COMMENTARY

Balancing the perceptions of NK modelling with critical insights

Richard J. Arend[*]

*Correspondence:
richard.j.arend@gmail.com
L.L. Bean/Lee Surace Chair
in Strategic Management,
Business Administration,
School of Business,
College of Management &
Human Service, University
of Southern Maine, P.O.
Box 9300, Portland, ME
04104, USA

Abstract

NK models are agent-based simulations of market evolution generated through new entry and firm innovation, and are often focused on better understanding complex interdependencies in organizational phenomena. We provide a counterpoint to the mostly optimistic descriptions of the advantages and of the requirements for these fitness-landscape-based analyses. We do so by offering a comprehensive list of limitations of that modelling technique as well as a critical analysis of two recent applications of the NK model—one of a theory-testing application and one of a theory-building application. Our analysis reveals that when care is not taken to capture the essential parts of a phenomenon, the NK approach may be unnecessary at best, and misleading at worst. We discuss the implications of these analyses and update past suggestions for future uses of NK simulations in organizational research.

Keywords: NK models, Organizational science, Theory building, Theory testing, Simulations

JEL Classification: C63, C61, C53, D02, D21, C73

Introduction

We offer this *commentary* on NK landscape-based modelling to balance the overly rosy picture of this unique type of research method (Wall, 2016) that currently exists in the literature (e.g., Csaszar, 2018). We offer the first explicit list of the limitations and dangers of NK landscape models. We also illustrate in detail what can go wrong when such modelling is misapplied by analyzing two examples of research that have appeared in high-quality outlets. It is important to consider the downsides of this methodology, because, even though there are over 70 published studies using the methodology since 1997 (Baumann et al., 2019), with such research continuing to appear in top journals, there remains a lot of misunderstanding on how it works. The main reason for such misunderstanding is that it is a complex and unfamiliar methodology, so personal experience using it is indispensable to properly review any contribution that applies it. However, unfortunately, such experience is not generally provided in doctoral programs (Baumann, 2015), so having such work properly reviewed and understood can be challenging. This commentary exists to clear up misunderstandings. Specifically, we wish to balance the perspective that the NK-based methodology has only led to positive and

significant contributions to the relevant literatures in innovation, entrepreneurship and organizational design. We do so by offering a less positive counter-perspective, and one that concludes that its role should be delegated to acting as a complement rather than as a substitute to more traditional research methods (Wall, 2016).

As Fioretti (2013) explained several years ago, the NK-based methodology offers an agent-based simulation used for (pseudo-) quantitative research in organizational studies. It is known for providing tunable rugged landscapes that coded agents traverse in search of local and global fitness optima in an evolving manner. Originally conceived for theoretical biology (e.g., Kauffman & Weinberger, 1989), it has existed in its current form for over 30 years. As a tool of research, it has several advantages in specific applications, because it provides precise control over factors, a stable evolution of the agent population to greater fitness, and the ability to collect a vast amount of data based on thousands of runs involving hundreds of interdependent agents, all at a very low relative cost. It has been established as a legitimate methodology by famous scholars at famous institutions, starting with its application in the natural sciences, where the theoretical laws of primary interactions are stationary (unlike in business). However, as with all methodological tools, the NK model can be misused or stretched in use. Such stretches are especially problematic when such a tool crosses into new fields, such as management, innovation and entrepreneurship—fields that all involve social phenomena that often entail much different interactions than found in biology. Unlike many other tools in business, however, the reviewer pool for NK simulations is relatively shallow in our fields, given it takes substantial extra training to understand and deconstruct these models and their mechanics. Due to that shallow reviewer pool, the possibility for greater tolerance of the misuse and stretch of this tool may be higher than for other methodologies. Regardless, NK modelling remains a powerful and growing niche of research in business.

