Received: 2 June 2017 | Revised: 8 December 2018 | Accepted: 11 December 2018
DOI: 10.1002/sres.2574

RESEARCH PAPER

WILEY SYSTEMS and BEHAVIORAL RESEARCH SCIENCE

Antifragility as a design criterion for modelling dynamic systems

Harald de Bruijn[1] | Andreas Größler[2] | Nuno Videira[3]

[1] Department of Business Administration, Nijmegen School of Management, Nijmegen, The Netherlands

[2] Department of Operations Management, Faculty of Management and Social Sciences, University of Stuttgart, Stuttgart, Germany

[3] CENSE – Center for Environmental and Sustainability Research, Departamento de Ciências e Engenharia do Ambiente, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, Caparica, Portugal

Correspondence
Harald de Bruijn, Department of Business Administration, Nijmegen School of Management, Thomas van Aquinostraat 3, Nijmegen 6500 HK, The Netherlands.
Email: harald-de-bruyn@live.nl

Funding information
EACEA (Education, Audiovisual and Culture Executive Agency of the European Commission); Portuguese Foundation for Science and Technology, Grant/Award Number: UID/AMB/04085/2013

Abstract
Highly improbable events can have a substantial impact on complex socio-economic systems and are frequently difficult to predict beforehand but easy to explain afterwards. Antifragile systems can withstand and benefit from this kind of outlier events, whereas merely robust systems cannot in any case. Yet the aim to design robust systems is almost as old as the system dynamics field itself. This research therefore aims to investigate the extent to which an antifragile system design criterion is more valuable than a robust one. By means of an extensive literature review, a simulation model was constructed, which is demonstrated to be antifragile. Comparing the antifragile and robust versions of the model shows that the former—as theorized—yields more favourable results in an environment with impactful outlier events. Implementing antifragility in systems involves the difficult task of changing policies (and, eventually, the mental models) of decision-makers. Consequently, this research concludes that antifragility should not and cannot always be attained; its feasibility is to be assessed at the start of a system dynamics modelling project.

KEYWORDS
antifragility, black swans, system dynamics

1 | INTRODUCTION

The meta-assumption behind the system dynamics methodology stresses the importance of explaining the behaviour of a system by means of its endogenous structure (Forrester, 1968; Richardson, 2011). Within the field, the ultimate aim of forming structural explanations is to attain robust system behaviour in which the system performs as well as the circumstances allow (Coyle, 1996; Sharp, 1976). It can be argued that the foundations of this robustness design criterion are grounded in the general preference of human beings for certainty and the corresponding aim for economic stability (Kahneman & Tversky, 1979; Quinn & Woolley, 2001). Typically, the robustness of a system is evaluated by building a formal model of its structure and conducting behaviour sensitivity tests such as simulating a worst- and a best-case scenario and/or conducting a Monte Carlo simulation (Sterman, 2000, p. 885). On the basis of such simulation experiments, potential changes to the system can be tested, and structural or policy amendments can be recommended to improve robustness. If the system's behaviour shows a satisfactory response to seemingly extreme conditions and is insensitive to changes in its parameter values within a plausible range, the system can be categorized as robust (Moxnes, 2005; Sterman, 2000, p. 883).[1]

[1] On the basis of this interpretation, robustness as used in this article refers to robustness as a design criterion for simulation models (and, subsequently, real systems) and not as part of model validation and/or extreme condition tests such as described in Barlas (1996) and Sterman (2000, p. 866).

According to Taleb (2007), such robustness assessment approaches lead to misplaced trust in the robustness of a system. Taleb argues that decision-makers apply a "Gaussian" way of looking at the world and the models they build of the world; they start to focus on the ordinary before investigating the exceptions. Applying this lens works well in a situation where there is "a rational reason for the largest [observation] not to be too far away from the average" (p. 256) and, consequently, allows for the application of the central limit theorem and law of large numbers.[2] Hence, decision-makers tend to assume a Gaussian probability distribution of events and base their decisions on this distribution.

However, the human designed world is complex and characterized by interdependencies "subjected to reinforcing positive feedback loops which cause fat tails[3]" (Taleb, 2007, p. 382). Taleb (2015) shows that in these complex environments, decision-makers are prone to estimating the wrong distribution because the central limit theorem and law of large numbers work very slowly or not at all.[4] This holds especially for decisions based on human-created data such as risk metrics (Paté-Cornell, 2012; Taleb, 2009). Consequently, decision-makers applying statistical distributions belonging to the thin-tailed domain, least square methods, or any method using variance as a measure of dispersion, will end up with correct estimations of the asymptote—but incorrect estimations of the tails of the distribution (Allington, McCombie, & Pike, 2012; Chichilnisky, 2010; Paté-Cornell, 2012; Taleb & Pilpel, 2004; Taleb and Douady (2013). Taleb and Pilpel (2004) show that it is also not possible to derive or fit the tails due to an unavoidable sampling error. The problem is especially large for negatively skewed probability distributions of an event as these always overestimate the mean and underestimate variance and risk.

Using incorrect estimates of the tails implies that in the advocated robustness assessment, abrupt and high impact outlier events might be missed. Following this reasoning, Taleb (2007) emphasizes that systems that are categorized as being robust can still "break" when faced with what he calls black swan events. Black swan events are extremely rare, are easily explainable in retrospect but not predictable beforehand, and have an extreme impact in complex environments subject to "positive reinforcing feedback loops" (p. 382). Taleb (2007, p. 334; 2012, p. 274) and Allen (1988, p. 129) even argue that policies aimed to improve the robustness of a system are insufficient and dangerous as they often unintentionally "fragilize" the system and create a misplaced confidence in its robustness. Furthermore, effects of black swans are becoming severer due to an increasingly globalized and connected world where reinforcing loops cause interdependencies across spatial and temporal borders, and redundancies in systems are removed for efficiency reasons (Allen, Varga, & Strathern, 2010; Estrada, 2009; Taleb, 2007). Examples of the impacts of these black swans are ample and include the fall of the bank IndyMac and the failure of the hedge fund Long Term Capital Management as a consequence of unexpected economic volatility (Allington et al., 2012; Taleb, 2012). Note that in these examples, the black swan was the unexpected economic volatility and not the failure of the institutions.

To reduce and even utilize the impact of black swans, Taleb (2012) introduces the concept of antifragility as an alternative design criterion to robustness. Antifragility refers to systems that gain from volatility and disorder and show an improvement in behaviour when subjected to large and implausible changes in parameters. The concept of antifragility is thus related to an evolutionary understanding of the "resilience" concept (see Dalziell & McManus, 2004; Walker, Holling, Carpenter, & Kinzig, 2004). Moreover, an antifragile system shows ex ante and ex post adjustment to black swans, whereas a robust system focuses only on ex post adaptation, which often comes too late.

The purpose of this article is to combine the notion of antifragility and system dynamics modelling and to explore the extent to which an antifragile model design criterion is more valuable than a robust one. First, it offers a conceptual representation of antifragility and its origins by a simulation model. Such a conceptualization is not possible with methods lacking the ability to deal with black swan events, feedback loops, non-linearities, and endogenous explanations as these are crucial components (Taleb, 2007, pp. 367, 383; Taleb, 2012, p. 287). This is one of the reasons why most knowledge about black swans and antifragility is anecdotal only. Second, it allows to compare the antifragile model to a robust one and to assess the proposition that the antifragile design criterion is preferred to a robust one. The comparison confirmed the proposition by showing that, in an environment with black swans, the antifragile yields more favourable simulation results than the robust model. Hence, the results show that the use of the robustness

[2]The central limit theorem shows that the behaviour of the sum of random variables makes it possible to derive the asymptotic probability distribution and the properties of this distribution (Feller, 1971). The dispersion of the distribution is lowered by the law of large numbers.

[3]Fat tails emphasize "the outsized role in the total statistical properties played by one single observation—such as one massive loss coming after years of stable profits or one massive variation unseen in past data" (Taleb, 2009, p. 745).

