complexity are capable of reproducing those empirically observed regularities.

A good theory of the physical word should not only be rigorous, but also relevant, i.e. useful for improved action and problem solving. Parallel work in Menezes and Ruiz-Hernández (2019) establishes the empirical validity of the *pars*-complexity measure by applying it to the study of twenty-seven business units of a large firm. In particular, we analyse the relation existing between our notion of structural complexity and some financial indicators of the firm, finding a particularly strong link between the proposed measure and the operating margin. Moreover, we rise arguments for the practical relevance of the measure as a tool for allocating costs among business units, on one hand, and among

SKUs, on the other.

We are convinced that, by proposing a definition of structural complexity, providing a measure for quantifying it, and developing a solid framework for the mathematics of *pars*-complexity, this article is an important step towards the understanding, and therefore solution, of the complexity crisis.

References

Abdelkafi, N. (2008). *Variety Induced Complexity in Mass Customization: Concepts and Management*, volume 7. Erich Schmidt Verlag.

Adams, C., Alldredge, K., Mueller, C., and Whitmore, J. (December, 2016). Simpler is (sometimes) better: Managing complexity in consumer goods. McKinsey & Co. INC.

Aitken, J., Bozarth, C., and Garn, W. (2016). To eliminate or absorb supply chain complexity: a conceptual model and case study. *Supply Chain Management: An International Journal*, 21(6):759–774.

Allesina, S., Azzi, A., Battini, D., and Regattieri, A. (2010). Performance measurement in supply chains: new network analysis and entropic indexes. *International Journal of Production Research*, 48(8):2297–2321.

Anderson, B., Hagen, C., and Reifel, J. (2007). Battling the beast: Complexity vs. customization. ATKearney.

Aries, O. and Hu, M. (2009). Untangling product complexity in M&A. ATKearney.

Baldwin, C. and Clark, K. (1997). Managing in an age of modularity. *Harvard Business Review*, 75(5):84–93.

Bernstein, F., Kök, A., and Xie, L. (2011). The role of component commonality in product assortment decisions. *Manufacturing & Service Operations Management*, 13(2):261–270.

Blecker, T. and Abdelkafi, N. (2006a). Complexity and variety in mass customization systems: analysis and recommendations. *Management Decision*, 44(7):908–929.

Blecker, T. and Abdelkafi, N. (2006b). Complexity in variety-rich-production systems. *Complexity Management in Supply Chains. Springer, Berlin Heidelberg New York*, pages 67–83.

Blecker, T. and Abdelkafi, N. (2007). The development of a component commonality metric for mass customization. *IEEE Transactions on Engineering Management*, 54(1):70–85.

Blecker, T., Kaluza, B., Abdelkafi, N., and Kreutler, G. (2004). Mass customization vs. complexity: A gordian knot? In *An Enterprise Odyssey. International Conference Proceedings*, page 890. University of Zagreb, Faculty of Economics and Business.

Blecker, T. and Kersten, W., editors (2006). *Complexity management in supply chains: concepts, tools and methods*. Erich Schmidt Verlag.

Bode, C. and Wagner, S. (2015). Structural drivers of upstream supply chain complexity and the frequency of supply chain disruptions. *Journal of Operations Management*, 36:215–228.

Bozarth, C., Warsing, D., Flynn, B., and Flynn, E. (2009). The impact of supply chain complexity on manufacturing plant performance. *Journal of Operations Management*, 27(1):78–93.

Brunoe, T. and Nielsen, K. (2016). Complexity management in mass customization smes. *Procedia CIRP*, 51:38–43.

Calinescu, A., Efstathiou, J., Schirn, J., and Bermejo, J. (1998). Applying and assessing two methods for measuring complexity in manufacturing. *Journal of the Operational Research Society*, pages 723–733.

Choi, T., Dooley, K., and Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: control versus emergence. *Journal of operations management*, 19(3):351–366.

Choi, T. Y. and Krause, D. R. (2006). The supply base and its complexity: Implications for transaction costs, risks, responsiveness, and innovation. *Journal of Operations Management*, 24(5):637–652.

Cook, M. (2001). The complexity of managing complexity. Bain & Company.

Da Silveira, G., Borenstein, D., and Fogliatto, F. (2001). Mass customization: Literature review and research directions. *International journal of production economics*, 72(1):1–13.

de Leeuw, S., Grotenhuis, R., and van Goor, A. (2013). Assessing complexity of supply chains: evidence from wholesalers. *International Journal of Operations & Production Management*, 33(8):960–980.