Why is that important? It seems self-evident, but worth stating that the question of *whether or not a research method is applied properly* is one that can significantly affect a study's results and prescriptions. For example, if the NK model is *in*correctly applied, then any observed outcomes, along with the pseudo-empirical supports and the policies that follow them, are all likely to be erroneous as well. Such new methods should not be thought of simply as involving nothing but upside value, but should also be considered as potentially dangerous solutions that, if applied incorrectly, can mask the damage they can do. Our contribution is to offer a counterpoint to optimistic descriptions of the NK model that currently exist, so that researchers and readers can take the results from NK model-based analyses with much greater and justified skepticism.

Methods and background

The method we use to provide our counterpoint is critical and constructive analysis of the NK model approach. That analysis is first aimed at the standard NK model and leads to a comprehensive list of its limitations. The analysis is then aimed at two illustrative recent examples involving issues with NK model applications and the efficacy of the results. To move forward, we also add some advice for improving upon future NK model-based research.

The NK model and its two most powerful uses

We recap the basics of the NK model methodology prior to describing its powerful applications. In an NK simulation model, each automaton (*aka* agent, player or firm) is represented by a string of N genes. For simplicity, the usual coding is binary for each gene (it has a 0 or 1 value). For each string, a fitness score (scalar) can be computed from a function (usually additive) that involves the value of each gene combining with the values of its K neighboring genes. A landscape can then be calculated from all possible strings and their corresponding fitness values; the smoothest occurring when $K=0$ and the most rugged when $K=N$-1. One can then visualize an agent with a specific string innovating so as to alter their string to move higher in their fitness landscape (where higher means better—e.g., being more profitable). The evolution of the population of these agents occurs by coding how they can alter their genes (with varying constraints and/ or costs) as each agent explores (and/ or imitates) a restricted set of local (or rival or random) gene variants. These agents do so simultaneously and independently. In addition, it is standard to replace the agents having the lowest fitness with new agents, each of which is provided a new string of randomly assigned gene values. Over time, given the coded-in ultimatum to increase fitness, the population moves towards stability at local or global optima (i.e., as embodied by the higher points in the rugged landscape).

It is important to note that the coupling of attributes in the landscape and across each agent's gene string (denoted by K) does *not* capture the complex interdependencies inside the firm with which a manager or entrepreneur must deal. Instead, the standard NK landscape used in these studies refers to a technical set of binary choices over a firm's nature (i.e., the steps in a process it uses, or the attributes of a product it sells) with a static, universal-to-all-firms optimal innovative design (of the process or product) that can be exogenously pre-determined by the coder (but not by the simulated firms). It is also important to note that the NK simulation does *not* actually provide an intuitive 3D (x–y–z) landscape to traverse virtually, although it is almost always depicted as such. Instead, each location [for $N>3$] is actually a corner on a difficult-to-visualize N-dimensional hyper-cube that projects payoffs outwards. With those basics of the NK model methodology covered, we can now consider its relevant applications.

The two most powerful uses of the NK model in entrepreneurial, innovation and management research involve activities supporting theory-building (Baumann, 2015; Wall, 2016). In the first use, the NK model provides a means to test or extend an existing theory. In the second use, the NK model provides a means to induce a new partial theory. Each use embodies a mediating role between organizational *reality*—real-world observations or practices—and organizational *theory*—abstractions of the key relationships among factors that are believed to drive the outcomes observed. In each use, the NK simulation generates a lot of life-like data in a controlled and particular context. When the agents are coded to act as if they are following the principles of an existing theory, doing so in a specific and difficult-to-replicate-in-reality context, then an empirical analysis of the simulated outcomes *can* provide some support for a theory's predictions. Alternatively, when an existing NK simulation is tweaked and added to in new ways, where some of these experiments provide simulated outcomes

that mimic real-life-observed patterns, then the driving forces represented by those tweaks can be assessed to see if they provide a coherent alternative explanation (*aka* a proposed new induced partial theory) for that seemingly mimicked phenomenon.[1]