[4]For mathematical evidence, the reader is referred to Taleb (2015, chapter 7).

criterion (as usually applied in system dynamics) is not axiomatic. However, also the use of an antifragility design criterion is not self-evident as it requires a policy process, which is focused on different leverage points, including the mental models of the decision-makers. These findings are of special interest to policymakers and system dynamic modellers involved in scenario planning. They provide an alternative way to be prepared for the unexpected and pinpoint the importance of accounting for the mental models in organizations using scenario planning.

The structure of this paper is as follows. In Section 2, we briefly summarize the notion of black swans and antifragility and relate it to system dynamics. In Section 3, a generic system dynamics model structure is inferred from the verbal descriptions of robust and antifragile systems. In Section 4, we test the reaction of these two model versions on random input, including black swan events. Section 5 compares robustness versus antifragility as a design criterion, based on the findings from the model analysis. Section 6 concludes this article with a discussion of its practical relevance and of scientific implications.

2 | BLACK SWANS AND GENERIC SYSTEM TYPES

2.1 | Modelling black swans in System Dynamics

As black swans cannot be predicted, being prepared on a general level is the only means we have to protect ourselves against these events (Makridakis, Hogarth, & Gaba, 2009; Makridakis & Taleb, 2009). The first step in preparing for black swans is formed by widening mental models so that they include the potential occurrence of extreme events (Aven, 2013), even when it is unknown what such extreme events could be. Moreover, as reactions of systems to black swans are complex, there also exists a strong incentive to include computer simulations to facilitate learning about these reactions (Moxnes, 2005). The system dynamics approach seems extremely useful in this respect as it aims to change mental models by means of modelling and simulation and provides policy recommendations based on an analysis of the system (Forrester, 1968; Richardson, 2011). With these aims in mind, mental models can be improved, the understanding of reinforcing feedback loops and non-linearities as described by Taleb (2007, pp. 104, 382; 2012, book V) can be created, and leverage points for making the system antifragile can be discovered.

Although no system dynamics study explicitly refers to the term “antifragility,” it might have been Coyle (1977 as cited in Coyle 1996) who made one of the first references to the mechanism behind antifragility. Coyle noted that “a managed system should be able to defend, recover from, and create and *exploit* shocks” (p. 6, italics added). The inclusion of benefitting from shocks is crucial to point out here, as it implies that the system is not merely defended—and therefore robust—against shocks but is also antifragile. Coyle (1996, p. 10) later switches to an alternative definition of managing systems in which the benefitting from shocks appears to be absent.

Two issues deserve further attention in this regard: (a) the investigation of how black swans should be represented in a model and (b) what an antifragile system dynamics model structure could look like. Concerning the first point, there has been a call to give more attention to simulating system dynamics models under uncertainty as current efforts were deemed insufficient (Pruyt, 2007). Kwakkel and Pruyt (2013) were the first who specifically differentiated between different kinds of uncertainty in a system dynamics context. In line with the previous paragraph, those authors acknowledged that modelling can be used by decision-makers to be at least less surprised by a black swan, hereby being prepared and reducing the impact. Yet, because of their wide definition of uncertainty, Kwakkel and Pruyt (2013) concluded that modelling under deep uncertainty is extremely difficult. Our research takes a different standpoint as—based on the original definition by Taleb and Pilpel (2004, p. 7)—the main uncertainty comes from identifying the black swan event and the probabilities attached to it and not from specifications inside the model boundary. This means that the system's structure is fixed. The uncertainty stems from the likelihood and type of event impacting on the system and not from boundaries, structures, the way decisions are modelled, non-linear relationships, and/or table functions. The latter approach would call for, as indicated by Kwakkel and Pruyt (2013), a set of plausible models.

Regarding the question of what an antifragile system dynamics model can look like, we follow the characteristics laid down by Taleb: (a) the model's exogenous black swan structure needs to be able to represent a fat-tailed distribution as this replicates an environment similar to the ones experienced by many decision-makers (Taleb, 2009, pp. 746–748; 2015, chapter 4); (b) outlier events need to have an extreme impact as “in fat tailed domains harm comes from the largest single event” (Taleb & Douady, 2013, p. 5); (c) outliers need to trigger a reinforcing loop as “mechanisms [in complex systems] are subjected to positive, reinforcing feedback loops” (Taleb, 2007, p. 382); and (d) values around the mean of the distribution need to have a relatively small impact, as these are captured by traditional methods hereby having modest consequences only (Taleb, 2009, p. 755).

In terms of fulfilling the requirements for black swan representations, a Monte Carlo element is used as it generates random numbers, which make it impossible to have perfect insight in future events that will impact on the model, thereby addressing the inabilities of risk estimation methods to predict black swans (Dhawan, 2006; Taleb & Pilpel, 2004). The Monte Carlo element is based on a standard Cauchy distribution[5] as this distribution fulfils the outlined requirements by Taleb and Douady (2013, p. 5). The element hereby creates a situation of ontological uncertainty. That is, a situation based on a non-ergodic assumption implying that no amount of past data can determine a probability distribution that is a function of changes that occur in the future (Allington et al., 2012). Consequently, an increase in the sample of simulation runs does not lead to a more correct estimation of the mean (Taleb & Douady, 2013).

2.2 | Generic system types: Robust, resilient, and antifragile

Taleb (2007) outlines a sequence of generic system types: fragile, robust, and antifragile. However, the scope of this paper will be on a variation of Taleb's sequence, namely, on robust, resilient, and antifragile system types. The adaptation is made based on the following reasons: (a) Resilience is included in our analysis—even though Taleb does not classify it as a system type—because it is studied frequently and applied to a large number of systems (Carpenter, Walker, Anderies, & Abel, 2001; Rydzak & Chlebus, 2007). Resilience addresses the dynamics of how a system absorbs the impacts of stress or shocks and how it reorganizes afterwards (Carpenter et al., 2001). Omitting resilience would imply that the paper neglects a significant amount of literature (e.g., Bueno, 2012) focusing on system reactions to randomness. (b) Fragility is excluded from the generic system types because human beings (including modellers) search for certainty and robustness (Kahneman & Tversky, 1979) to improve systems. This means that fragile systems are normatively not desired as they break from randomness and, hence, do not provide the desired robustness.[6] Conclusive reactions of each of the identified generic system types to randomness are shown in Figure 1. On the basis of the system reactions, possible behaviour pathways are identified in Figure 2.

[5]More information concerning the use of Cauchy distribution and specification specifics of the used can be found in part F of the Supporting Information.

[6]For a system reaction function of the different fragility types, the reader is referred to Taleb and Douady (2013, p. 1679).

Robustness was already recognized by Sharp (1976) and Senge and Forrester (1980) as an important design criterion for system dynamics models. A system is classified as robust if it (a) shows satisfactory responses when subjected to a wide variety of inputs; (b) performs satisfactory over the range of parameter values considered plausible; and (c) is relatively unaffected by a considerable amount of noise usually found in socio-economic systems. Important to note here is that responses are to be evaluated in terms of behaviour patterns and not in numerical values (Hekimoğlu & Barlas, 2010). Hence, a robust system shows no significant changes in behaviour patterns when it is subjected to a variety of shocks. Consequently, it can be observed from Figure 1 that moving an equal degree to the left and to the right from the "system's initial level" point always leads to a net effect of zero, as the reaction function in the normally distributed space is symmetrical. Hence, in a robust system situated in the normally distributed space, the reaction function ensures that the gains and losses cancel out over time. Taking this observation, it is possible to show that the robust system's level in Figure 2 starts and ends at the same level. Reality, however, shows that systems that initially appear to be robust are at some point in time subjected to outlier events such as black swans (Taleb, 2012). In line with Taleb's argumentation, Figure 1 therefore displays that events occurring in the fat tail area impacting on the "initial system's level" lead to a negative concave reaction of a robust system. This negative concavity results from the negative asymmetry in the reaction function at that point. From this, it follows that a system might appear to be robust in the short run—because it is only exposed to normally distributed randomness—but it is fragile in the long run, as time often equals not-normal distributed randomness (Taleb, 2007).