Deshmukh, A., Talavage, J., and Barash, M. (1998). Complexity in manufacturing systems, part 1: Analysis of static complexity. *IIE transactions*, 30(7):645–655.

Fan, G., Li, A., , Moroni, G., Xu, L., and Liu, X. (2017a). Operation-based configuration complexity measurement for manufacturing system. *Procedia CIRP*, 63:645–650.

Fan, H., Li, G., Sun, H., and Cheng, T. (2017b). An information processing perspective on supply chain risk management: Antecedents, mechanism, and consequences. *International Journal of Production Economics*, 185:63–75.

Fisher, M., Gaur, V., and Kleinberger, H. (2017). Curing the addiction to growth. *Harvard Business Review*, 95(1):66–74.

Fisher, M. L. and Ittner, C. D. (1999). The impact of product variety on automobile assembly operations: Empirical evidence and simulation analysis. *Management science*, 45(6):771–786.

Frizelle, G. and Suhov, Y. (2008). The measurement of complexity in production and other commercial systems. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 464, pages 2649–2668. The Royal Society.

Frizelle, G. and Woodcock, E. (1995). Measuring complexity as an aid to developing operational strategy. *International Journal of Operations and Production Management*, 15(5):26–39.

Garbie, I. and Shikdar, A. (2011). Analysis and estimation of complexity level in industrial firms. *International Journal of Industrial and Systems Engineering*, 8(2):175–197.

George, M. and Wilson, S. (2004). *Conquering complexity in your business*. McGraw Hill.

George Group (2006). Unraveling complexity in products and services. Wharton, University of Pennsylvania.

Golfmann, J. and Lammers, T. (2015). Modular product design: reducing complexity, increasing efficacy. *Performance*, 7(1):56–63.

Gong, Q., Yang, Y., and Wang, S. (2014). Information and decision-making delays in mrp, kanban, and conwip. *International Journal of Production Economics*, 156:208–213.

Hendricks, K. and Singhal, V. (2016). An empirical analysis of the effects of excess inventory on operating performance. Mimeo.

Herd, T. and McClelland, M. (2017). Sizing up M&A value now. Accenture Strategy.

Heylighen, F. (1999). The growth of structural and functional complexity during evolution. *The evolution of complexity*, pages 17–44.

Heywood, S., Spungin, J., and Turnbull, D. (2007). Cracking the complexity code. *McKinsey Quarterly*, 2:84.

Hirose, R., R. Sodhi, D., and Thiel, A. (April, 2017). How do winning consumer-goods companies capture growth?

Isik, F. (2010). An entropy-based approach for measuring complexity in supply chains. *International Journal of Production Research*, 48(12):3681–3696.

Isik, F. (2011). Complexity in supply chains: A new approach to quantitative measurement of the supply-chain-complexity. In L., P., editor, *Supply chain management*. InTech.

Jacobs, M. (2008). *Product Complexity: theoretical relationships to demand and supply chain costs*. PhD thesis, Michigan State University.

Jacobs, M. A. (2013). Complexity: Toward an empirical measure. *Technovation*, 33(4-5):111–118.

Jacobs, M. A. and Swink, M. (2011). Product portfolio architectural complexity and operational performance: Incorporating the roles of learning and fixed assets. *Journal of Operations Management*, 29(7-8):677–691.

Jacquemin, A. P. and Berry, C. H. (1979). Entropy measure of diversification and corporate growth. *The journal of industrial economics*, pages 359–369.

Kaivola, A. (2017). Decreasing product variety-driven costs: The view point from complexity management and component commonality in batch-process context. Master's thesis, Tampere University of Technology, Finland.

Kaluza, B., Bliem, H., and Winkler, H. (2006). Strategies and metrics for complexity management in supply chains. In Kersten, W. and Blecker, T., editors, *Complexity Management in Supply Chains*, pages 3–19. Erich Schmidt Verlag, Göttingen.

Karp, A. and Ronen, B. (1992). Improving shop floor control: an entropy model approach. *International Journal of Production Research*, 30(4):923–938.

Lindemann, U., Maurer, M., and Braun, T. (2008). *Structural complexity management: an approach for the field of product design.* Springer Science & Business Media.

Mariotti, J. (2008). *The complxity crisis. Why too many products, markets, and customers are crippling your company.* Platinum Press.