The reason why such *mediating* roles are sought in social science fields such as innovation management and entrepreneurship is that methodologies such as NK modelling can provide what-looks-like-data when real-world data is much more difficult or impossible to obtain. In the social sciences, we cannot always feasibly experiment on our phenomena-of-interest directly (e.g., it is impossible to run a controlled, repeated experiment on the economy to see how it reacts to a number of alternative entrepreneurship-promoting policies). Therefore, we turn to indirect ways—and that means using models. Sometimes a formal model can provide insights into outcomes through logical arguments. Sometimes a closed-form mathematical model can provide optimizations and explanations for more complex systems of relationships and their outcomes. However, when the phenomenon cannot be properly modelled in a box-and-arrow form, or with solvable mathematics, then more sophisticated methods such as simulations can be effective. Such computational simulations generate sort-of-real data, especially when the simulation is calibrated to a set of real observations for specific conditions and when the coding is based on accepted theoretical premises. In the best cases, a simulation acts as a scaled-down and simplified model, just as a wind-tunnel provides a small and controllable model of real fluid dynamics, although at a different Reynolds number and enclosed by tighter boundary conditions. Here, the NK simulations are the wind-tunnels that can provide insights on the performance of existing organizational, product and process designs under new conditions (i.e., to test existing theories under various new constraints) as well as on possible new drivers of established, recognized outcomes (i.e., to induce new theory based on possible new inputs into processes that are controllable and visible in a simulated world).

Research scholars in entrepreneurship, management, process innovation and strategy have indeed attempted to use the NK model in such ways. It has been used to translate between reality and theory (i.e., to identify potentially new simulated drivers of outcomes that produce near-real-world observations) and between theory and reality (i.e., to test whether theoretical predictions generate their expected outcomes in a simulated realistic economy) when real phenomena were hard to control, idiosyncratic, adaptive, subject to deception, and so on.

The option of using the NK model in such mediating roles is one of the more attractive choices for several reasons: it is inexpensive. It generates a magnitude of data for high statistical power, including data that can easily be used to visualize the evolution of fitness and identify the eventual equilibria under different assumptions (e.g., assumptions about the firm's internal coupling of processes, and about the firm's types of interactions with its environment). In addition, it *looks* like hard science. It does so, because it involves coding (i.e., with explicit statements of the model assumptions), mathematics,

---

[1] As alluded to, we consider the NK model as a useful tool, but more as a complementary rather than a validly independent methodology. In that complementary role, it has value in testing boundaries, discovering discontinuities (and other unusual nonlinearities) in existing theories, and in testing the robustness of proposed possible real-world measures, all when data is not easily available (e.g., Ganco, 2017; Wall, 2016).

and many principles established in the natural sciences (e.g., evolution); in addition, it carries legitimacy from being a method established in a hard sciences.

Results

We now describe the outcomes of our analysis. Unfortunately, it appears that the application of NK models to business research has, at times, failed. However, this has *not* been sufficiently acknowledged and analyzed. To address that deficit, we explain several reasons for such failures, first by critiquing the NK model generally, and second by critiquing two specific but representative example applications.

Limitations of the standard NK model in research

In Table 1, we list the main reasons as to why the depiction of business reality through an NK model can be poor. Note that it should not be surprising that it can be poor, given the NK simulation methodology was *not* created to model managerial and entrepreneurial phenomena. The NK model was not written to capture human-designed organizational behavior or structure, nor was it written to depict how such organizations innovate and compete. In addition, it was not written to track interdependent, multiple performance measures but rather to assess only one landscape-terrain-shape-defining payoff output (i.e., the fitness scalar). Instead, the NK model was written to describe groups of similar individual entities with the same potential capabilities and constrained by the same length of one genetic string competing on an inert and stationary landscape independently, where survival can be relatively (and often absolutely) dependent on just one instantaneous performance measure.