Carpenter et al. (2001) outline that resilience addresses the dynamics of dealing with disturbances, how a system absorbs the impacts of stress or shocks, and how it reorganizes afterwards. Up to a certain threshold, a resilient system deals with disturbances in the same system configuration. However, if the threshold is breached, the system's recovery process based on self-organization is started. This capability reflects the ability of a resilient system to return to its original state after randomness. A resilient system can do this because of its defence mechanism. This mechanism learns from different circumstances and adapts its capacity towards future stressors impacting on the system (Tseitlin, 2014). This means that a resilient system builds up redundancy, which allows the system to survive future similar stressors. Comparing it to a robust system, the resilient is similar in the sense that it is capable of handling mild randomness (Bueno, 2012), but, additionally to the

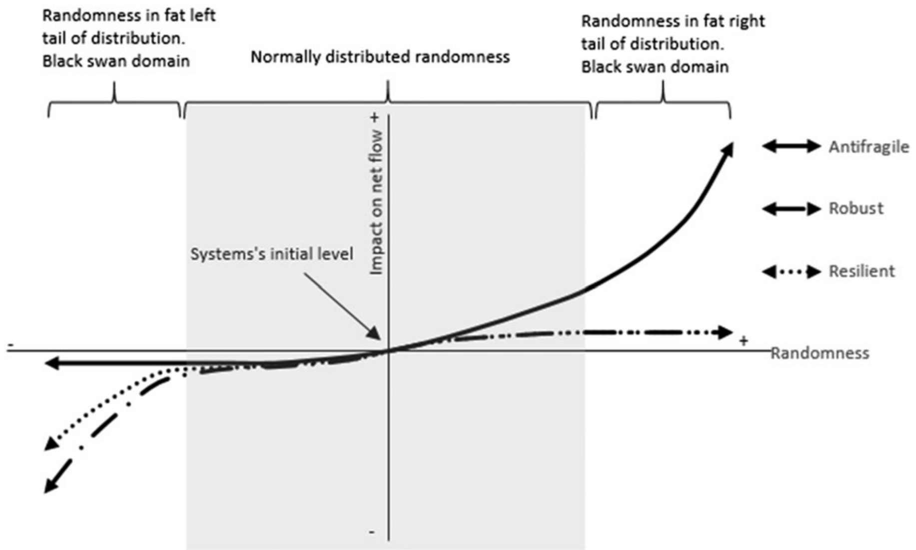

FIGURE 1 System reaction functions to randomness imposed from the "systems initial level" point. Moving from the left to the right from the systems initial level point leads to a gain or a loss for that initial level. Because gains and losses for levels occur via flows, the *y*-axis can be interpreted as the magnitude of the impact in a net flow for the system level. If we take a business as example, a black swan event in the left tail of the probability distribution will, for a robust system, lead to a large negative change in the net flow of profits. Note that the figure does not display values over time

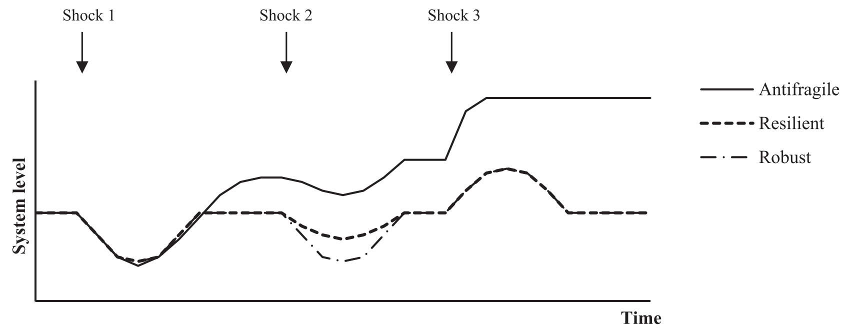

FIGURE 2 Possible behaviour paths over time of the system's levels as a result of three equally strong shocks in the normal distribution space. Shocks 1 and 2 are negative, and Shock 3 is positive

robust, the resilient learns from previous shocks and is therefore harmed less by a similar shock in the future. Figure 1 shows this by displaying a reaction function for the resilient, which decreases less severely than the function of the robust and, consequently, is asymmetrical. This system reaction leads to a resilient behaviour pattern, displayed in Figure 2, in which the resilient is harmed less than the robust at the second shock. Though, like the robust, large deviations such as exposure to black swan events can lead to a concave system's reaction as shown in Figure 1.

On the other hand, an antifragile system can be identified by a reaction to randomness as displayed in Figure 1. It can be observed that an antifragile system is characterized by a positive convex-asymmetric reaction function. Therefore, imposing randomness impacting both positive and negative on the initial system's level point leads to more gains than losses. This means that for the antifragile system, the gains are always bigger than the losses notwithstanding the size of the randomness impacting on the system (Taleb, 2012). A system that benefits from randomness is then also characterized by long-term survival. Table 1 provides an overview of the main properties of each generic system type.

As observed in Table 1, robust systems are characterized under normal randomness by a symmetric reaction function in which, over time, the effect of positive and negative random events impacting on the system will cancel out. This means that a robust system will show an outcome probability distribution, which is centred around the mean and characterized by thin left and right tails. Resilient systems are characterized by an asymmetric reaction function, as the losses under normal randomness are smaller for the resilient than for the robust. Consequently, the probability distribution for the outcome of a resilient system must have a higher mean than the robust, a shorter left tail, and a longer right tail. Lastly, the antifragile system is characterized by a positive

TABLE 1 Aggregated system properties

System reaction	Reaction function to normally distributed randomness	Effect of normally distributed randomness	Reaction function shape to not normally distributed randomness	Sensitivity to left tail uncertainty of environment	Sensitivity to right tail uncertainty of environment	Left tail outcome distribution	Right tail outcome distribution
Robust	Symmetric	Neutral	Negative concave	Breaks at a higher point than fragile systems	No effect	Thin	Thin
Resilient	Asymmetric	Neutral and prepares for similar volatility	Negative concave	Breaks at a higher point than robust systems	No effect	Thin (thinner than robust)	Thin
Antifragile	Asymmetric	Benefits	Positive convex	Robust	Benefits	Thin	Fat

Note. Left and right tails refer to the thickness of the tails of the probability distributions belonging to the level of each system resulting from the system reactions as displayed in Figure 1.

asymmetric reaction function leading to a convex response to randomness. Hence, antifragile systems are not sensitive to volatility in the left tail and positively exposed to right tail uncertainty. This means that the outcome probability distribution of an antifragile system is skewed to the right (Taleb, 2015).

3 | INFERRING A SYSTEM DYNAMICS MODEL OF BLACK SWANS AND ANTIFRAGILITY

A robust system shows—due to the strength of its balancing loops—no significant changes in its behaviour pattern when subjected to randomness (Coyle, 1996; Hekimoğlu & Barlas, 2010; Moxnes, 2007; Sharp, 1976). However, as explained earlier, humans are prone to wrongfully estimating the tails of the distribution, which form an important input to estimating the robustness of a system. Focusing on designing robust models can then also seen to be misleading, as the balancing loops might be strong enough in the tests, whereas the real-world counterpart will break due to the occurrence of a black swan (Taleb, 2012). An example of this is a failing balancing loop in which the cash reserve of a bank is unsuccessfully used during a bank run.

The system dynamics methodology has only accounted for black swans to a limited extent. First, system dynamics is in its robustness assessment merely focused on plausible —and therefore limited—parameter variations (Sterman, 2000, p. 884). As mentioned earlier, this approach is defendable for parameters that follow a thin-tailed distribution and parameters for which the minimum and maximum values are known. Examples of these are variables that are bound by nature, such as a person's height. The approach however fails for parameters based on human-created data (Paté-Cornell, 2012; Taleb, 2009, 2015). The system dynamics methodology recognizes that statically derived parameter estimations can be flawed and are characterized by overconfidence (Sterman, 2000, p. 884). Yet, as a solution, a rule of thumb is provided in which it is suggested to use at least a parameter variation range, which is twice as wide as statistical tests and judgement suggest. A more robust approach was suggested by Miller (1998) by introducing automated nonlinear tests to "search across a set of *reasonable* model perturbations with the objective of maximizing the deviation between the original model's prediction and that obtained from the perturbed model" (p. 820, italics added). However, also this approach relies on plausible parameter variations to discover worst and best cases. Taleb, Canetti, Kinda, Loukoianova, and Schmieder (2012) show that it is exactly this limited amount of variance assumed that creates a mere illusion of a robust system. Sniedovich (2012) supports this by showing that simulating a worst-case scenario—which is also a frequent activity in system dynamics modelling—leads to too conservative policies and hence potentially costly errors. By focusing on sensitivity analysis with plausible parameter variations and worst-case scenarios, improbable events are missed, and system dynamics falls subject to the so-called inverse and the Lucretius problem. In short, the inverse problem holds that the derived probability distributions are always erroneous due to an unavoidable sampling error. The asymptote of the parameter distribution can be estimated correctly, whereas estimations for the tail will be incorrect (Allington et al., 2012; Chichilnisky, 2010; Paté-Cornell, 2012; Taleb & Douady, 2013; Taleb & Pilpel, 2004); the Lucretius problem refers to the misplaced confidence in worst-case scenario analysis as it is not possible to know what the worst-case is (Taleb, 2007, 2012).