Marti, M. (2007). *Complexity management: optimizing product architecture of industrial products.* Springer Science & Business Media.

Menezes, M. B. C. and Ruiz-Hernández, D. (2019). Empirical analysis of an information-content based measure for structural complexity. Mimeo.

Novak, S. and Eppinger, S. D. (2001). Sourcing by design: Product complexity and the supply chain. *Management science*, 47(1):189–204.

Palepu, K. (1985). Diversification strategy, profit performance and the entropy measure. *Strategic management journal*, 6(3):239–255.

Perona, M. and Miragliotta, G. (2004). Complexity management and supply chain performance assessment. a field study and a conceptual framework. *International Journal of Production Economics*, 90(1):103–15.

Rutenberg, D. P. and Shaftel, T. L. (1971). Product design: Subassemblies for multiple markets. *Management Science*, 18(4-part-i):B–220.

Saeed, B. and Young, D. (1998). Managing the hidden costs of complexity. Boston Consulting Group.

Schaffer, J. and Schleich, H. (2008). Complexity cost management. In Parry, G. and Graves, A., editors, *Build to Order: The Road to the*, pages 155–176. Springer.

Scheiter, S., Scheel, O., and Klink, G. (2007). How much does complexity really cost? ATKearney.

Serdarasan, S. (2013). A review of supply chain complexity drivers. *Computers & Industrial Engineering*, 66(3):533–540.

Shannon, C. (1948). The mathematical theory of communication. *Bell System Technical Journal*, 27:379–423.

Shannon, C. and Weaver, W. (1949). The mathematical theory of communication. *University of Illinois Press Chicago, IL.*

Simon, H. A. (1962). The architecture of complexity. In *Proceedings of the American Philosophical Society*, volume 106, pages 457–476. American Philosophical Society.

Sivadasan, S., Efstathiou, J., Calinescu, A., and Huatuco, L. (2006). Advances on measuring the operational complexity of supplier–customer systems. *European Journal of Operational Research*, 171(1):208–226.

Sivadasan, S., Efstathiou, J., Frizelle, G., Shirazi, R., and Calinescu, A. (2002). An information-theoretic methodology for measuring the operational complexity of supplier-customer systems. *International Journal of Operations & Production Management*, 22(1):80–102.

Sivadasan, S., Efstathiou, J., Shirazi, R., Alves, J., Frizelle, G., and Calinescu, A. (1999). Information complexity as a determining factor in the evolution of supply chains. In *Proceedings of the International Workshop on Emergent Synthesis*, pages 237–242.

Sivadasan, S., Smart, J., Huatuco, L., and Calinescu, A. (2010). Operational complexity and supplier–customer integration: case study insights and complexity rebound. *Journal of the Operational Research Society*, 61(12):1709–1718.

Subramanian, N. and Rahman, S. (2014). Supply chain complexity and strategy. In Ramanathan, U. and Ramanathan, R., editors, *Supply chain strategies, issues and models*, pages 1–27. Springer.

Suh, N. P. (1999). A theory of complexity, periodicity and the design axioms. *Research in Engineering Design*, 11(2):116–132.

Tolman, R. C. (1938). *The Principles of Statistical Mechanics*. The Clarendon Press.

Vachon, S. and Klassen, R. (2002). An exploratory investigation of the effects of supply chain complexity on delivery performance. *IEEE Transactions on Engineering Management*, 49(3).

Waldrop, M. (1993). *Complexity: The emerging science at the edge of order and chaos*. Simon and Schuster.

Wang, H., Gong, Q., and Wang, S. (2017). Information processing structures and decision making delays in mrp and jit. *International Journal of Production Economics*, 188:41–49.

Wang, Q. and von Tunzelmann, N. (2000). Complexity and the functions of the firm: breadth and depth. *Research Policy*, 29(7):805–818.

Wazed, M., Ahmed, S., and Nukman, Y. (2009). Commonality and its measurement in manufacturing resources planning.

Weber, C. (2005). What is "complexity"? In *DS 35: Proceedings ICED 05, the 15th International Conference on Engineering Design, Melbourne, Australia, 15.-18.08. 2005*.

Weiser, A., Baasner, B., Hosch, M., Schlueter, M., and Ovtcharova, J. (2016). Complexity assessment of modular product families. *Procedia CIRP*, 50:595–600.

Wilding, R. (1998). The supply chain complexity triangle: Uncertainty generation in the supply chain. *International Journal of Physical Distribution & Logistics Management*, 28(8):599–616.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus.* Routledge Classics, London. (Translated by Pears D.F. and McGuiness, B.F., 1975).