In Table 1, we provide two dozen concerns about the basic NK modelling method. While we acknowledge (in the table) that there exist individual exceptions to many of the specific critiques, the existence of such exceptions actually reinforces the point we are making in this commentary. If the modifications we suggest are important enough to be published (in the pieces that we consider exceptions), then the critiques of the unmodified base model must hold significant validity. In addition, if those modifications continue to be exceptions—if they are not systematically adopted in each future version of the updated-new-base-model—then that provides an additional challenge to the efficacy of a (insufficiently changing) methodology that continues to be applied to business phenomena.

Every formal modelling method involves the sacrifice of realism (McGrath, 1981), and the NK model methodology is no exception. However, our concern is more nuanced, and is offered to counter-balance the often unquestioned portrayal of NK models as providing only legitimate data-as-evidence in their applications. We raise the possibility that such data can less-than-reliable when, for example, poor coding fails to capture important aspects of the phenomenon being studied. The sacrifice of realism is only warranted when the model's product is useful either in the abstract or the real, and that is not always the case with this often less-understood methodology. As such, we see it *not* as an independent solution to better understanding complex phenomena, but rather as one of the complementary research methods that allows researchers to engage in a process of strong inference when confronting such phenomena with a *mix* of methods (Platt, 1964).

Table 1 Analysis of inappropriate NK assumptions

Questionable Assumptions for Managerial Phenomena	Notes
Managers can only see better positions within a specific, local neighborhood (the standard *small jumps* restriction) [see below for when *big jumps* are allowed] Fixed behavioral rules (oftentimes conditional) that are not modified in response to feedback (Baumann, 2015)	Makes more sense for genetic improvements than for business ones. Search cost functions don't appear to affect this There is empirical support for modelling local search as limited to the local neighborhood (e.g., Conlisk, 1996) Is a very specific way to model bounded rationality; humans are more intelligent than the limited adaptive automata modelled (Baumann, 2015; Csaszar, 2018). The few exceptions to this restriction include mental model based searching (Csaszar & Levinthal, 2016; Gavetti & Levinthal, 2000) The conditional part is affected by feedback (e.g., by a failure to improve—Csaszar & Siggelkow, 2010; by too slow an improvement—Csaszar & Levinthal, 2016), but the rule itself does not change
Managers can immediately act to exploit an identified better position	It is unrealistic to model, for a simulation, that no constraints, frictions or delays exist for firms to alter a product or process
Managers cannot follow the steepest gradient for improvement; instead, they retain no path memory and simply jump to the next better location (if one exists). There is no learning (Puranam et al., 2015)	Organizations and managers have memories (and path dependencies), and their consistency along a tactical path is usually expected for planning purposes. They learn Alternative performance feedback responses—that may be more realistic—remain unconsidered Few exceptions to this restriction, some involve explicit memory modelling (e.g., Jain & Kogut, 2014)
When *big jumps* are allowed, it is modelled as a random draw (from a uniform distribution) or as a costless and perfect imitation of a more successful rival (Csaszar & Siggelkow, 2010)	Real firms cannot do random full transformations, let alone without frictions or extra costs (so why not model some restrictions?)
No direct costs to altering the DNA of the organization/ entity. All changes cost the same, and are equally effectively completed	Change is costly in the real world, and dependent on the type, timing and technique of the change. Change is also often considered a source of competitive advantage (e.g., in the DCV), but that is ignored here
The total number of entities on the landscape remains constant. (Few exceptions—these include Adner et al., 2014.)	Yes, this is easier to code, but the context here does not support it. The idea that all deaths are replaced by births (or potential entrants) is not realistic, nor are the implicit restrictions on firm growth in scale, or via franchising or buyouts in the NK simulation proper
New firms (births) have random DNA or imitate currently successful DNA	This isn't realistic for business. Such births would not get investment, because they have no expected advantage over incumbents
It is a simultaneous move game, where all other information (i.e., about the opponents, payoffs and random draw functions) is known with certainty. Any one firm can do what all other firms can do if it is in the same location	This type of information set and this type of homogeneity is unrealistic in business. It can provide a benchmark in isolation, but with all the other assumptions added on, it stretches what relevance the simulated output provides. Furthermore, this assumption provides little room for managerial discretion or function at all Often, knowledge of payoffs is inaccurate (Puranam et al., 2015) Very few exceptions [e.g., Rivkin and Siggelkow (2003) include heterogeneity in firms in terms of the number of alternatives seen]