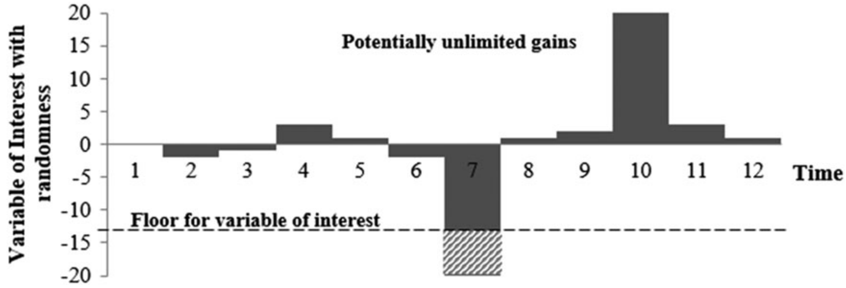

FIGURE 3 A barbell strategy over time. Flooring negative payoffs for the variable of interest while keeping upside effects. Adapted from Taleb, 2015, p. 436)

Second, there is a tendency in the system dynamics field to search for optimal policies in situations of uncertainty (e.g., Kwakkel & Pruyt, 2013; Moxnes, 2005). However, optimization leads to the creation of fragile systems (i.e., a single large shock has a greater effect than the cumulative effect of equivalent small shocks) due to its usual focus on a single objective (Allen, 1988; Taleb, 2012). European railroads are, for example, overly optimized, as a small unexpected event will lead to a disproportionate amount of chaos (Taleb, 2012, p. 274). Additionally, Taleb (2012, p. 45) points out that optimizing inventory levels is inefficient as it makes companies vulnerable to hazardous unusual events but also prevents them from exploiting the opportunities that arise during these events. Using antifragility as a design criterion addresses both aspects but requires changes in the way models are designed.

Taleb (2012) outlines five characteristics that are conditional on the reaction of an antifragile system: the "barbell strategy," hormesis, redundancy, unlimited gains, and small-scale experimentation. These characteristics form starting points for the dynamic hypothesis of an antifragile system dynamics model and are translated subsequently into six propositions to guide model building. Throughout the modelling and analysis process, an increase in the state of the variable of interest was considered as beneficial.

In the processes of inferring a system dynamics model from the concepts of Taleb, the paper aims to remain as much as possible grounded on the original concepts. This has the advantage that translation errors are reduced; the disadvantage is that the translation results in abstract concepts. Therefore, real-world examples of the described structure are included in the Supporting Information, part G. These examples are focused on the banking system, as explored by Taleb, and also include socioecological systems to show the broad applicability of the concepts. Note that the examples do not aim to form any proof of applicability, they merely facilitate the reader's understanding of the abstract concepts.

First, an antifragile system is characterized by operating under a "barbell strategy." This strategy controls the dispersion of a system's outcome by focusing on the avoidance of risk in areas where potentially large negative outcomes can occur while at the same time taking a lot of small risks in areas where positive outcomes can be potentially of a large scale (Derbyshire & Wright, 2014). Figure 3 shows how a barbell strategy limits the harm from negative black swans by implementing a floor that the variable of interest cannot cross but keeps the possibility open for positive black swans. A simple example of this is a bet with limit liability where the outcome cannot go below zero (Taleb, 2011, p. 7; Taleb & Douady, 2013, p. 1685).[7]

To control the dispersions of outcomes of the variable of interest, a balancing loop needs to dominate after a negative black swan. This loop pushes the variables of interest back to its desired state and hereby creates a floor for the possible negative values. Also, it can be deduced by the convex reaction function of antifragility that in case of a positive black swan, the system needs to be dominated by a reinforcing loop, which creates further gains in the variable of interest. This leads to the following two propositions:

> **Proposition 1.** *In case of a negative black swan, antifragility requires a dominating balancing feedback loop, which limits the negative reinforcing spiral caused by the black swan and gets the variable of interest back to its desired state.*

In terms of model operationalization, this implies that a balancing loop needs to influence the variable of interest in the case of a strong decrease. To attain this goal, the most basic balancing loop such as described in Sterman (2000, p. 274) will be used.

> **Proposition 2.** *In case of a positive black swan, antifragility requires a dominating positive reinforcing feedback loop, which enhances the positive impact caused by the black swan.*

This proposition implies the need for a link between the variable of interest and the associated inflow. Such a

[7]For more complex examples such as portfolio construction, the reader is referred to the mentioned references.

structure could function as a reinforcing loop, which, once triggered by randomness, allows for a further increase in the variable of interest by means of an increase in the inflow.

Second, an antifragile system is characterized by hormesis. Hormesis in an antifragile system means that the reaction to a source of randomness includes the overcompensation in a system's state as well as overcompensation in terms of preparedness for the next shock. An antifragile system therefore goes in an overshoot mode, meaning that it is not only fitted for the environment it is in but also prepared for more randomness. As this requires an increase in the system's state, the following proposition can be deduced:

> **Proposition 3.** *The desired level of the variable of interest controlled by the balancing loop identified in Proposition 1 needs to be flexible in such a way that the level can be increased after a shock.*

This proposition can be established by the inclusion of a link between the source of randomness and the desired level attached to the variable of interest. The connection includes a delay as the system initially tries to return to its original state and only then increases its desired level.

Third, antifragility is characterized by redundancy. According to Makridakis and Taleb (2009), redundancy is "a strategy that increases survival in complex systems" (p. 843) such as a reserve build up because of overcompensation, resulting from hormesis. Makridakis and Taleb argue that a redundancy buffer needs to fulfil two functions to serve as overcompensation: (a) It needs to ensure survival of the system in difficult situations and (b) it needs to be proactive in building reserves that are used for negative events that will repeat themselves. It is then also in extreme circumstances that redundancy ensures survival of a system by facilitating the regeneration of the system itself. This assurance is per Taleb (2009) more important than inefficiencies attached to redundancy in the short run. From this, the following two propositions can be deduced:

> **Proposition 4.** *Antifragility requires a stock of redundancy, which stays relatively at the same level compared with its desired state and increases disproportionately to randomness negatively influencing the variable of interest.*

> **Proposition 5.** *If the balancing loop—identified in Proposition 1—aims to influence the inflow, the stock of redundancy needs to increase the variable of interest by means of an inflow.*

Proposition 4 requires that a link exists between the randomness in the model and the disproportional increase in redundancy. This mechanism can be established in a system dynamics model with the inclusion of a non-linear look-up function, which connects the source of randomness to the inflow associated with the redundancy stock. Regarding Proposition 5, the outflow of the redundancy needs to flow into the variable of interest.

Fourth, antifragility is—with exceptions of natural processes—defined as having the potential to generate unlimited gains. However, in line with the limits to growth study (Meadows, Meadows, Randers, & Behrens, 1972), every system is limited in terms of size if it is considered in its wider environment. Therefore, this research will follow the limits to growth assumption and divert from Taleb's original ideas in this case, which leads to the inclusion of the following proposition:

> **Proposition 6.** *A balancing loop needs to establish an upper limit to the value of the variable of interest.*

Limiting growth is commonly modelled by the inclusion of a carrying capacity and the evaluation of a ratio between this carrying capacity and the actual variable of interest (Sterman, 2000, p. 118). As soon as the variable of interest comes close to the carrying capacity, the outflow is increased, which reduces the variable of interest.