Wu, Y., Frizelle, G., and Efstathiou, J. (2007). A study on the cost of operational complexity in customer–supplier systems. *International Journal of Production Economics*, 106(1):217–229.

Appendix

One of the main issues that arise when a new measure or technique is proposed is its comparison to other available measures targeting the same problem. In many cases such comparison is possible (e.g. scales measuring psychological or personality traits, facial recognition techniques, algorithms for the solution of certain combinatorial problem, and so on) provided that either the sample sets for previous experiments are available or, alternatively, that the benchmark techniques can be applied over the same population where the new technique is tested. In any case, a fundamental assumption is that both, old and new techniques are aimed at measuring the same phenomenon and that the parameters and/or explanatory variables are the same in both approaches.

Notwithstanding how desirable such comparison can be, we have found that it is virtually impossible to compare our measure of complexity against other measures available in the literature. Moreover, it seems to us that this is a problem affecting all proposed measures, as we have not been able to find a reference where the authors compare their results against the ones of other authors or even against their own previous results. We think this is inconvenience is inherent to the nature of the problem under study, stemming from the absence of a unique and generally accepted definition of complexity and that different studies target diverse aspects or conceptualisations of complexity.

Table 4 summarises (some of) the information requirements of a number of articles addressing the problem of measuring complexity. The first column indicates the scope of the study or, equivalently, the type of complexity they are concerned about, the second one indicates whether the measure is entropy based, algebraic or based in statistical or regression analysis. The third column (without being exhaustive) summarises the main variables discussed in each study, and the fourth one indicates whether or not the raw data are provided in the study. The fifth column indicates if the proposed measure (and information required) allows for look-ahead (or what-if) analysis. The last column provides the reference.

As it can be observed in the table, most of the available work is based on time series data collected on a large number of variables, whereas our measure is based on the distribution of one single variable: sales. Moreover, while most studies focus on flows, *pars*-complexity is a static measure. Additionally, it must be taken into consideration that, unlike other studies referred in Table 4, our measure is aimed at measuring structural complexity -leaving aside issues of complicatedness or uncertainty. We hope that this argument suffices for justifying the absence of inter-measures comparisons in this manuscript.

Target Complexity	Type of Measure	Variables	Data Provided	Look-Ahead	Reference
Product (diversification)	Entropy	Segment's share over total sales of the firm	Partially	No	Palepu (1985)
Manufacturing	Entropy based	Bills of materials, routings, work centres, demand pattern. Periodical data.	No	No	Frizelle and Woodcock (1995)
Supply chain	Entropy based	Imperfectly specified. Forecasted, requested, scheduled and confirmed deliveries. Target and actual production. Materials and information flow. Periodical data.	No	No	Sivadasan et al. (1999)
Supplier-customer systems	Entropy based	Imperfectly specified. Flow focused. Demand, production, deliveries. Variations in time and quantities. Periodical data.	No	No	Sivadasan et al. (2002)
Supply chain	Regression analysis	Lead time, throughput time, late delivery, tardiness, AMT investment, vertical integration, quality failures, firm size, etc.	Partially	No	Vachon and Klassen (2002)
Supply chain, processes	Algebraic	Total value-add time in the process, percentage of defective products, unit processing time, total demand.	Examples	No	George and Wilson (2004)
Supplier-customer systems	Entropy based	Flow variations (order-forecast, delivery-order, actual-scheduled production), time and/or quantity variations.	Partially	Yes	Sivadasan et al. (2006)
Supply chain	Algebraic	Number of SKUs, markets served, company legal entities, facilities, employees, suppliers, customers. Sales revenues.	Examples	Yes	Mariotti (2008)
Supply chain	Regression analysis	Number of customers, life cycle, number of active material parts, number of products, number of suppliers, percentage of purchases imported, etc.	Partially	No	Bozarth et al. (2009)
Supply chain	Entropy based	Expected and actual orders per month (flow based).	Example	Yes	Isik (2010, 2011)
Supply chain	Statistical analysis	Expert survey. Number of stock keeping units, stock locations, employees and years active in market. Several questionnaire topics.	No	No	de Leeuw et al. (2013)
Product / Portfolio / Supply Chain	Algebraic	Number of variants, common elements of components, number of connections.	Yes	Yes	Jacobs (2013)

Table 4: Data requirements of different complexity measures proposed in literature