Table 1 (continued)

Questionable Assumptions for Managerial Phenomena	Notes
Travel across the landscape is only based on either pure path dependency (is continuous) or luck (is discontinuous)	Where is the room for managerial strategy rather than simple heuristics in this model? It seems inconsistent to assume simple decisions to analyze outcomes so as to prescribe non-simple decisions Heterogeneity in policies emerges as stable due to path dependencies in rugged landscapes, which may not be realistic for rational, informed decision-makers (Puranam et al., 2015) Few exceptions (e.g., Csaszar & Levinthal, 2016 include a parameter for heterogeneity in landscape attribute attention that affects travel)
The initialization of the landscape and the initial population are based on random draws from a uniform distribution (i.e., for the DNA elements and for K-type interactions)	Why is there no symmetry imposed for the K-type interactions among the same elements, and why are no population dynamics taken from related landscapes? Yes, it is easier to code and may maximize initial entropy (Jaynes, 2003), but those justifications are questionable in a business context, where structures do exist. The few exceptions to the base case appear to recognize that fact [e.g., Posen and Martignoni (2017), where the initial population imitates good performers; Albert and Siggelkow (2022) and Rivkin and Siggelkow (2003) control initial populations for specific characteristics]
The landscape usually remains fixed throughout the analysis (There are exceptions for an NKC version of the simulation, and for models that test against shocking the system [by altering the landscape during a run] for checking the robustness of specific strategies) Static environments are too abstract for modelling some important problems (Baumann, 2015; Ganco & Hoetker, 2009) Payoff structures are exogenous (Gavetti et al., 2017) The search space is exogenous (Ganco et al., 2020)	Where is the *co-evolution* of the environment with the players interacting with it? The *C* in NKC includes modelling of some effects of cooperation and competition on rival landscapes. In addition, the shocks can capture some meta-phenomenon effects. However, single landscape co-evolution is missing in the basic NK approach, although such co-evolution may be a more realistic accounting of many management phenomena Some endogeneity of payoffs and the space is more realistic Few exceptions exist [e.g., Rivkin and Siggelkow (2003) model some turbulence in the landscape; Gavetti et al. (2017), and Li and Csaszar (2019) include some limited ways to shape the landscape]
The survival rule is imposed with immediacy, eliminating the current lowest performers (with either certainty or high probability)	Where would *Amazon* and other long-play-strategy firms [firms that did not report profitability for years] be in this model? Does it seem proper to exclude such major recent success stories with this approach? Very few exceptions [e.g., Csaszar and Siggelkow (2010) do not eliminate firms]
Firms can engage in both local and distant search	Why is it proper to assume this kind of ambidexterity? Regardless if search impact comparisons are made, why not assume that there are specialists in each search type instead, as each is likely to involve different skills?
The most common steps repeated in the simulated timeline are: identify deaths, conduct survivor choice searches, replace deaths, allow action, calculate outcomes, and repeat.	The steps lack interdependence (unless done in a NKC simulation, where there is a Stackelberg-like sequencing of move-countermove). This homogenizes search and action (e.g., in terms of efficiency) across all players, which is not realistic Such steps highlight the ecological roots of a method that assumes such ordered and linear processes (Hannan & Freeman, 1977)
Models the on–off switches (0–1) in the DNA instead of qualitatively different choices for each factor in organizational management (A–B–C…). N-dimensional binary vector optimization constitutes a strong abstraction from real world problems (Wall, 2016).	Easy-to-code, but misses the point of *interior optima* (and the tradeoffs involved among factor levels) that occur in the real world more than extreme optima (e.g., hitting boundary conditions) Lack of external validity (Wall, 2016) Very few exceptions to the two-level model (e.g., Rahmandad, 2019)