Fifth, antifragility is characterized by small-scale experimentations where—by means of trial and error—antifragility comes from the fragility of its components (Taleb, 2012). In this process, the smaller subunits of the system fail, hereby sending information about failures to the overall antifragile system. An example of this would be a start-up ecosystem. In such a system, information of failing or succeeding start-ups is send to other new and established institutions as to what might be fruitful paths to pursue. This mechanism is omitted in this research as it requires the interaction of micro- and macro-level systems and emergences of new structures. Besides the fact that such a process is difficult to capture in system dynamic simulations, it would also make the configuration of the model overly complex. Thus, our study rather underestimates the effect of antifragility since we left out this characteristic.

The antifragile model structure is the result of translating the six propositions into system dynamics lingo. The causal loop diagram of the model, including the causal representation of each proposition (1 to 6), is displayed in Figure 4. For an exact description of the model structure based on the antifragility literature, the reader is referred to part A of the Supporting Information accompanying this article. Please note that the graphical functions and adjustment times used in the model are not empirically grounded due to the model's level of abstractness.

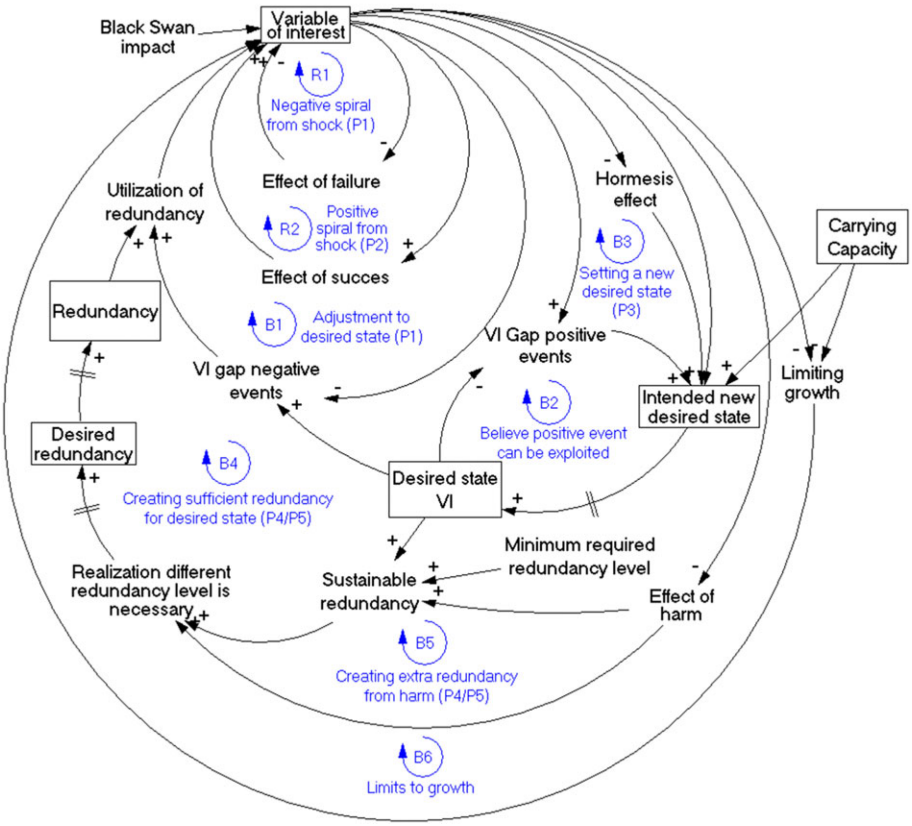

FIGURE 4 Causal loop diagram representing the main loops of the antifragile model structure [Colour figure can be viewed at wileyonlinelibrary.com]

4 | MODEL ANALYSIS

The degree of antifragility resulting from the model structure will first be analysed by using a heuristic identified by Taleb (2011) and Taleb, Canetti, Kinda, Loukoianova, and Schmieder (2012). This heuristic is displayed in Equation (1) and allows assessing whether the system is antifragile.

$$H = \frac{f(a-\Delta) + f(a+\Delta)}{2} - f(a) \tag{1}$$

Equation (1) is the detection of a system's response to randomness where f(x) is the profit or loss for a certain system level a in the state variable and Δ is a change in a, representing a certain deviation from the variable's mean.

The heuristic works intuitively; when $H = 0$, the imposed randomness (represented as a change in the level a) has no effect, and the imposed gains and losses are cancelled out. When $H < 0$, the losses are larger than the gains implying fragility, and when $H > 0$, the gains are larger than the losses implying antifragility. However, as the use of Equation (1) is limited to only two data points, also, a sensitivity analysis for the whole model is necessary. This allows the simulation to generate a large variety of randomness, hereby testing the model under different circumstances. In our study, if the average level of the variable of interest after 2,000 Monte Carlo sensitivity simulations is higher than its original value, the model is evaluated as being antifragile.

Running the model without a sensitivity analysis does not yield enough black swan events to demonstrate whether the model is antifragile or not, which is in line with the assumptions in Taleb and Pilpel (2004). Therefore, we first use a "manually" generated black swan to investigate its influences on the model. This was done by implementing a negative PULSE function at $t = 10$ in the *decrease from randomness outflow* for the negative shock and implementing a positive PULSE function in the *increase from randomness inflow* at $t = 10$ for the positive shock. The resulting model behaviour is displayed in Figures 5 and 6. The figures show that irrespective of the size and the direction of the shocks, the new model equilibrium is higher than the initial equilibrium of 2,000. Consequently, the accompanied heuristic measure is positive, and the system can be considered as antifragile. The extent to which Figures 5 and 6 match with the theorized

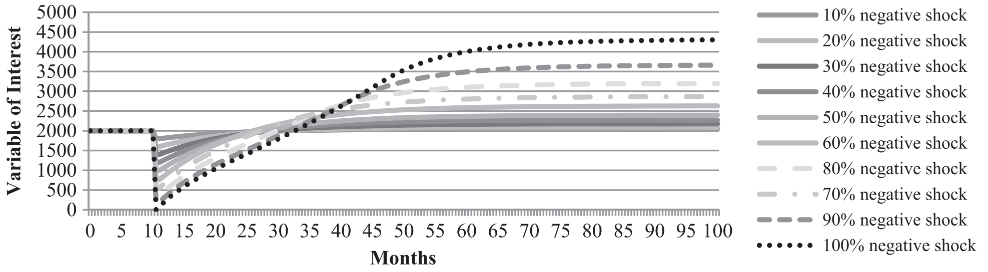

FIGURE 5 Antifragile behaviour patterns for different sizes of negative shocks. The time frame remains merely arbitrary and is set to 100 months, in order to limit the amount of data

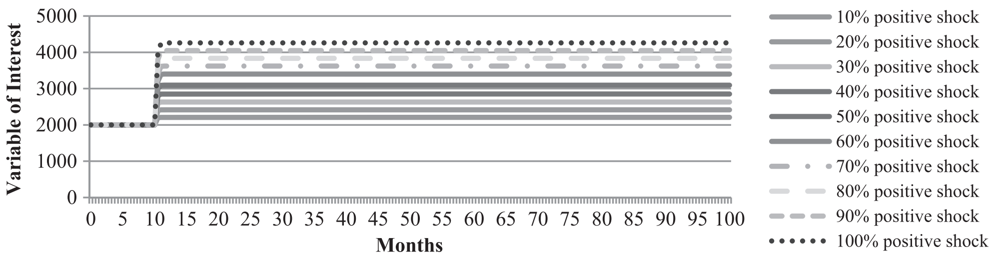

FIGURE 6 Antifragile behaviour for different sizes of positive shocks. The time frame remains merely arbitrary and is set to 100 months, in order to limit the amount of data

behaviour in Figure 2 further increases the evidence that the system can be considered antifragile. In addition, a plot of the heuristic measure in Figure 7 reveals that, as theorized in Figure 1, the measure increases with a convex shape for different pulse sizes. This implies that the system is characterized by a positive asymmetric reaction function leading to a convex response to randomness (Taleb, 2011, 2012; Taleb et al., 2012 In other words, the higher the volatility imposed on the variable of interest, the more this variable grows.