Table 1 (continued)

Questionable Assumptions for Managerial Phenomena	Notes
Changing one of the N-genes does not affect another directly; it affects payoffs through the K-function that involves the other genes.	This does not seem realistic for product or process design alterations (e.g., with power supplies, platform choices, and so on), where the effects are seen in the realizations of the product itself rather than in its gross revenues.
The K-based function is *not* in a universal form, but entails a different sub-function for each subset of genes (and is not even symmetrical in those effects between gene pairs).	This seems like an arbitrary choice of functional form rather than one that has parallels in business or engineering (and is inexplicably restricted to effects on the closest other genes without any check on why that closeness appears in the first place).
It is possible to have more than one global maximum (e.g., the neutrality modelled in Jain & Kogut, 2014).	Often unrealistic in business (e.g., in standards wars).
Restricted to one dependent variable (DV)—i.e., the landscape-height-as-payoff fitness measure.	Other DVs are important (e.g., speed to payoff) that are both simulation-based and reality-based (e.g., market share; brand; corporate social responsibility; carbon footprint; and, so on). Real organizations face a concurrency of multiple and conflicting performance measures (Baumann, 2015; Puranam et al., 2015). Very few exceptions that model multiple fitnesses (e.g., Adner et al., 2014).
Involves the trick of a 3D landscape representation of an N + 1-dimensional game, and the power of being able to envision both rough versus smooth terrains and the physical traverse of that landscape to higher ground	While the analogy appeals well to basic human experiences and visual abilities, simplifying very complex competitive strategic management decisions in this way is likely to be misleading (and dangerously confidence-building)
It is possible—with non-structured models—to rig them to produce the desired results (Wall, 2016). NK models are often seen as a black boxes to most readers and especially to most practitioners. Model specifications are seen as idiosyncratic to the researcher (Ganco & Hoetker, 2009).	Given the standard, structured NK model has limits and has been extensively used, modified models are becoming more popular. This increases the rigging worries because of the black box effect (from the unfamiliar modifications). Behavioral rules are sometime then introduced in an ad hoc basis, without empirical validation, and sometime only based on stylized facts (Wall, 2016).
The analysis of the model's substantial data output is done through extensive numerical derivation (e.g., filtering, smoothing, regression, and so on) to identify reported patterns.	Low traceability of variance, such that reported data does not represent all outcomes (Wall, 2016). Difficult to isolate complementarities (Ganco et al., 2020) that are of managerial interest, and which may be discoverable in reality.
Firms (as simulated agents) are not competing for resources with each other (Ganco et al., 2020).	In reality, firms do compete for resources—horizontally and vertically.

As Table 1 details, there are many concerns with a sole reliance on the NK model methodology. For example, there exist fundamental incompatibilities with its application to organizational phenomena (e.g., to the characteristics of the actors, optimizations and interactions involved) that severely limit how effectively it, alone, can capture any real focal research issue. In addition, even if that methodology could capture a real phenomenon in any given run of a simulation, verifying that that specific model run's variable values were applicable to a given manager's specific decision problem is generally impossible.[2] Instead, the power of the NK simulation most often emerges from the visualizations of the main outcome *patterns* that are revealed over thousands of specific runs as

---

[2] Note that it is unusual for any NK model's parameter values to be synchronized to those in the real world; it is unusual because such parameters – e.g., organizational genes – almost never exist.

one focal factor value at a time is varied to see its average effect. The problem with such powerful visualizations is that any one individual organization is *not* represented by a de-noised version of an average firm in a context, where all other factors are held constant. Rather, in reality, each firm is unique in space and time, and that uniqueness often drives its relative performance as well as the identity of what action is best for a manager to choose in that particular dynamic situation.[3] The average pattern does not correspond to any one firm's choices and performance, so extrapolating from the average—in the midst of likely contingencies—provides *little* to no guarantee of improved performance.