However—as pointed out before—antifragility does not merely involve benefitting from shocks but also the protection against future shocks. Therefore, a second shock should have less impact on the variable of interest than the first shock and a third shock less than the second shock. Tests in which the model was subjected to a series of external shocks confirmed that this is indeed the case as the variable of interest decreased less from every consecutive shock. An additional consequence of this effect is that the more pulses the model receives, the higher the future pulse needs to be in order to harm the system to the same extent. Moreover, it was found that the model, as proposed by the barbell strategy, indeed eliminates the negative consequences of black swans by means of a payoff floor (Taleb, 2012). All the above stated findings reinforce the result of the heuristic measure, indicating that this model is antifragile indeed. Specific information and results concerning these tests can be found in part B of the Supporting Information accompanying this article.

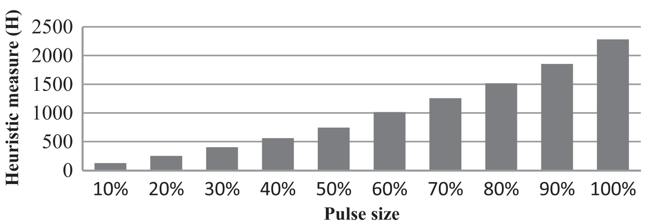

FIGURE 7 Increase in the heuristic measure (H) for the variable of interest for different sizes of shocks. Note that the limit to the increase can be varied by changing the carrying capacity

Finally, to assess the whole model in combination with the black swan structure, the model was subjected to a sensitivity analysis. Figure 8 shows that the black swan structure satisfies all the conditions with respect to black swan characteristics. The structure generates most variance relatively close to the mean, whereas it displays at rare occasions extremely high or low values at unpredictable points in the simulation. Also, these extreme values have a high impact on the model, contrasting the low impact of values around the mean. Thus, this structure replicates an environment where decision-makers are prone to be affected by black swans (Taleb & Douady, 2013). From this, it follows that model users who base their conclusions on a simulation time, which is too short or on too few simulation runs, potentially draw wrong conclusions about the environment in which they need to decide (Taleb, 2009, 2015).

Figure 9 confirms Figure 8 and shows that only in a small number of simulation runs black swans occur

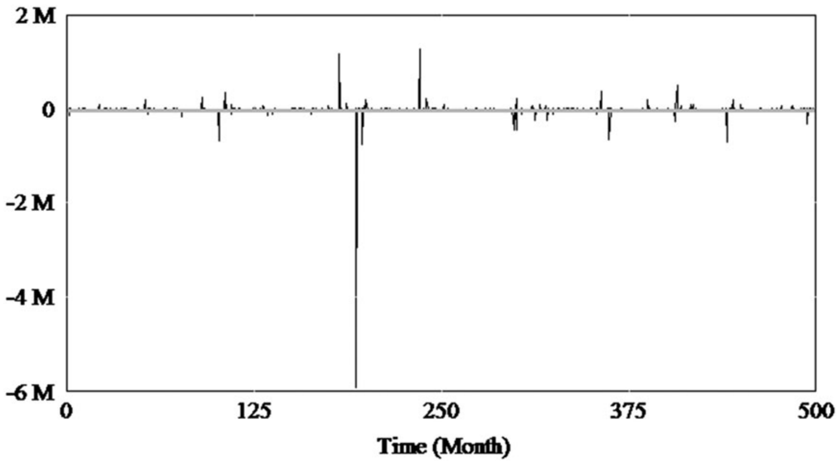

FIGURE 8 Individual traces for the "ratio actual over expected noise" after a simulation run with 2,000 iterations where the seed variable for the random number generator based on the normal distribution is varied. The ratio was chosen because it indicates whether the expected randomness generated by a normal distribution is similar to the actual one generated by the Cauchy distribution. Only large deviations from the expected randomness cause a large effect on the model as it represents an unexpected event. The mean is displayed in horizontal in grey and does not change significantly. For the sensitivity settings, see part D of the Supporting Information

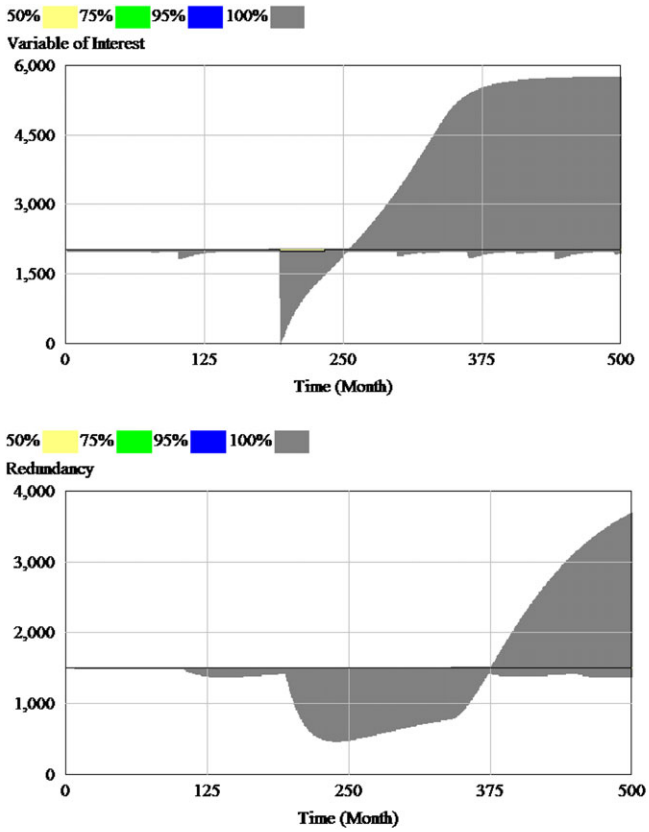

FIGURE 9 Confidence bounds for the variable of interest (top) and redundancy (bottom) in a sensitivity run where the seed variable of the normal distribution is varied 2,000 times [Colour figure can be viewed at wileyonlinelibrary.com]

(indicated by the large grey area signifying only 5% of all simulation outcomes; 95% of all simulations are bound close to the original values and hardly visible). In addition to this, it can be observed that, like the theorized system characteristics in Figures 1 and 2, model behaviour in Figure 9 (top) shows an increase after both negative and positive shocks. Thus, after 2,000 simulations runs, the mean for the variable of interest is increased. This result indicates that, complimentary to the antifragile evaluation after stand-alone PULSE shocks, the model can also be evaluated as antifragile when subjected to various unpredictable shocks. Furthermore, Figure 9 (bottom) shows that, initially, the redundancy level provides the source of growth for the variable of interest and only then increases its own level as the system grows.

5 | COMPARING ANTIFRAGILITY AND ROBUSTNESS

To investigate the difference between antifragility and robustness, the original antifragile model was adjusted to be "merely" robust. For an overview of the structural model changes, please refer to part E of the Supporting Information to this article. Figure 10 shows a comparison between the antifragile and the robust model version for the stock variable of interest. Figure 11 shows the same comparison but for the stock redundancy. In these simulations, both models were subjected to the following shocks; −1,800 at t = 10, 1,800 at t = 75, and −1,800 at t = 85. The shocks were implemented in the same way as the stand-alone black swan shock displayed in Figures 5 and 6.

Figures 10 and 11 show that the antifragile model learns from shocks and builds up its redundancy levels, whereas redundancy levels for the robust model can merely decrease from their initial values. Therefore, if the aim is to design a system that focuses on long-term growth, hereby being not only protected, but also benefiting from randomness, antifragility is clearly more desirable than robustness. A behaviour pattern indicator analysis (Richardson, 1995) and loop eigenvalue elasticity

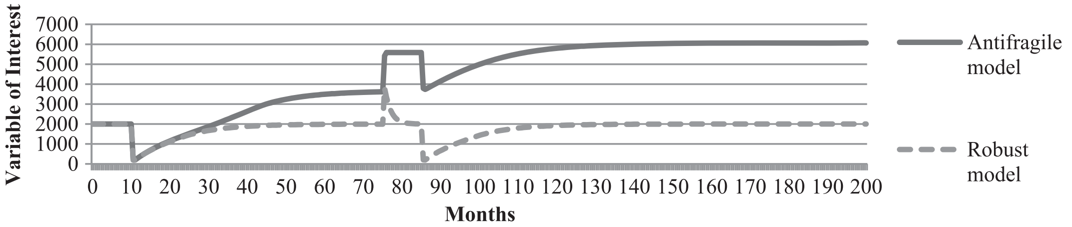

FIGURE 10 Comparison of model behaviours for the variable of interest for the robust and antifragile models

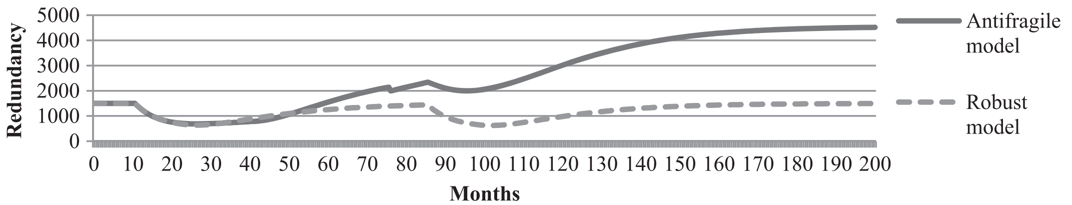

FIGURE 11 Comparison of model behaviours for the stock redundancy for the robust and antifragile models

analysis[8] (Kampmann & Oliva, 2008) revealed that the main differences in model behaviour between the robust and antifragile model are formed by the loops that were indicated in the antifragility dominance analysis as being necessary for antifragility: "Setting a new desired state (P3)," "Creating sufficient redundancy from harm (P4/5)," and "Believe positive events can be exploited." These loops require adjustments to the desired state of the variable of interest and, therefore, point to the decision-making processes and to the policies included in the system (Forrester, 1992). Thus, transforming a robust into an antifragile system requires significant attention to the way decisions are made and, in particular, to the way desired states are determined. This finding emphasizes the psychological aspect of decision-making, as desired states are often set and adjusted implicitly. Additionally, it pinpoints to the importance of mental models in decision-making as they govern the implicit rules in the decision-making process. A second significant attention point is formed by the advantages of antifragility when compared with robustness. Robust policies are based on predictability to increase efficiency (Makridakis & Taleb, 2009). Antifragile policies, on the contrary, cannot be based on predictability and consequently focus to benefit from unpredictable volatilities. The antifragile can, however, as argued by Markidakis and Taleb and demonstrated by the simulation model, only benefit when the system possesses sufficient redundancies. Reality shows that these redundancies, such as maintaining extra-large reserves, can carry significant costs. In transforming from the robust to the antifragile, it is thus key to evaluate the desire for long-term survival and the net effect related to benefits from volatilities and cost related to the required redundancies.

[8]For a presentation of the results of these analyses, please refer to part C of the Supporting Information.

6 | CONCLUSIONS

The result of using the antifragility design criterion is that —in an environment prone to black swans—our model shows more favourable behaviour than a "merely" robust model. It can therefore be concluded that antifragility as outlined by Taleb (2012) is useful. As a design criterion, antifragility yields additional value over the robustness criterion, though the antifragility criterion requires strong changes in the desired state of the system and hence the employed mental models of decision-makers. Additionally, redundancy needs to be present for an antifragile system to work. This research therefore arrives at the conclusion that antifragility as a design criterion cannot be given the status of a feature, which per definition needs to be attained for every system. Instead, it should be interpreted as a design criterion, which can lead to a normative system if the aim of the system is to increase a desired state level. Antifragility can, however, only be attained under certain conditions. Decision-makers in the system need to understand the mechanisms of antifragility and, therefore, be willing and able to invest heavily in the system's redundancy levels and increase the system's goal after black swan events. For modellers, this means that, contrary to merely focusing on the attainment of robustness, they should evaluate whether antifragility belongs to the possibilities for improving system performance.

This leads to the observation that the usage of both the design criteria robustness and antifragility should be context dependent. With respect to antifragility, several characteristics need to be met. First, it should be examined whether the environment of the system possesses thick tails and thus can generate black swans. Second, the system should already be robust as antifragility requires robustness to fluctuations in the left tail of the probability distribution (Taleb, 2011). Third, the aim of the operators of the system should be to let the system expand or at least to let it survive while maintaining at all times sufficient redundancy sources. Fourth, decision-makers in the system should possess some characteristics of personal mastery (Senge, 1990) allowing them to turn negative fluctuations to an advantage and to reduce structural conflict between lowering goals and perusing the vision (Garcia-Morales, Lloréns-Montes, & Verdú-Jover, 2007). Only then there is a chance that necessary changes to attain antifragility could diffuse into the organization as a whole. If the results of such an evaluation indicate that personal mastery is not present, there exists little chance that decision-makers support efforts to develop antifragility (Garcia-Morales et al., 2007). In such cases, antifragility should not be used as design criterion, and robustness should be the criterion to focus on.

Three directions for future research can be outlined. The first and probably most significant one comes from the notion that this research addresses antifragility without the inclusion of the notion that antifragility also comes from the fragility of its components. This would require the inclusion of shifts and emergences of new structures at lower levels of abstraction in the model. Thus, future research should focus on the relationship between fragility at a micro level and antifragility at a macro level, including the effects of small-scale experimentation mentioned above. Second, the exploratory character of this research implies that only the loops and the shapes of the graphical functions are precisely in line with the literature. The values in both the graphical functions and adjustment times are not empirically grounded. The next logical step for future research is then also to apply findings from this research to real-life cases and observe the bounds for the values used in the model in an empirical way. In particular, research focusing on the relationship between personal mastery and antifragility would be of value. Third, it is widely acknowledged that carrying redundancy usually results in costs and a loss of efficiency. Yet antifragility requires a high level of redundancy. Therefore, antifragility applied to a specific case needs to consider the cost-effectiveness of redundancy. This would imply investigating how redundancy costs relate to the antifragility of the whole system and develop strategies to minimize the cost of large redundancy sources, to guarantee survival of the system in the short term.

ACKNOWLEDGEMENTS

The authors acknowledge Radboud University Nijmegen (The Netherlands), NOVA University Lisbon (Portugal), University of Bergen (Norway) and University of Palermo (Italy) for capacitating the undertaking of this research within the European Master Programme in System Dynamics (EMSD). The first author was supported by the EACEA (Education, Audiovisual and Culture Executive Agency of the European Commission) during the EMSD programme. CENSE was supported by the Portuguese Foundation for Science and Technology through the strategic project UID/AMB/04085/2013.

ORCID

Harald de Bruijn https://orcid.org/0000-0001-5550-6736
Nuno Videira https://orcid.org/0000-0002-4514-1996

REFERENCES

Allen, P. M. (1988). Dynamic models of evolving systems. *System Dynamics Review*, *4*(1–2), 109–130. https://doi.org/10.1002/sdr.4260040107

Allen, P. M., Varga, L., & Strathern, M. (2010). The evolutionary complexity of social and economic systems: The inevitability of uncertainty and surprise. *Risk Management*, *12*(1), 9–30. https://doi.org/10.1057/rm.2009.15

Allington, N. F., McCombie, J. S., & Pike, M. (2012). Lessons not learned: From the collapse of Long-Term Capital Management to the subprime crisis. *Journal of Post Keynesian Economics*, *34*(4), 555–582. https://doi.org/10.2753/PKE0160-3477340401

Aven, T. (2013). On the meaning of a black swan in a risk context. *Safety Science*, *57*, 44–51. https://doi.org/10.1016/j.ssci.2013.01.016

Barlas, Y. (1996). Formal aspects of model validity and validation in system dynamics. *System Dynamics Review*, *12*(3), 183–210. https://doi.org/10.1002/(SICI)1099-1727(199623)12:3<183::AID-SDR103>3.0.CO;2-4

Bueno, N. P. (2012). Assessing the resilience of small socio-ecological systems based on the dominant polarity of their feedback structure. *System Dynamics Review*, *28*(4), 351–360. https://doi.org/10.1002/sdr.1476

Carpenter, S., Walker, B., Anderies, J. M., & Abel, N. (2001). From metaphor to measurement: resilience of what to what? *Ecosystems*, *4*(8), 765–781. https://doi.org/10.1007/s10021-001-0045-9

Chichilnisky, G. (2010). The foundations of statistics with black swans. *Mathematical Social Sciences*, *59*(2), 184–192. https://doi.org/10.1016/j.mathsocsci.2009.09.007

Coyle, R. G. (1977). *Management system dynamics*. Australia: John Wiley & Sons.

Coyle, R. G. (1996). *System dynamics modelling: A practical approach (vol. 1)*. Washington, DC: CRC Press. https://doi.org/10.1007/978-1-4899-2935-8

Dalziell, E. P., & McManus, S. T. (2004). Resilience, vulnerability, and adaptive capacity: Implications for system performance. Stoos, Switzerland: 1st International Forum for Engineering Decision Making (IFED), 5–8 Dec.

Derbyshire, J., & Wright, G. (2014). Preparing for the future: Development of an 'antifragile' methodology that complements scenario planning by omitting causation. *Technological Forecasting and Social Change*, *82*, 215–225. https://doi.org/10.1016/j.techfore.2013.07.001

Dhawan, R. (2006). Usefulness of probabilistic system dynamics in dynamic decision making. In *Proceedings of the 23rd International Conference of the System Dynamics Society* (pp. 1–13). Boston, MA: The System Dynamics Society.

Estrada, J. (2009). Black swans, market timing and the dow. *Applied Economics Letters*, *16*(11), 1117–1121. https://doi.org/10.1080/17446540802360074

Feller, W. (1971). An introduction to probability theory and its applications, vol. 2.

Forrester, J. W. (1968). *Principles of systems*. Waltham, MA: Pegasus Communications.

Forrester, J. W. (1992). Policies, decisions and information sources for modeling. *European Journal of Operational Research*, *59*(1), 42–63. https://doi.org/10.1016/0377-2217(92)90006-U

Garcia-Morales, V. J., Lloréns-Montes, F. J., & Verdú-Jover, A. J. (2007). Influence of personal mastery on organizational performance through organizational learning and innovation in large firms and SMEs. *Technovation*, *27*(9), 547–568. https://doi.org/10.1016/j.technovation.2007.02.013

Hekimoğlu, M., & Barlas, Y. (2010). Sensitivity analysis of system dynamics models by behavior pattern measures. In *Proceedings of the 28th International Conference of the System Dynamics Society* (pp. 1–31). Albany, NY: The System Dynamics Society.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*, 263–291. https://doi.org/10.2307/1914185

Kampmann, C. E., & Oliva, R. (2008). Structural dominance analysis and theory building in system dynamics. *Systems Research and Behavioral Science*, *25*(4), 505–519. https://doi.org/10.1002/sres.909

Kwakkel, J. H., & Pruyt, E. (2013). Exploratory modeling and analysis, an approach for model-based foresight under deep uncertainty. *Technological Forecasting and Social Change*, *80*(3), 419–431. https://doi.org/10.1016/j.techfore.2012.10.005

Makridakis, S., Hogarth, R. M., & Gaba, A. (2009). Forecasting and uncertainty in the economic and business world. *International Journal of Forecasting*, *25*(4), 794–812. https://doi.org/10.1016/j.ijforecast.2009.05.012

Makridakis, S., & Taleb, N. (2009). Living in a world of low levels of predictability. *International Journal of Forecasting*, *25*(4), 840–844. https://doi.org/10.1016/j.ijforecast.2009.05.008

Meadows, D. H., Meadows, D. L., Randers, J., & Behrens, W. W. (1972). *The limits to growth* (p. 102). New York: Chealse green publishing.

Miller, J. H. (1998). Active nonlinear tests (ANTs) of complex simulation models. *Management Science*, *44*(6), 820–830. https://doi.org/10.1287/mnsc.44.6.820

Moxnes, E. (2005). Policy sensitivity analysis: simple versus complex fishery models. *System Dynamics Review*, *21*(2), 123–145. https://doi.org/10.1002/sdr.311

Moxnes, E. (2007). Guidelines for the analysis of complex, dynamic systems. In *Proceedings of the 2007 International System Dynamics Conference* (pp. 1–31). Boston, MA: The System Dynamics Society.

Paté-Cornell, E. (2012). On "black swans" and "perfect storms": Risk analysis and management when statistics are not enough. *Risk Analysis*, *32*(11), 1823–1833. https://doi.org/10.1111/j.1539-6924.2011.01787.x

Pruyt, E. (2007). Dealing with uncertainties? Combining system dynamics with multiple criteria decision analysis or with exploratory modelling. In *Proceedings of the 25th International Conference of the System Dynamics Society* (pp. 1–22). Boston, MA: The System Dynamics Society.

Quinn, D. P., & Woolley, J. T. (2001). Democracy and national economic performance: The preference for stability. *American Journal of Political Science*, *45*, 634–657. https://doi.org/10.2307/2669243

Richardson, G. P. (1995). Loop polarity, loop dominance, and the concept of dominant polarity (1984). *System Dynamics Review*, *11*(1), 67–88. https://doi.org/10.1002/sdr.4260110106

Richardson, G. P. (2011). Reflections on the foundations of system dynamics. *System Dynamics Review*, *27*(3), 219–243. https://doi.org/10.1002/sdr.462

Rydzak, F., & Chlebus, E. (2007). Application of resilience analysis in production systems – Bombardier transportation case study. In *Proceedings of the 2007 International Conference Of The System Dynamics Society* (pp. 1–18). Boston, MA: The System Dynamics Society.

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. New York: Doubleday/Currency.

Senge, P. M., & Forrester, J. W. (1980). Tests for building confidence in system dynamics models. *System dynamics, TIMS studies in management sciences*, *14*, 209–228.

Sharp, J. A. (1976). Sensitivity analysis methods for system dynamics models. In Proceedings of the International Conference on System Dynamics, Geilo, Norway (Vol. 152).

Sniedovich, M. (2012). Black swans, new Nostradamuses, voodoo decision theories, and the science of decision making in the face of severe uncertainty. *International Transactions in Operational Research*, *19*(1–2), 253–281. https://doi.org/10.1111/j.1475-3995.2011.00790.x

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Boston: Irwin/McGraw-Hill.

Taleb, N. N. (2007). The black swan: The impact of the highly improbable. Penguin.

Taleb, N. N. (2009). Errors, robustness, and the fourth quadrant. *International Journal of Forecasting*, *25*(4), 744–759. https://doi.org/10.1016/j.ijforecast.2009.05.027

Taleb, N. N. (2011). A map and simple heuristic to detect fragility, antifragility, and model error. NYU Poly, https://doi.org/10.2139/ssrn.1864633.

Taleb, N. N. (2012). Antifragile: Things that gain from disorder. Penguin.

Taleb, N. N. (2015). Silent risk. Retrieved from https://drive.google.com/file/d/0B8nhAlfIk3QIR1o1dnk5ZmRaaGs/view

Taleb, N. N., Canetti, E. R., Kinda, T., Loukoianova, E., & Schmieder, C. (2012). A new heuristic measure of fragility and tail risks: Application to stress testing (IMF Working paper 12/216). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.259.1728&rep=rep1&type=pdf

Taleb, N. N., & Pilpel, A. (2004). On the very unfortunate problem of the nonobservability of the probability distribution. Unpublished Manuscript. Retrieved from http://www.fooledbyrandomness.com/knowledge.pdf (June 7th, 2006)

Taleb, N. N., & Douady, R. (2013). Mathematical definition, mapping, and detection of (anti) fragility. *Quantitative Finance*, *13*(11), 1677–1689.

Tseitlin, A. (2014). The antifragile organization. *Communications of the ACM*, *56*(8), 40–44.

Walker, B., Holling, C. S., Carpenter, S. R., & Kinzig, A. (2004). Resilience, adaptability and transformability in social–ecological systems. *Ecology and Society*, *9*(2), 5. [online] URL: http://www.ecologyandsociety.org/vol9/iss2/art5/

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: de Bruijn H, Größler A, Videira N. Antifragility as a design criterion for modelling dynamic systems. *Syst Res Behav Sci*. 2019;1–15. https://doi.org/10.1002/sres.2574