Regardless, the NK approach certainly looks like an attractive tool with a history of well-respected users. It is a methodology that adds to the diversity of our approaches for addressing complex problems, and that is valuable. It provides a cheaper way to generate data than surveying real firms struggling with real decisions. In addition, it can shorten the descriptions of real and complex problems by referencing foundational and related NK modelling in business (e.g., Levinthal, 1997).[4] Despite its many legitimate advantages, the application of this methodology is often flawed, as we point out in the two examples below.

Critiques of illustrative example NK model applications in business

Table 2 provides the main critical issues and faults relevant to two specific NK modelling method applications capturing its two most powerful uses. One example application relates to theory-testing; the other relates to theory induction. In the first example—a recent piece in entrepreneurship—Welter and Kim (2018) use the NK model to test a focal theory (i.e., of effectuation—see Sarasvathy, 2001). They employ the NK model to determine whether that theory's prescriptions perform better than alternatives across specific conditions. They code simulated firms to traverse an NK landscape using decision-making logic based on the focal theory's prescriptions, and compare the outcomes to when a different set of prescriptions, based on an alternative decision-making logic (i.e., of causation), are coded. They also code landscape shifts—i.e., sudden alterations in how fitness is computed from the N genes—to test the performance differences of those prescription sets across those various contextual conditions.

Using an NK model to test a theory requires, at a minimum, that its coding faithfully captures the original theorizing (Fioretti, 2013).[5] Here, it also requires that the coding accurately depicts the various competitive landscape shifts. Unfortunately, in the paper, *none* of these are captured accurately: the focal theory's relevant components are *not* all coded (i.e., many defining characteristics of the focal decision-making logic of

---

[3] One irony about NK modelling is that the core *local experimentation process* that is hard-coded into the decision-making of the automaton firms actually describes a better approach for real managers to take in most exploratory contexts than any approach based on the analysis of the model's output patterns emerging from those simulated experiments.

[4] That said, sometimes this shortening is inappropriate because it can cut down on much needed debate over the necessary modifications to any referenced foundational model that is the basis of an application of the method to a new organizational phenomenon. Also, note that every new method eventually gets stretched outside of its applicable bounds. So, it should *not* be surprising that that has happened with the NK model methodology. However, in its case, we suggest that those bounds are quite tight, and mostly limited to making specific points about evolution-based patterns arising from variations in organizationally-bound structural categories (e.g., internal coupling) or in contextual characteristics (e.g., local versus global search), as Nelson and Winter (1982) and others have shown in their versions of evolution-driven simulations.

[5] It would have been prudent to first test whether the focal theory's own full logic was internally consistent when fully coded prior to testing it against other logics. Unfortunately, that was not done.

Table 2 Analysis of examples of how the NK approach is a poor mediator in either direction

Reality-to-Theory Mediation Example using NK Simulation	Theory-to-Reality Mediation Example using NK Simulation
Lenox et al. (2007) *Management Science* piece offering an alternative explanation for observed patterns of industry evolution based on a simulation involving an NK model.	Welter and Kim's (2018) *Journal of Business Venturing* piece testing the logic of effectuation through an NK simulation.
Main Issues: Their NK model is used as a feeder-of-pseudo-data (i.e., of a firm's cost levels) into a static Cournot competition game (that updates the fitness outcomes in the NK model). This two-step procedure with feedback is repeated to mimic several generally observed patterns of industry evolution. In the paper, the NK model acts as an intermediary means to get from reality to theory-building (through its use in a process that appeared to mimic real outcomes). The first issues is that a more applicable model (e.g., the NKC simulation) for the phenomenon was not used. The second issue is that a less-confounding explanation was available (Ganco et al., 2020). These issues raise legitimate suspicions over the conclusions reached.	***Main Issue***: Their NK model does *not* capture the theory being tested (e.g., here, it does not actually model effectuation logic's five parts) nor the contexts in which it is being tested (e.g., known versus risky versus uncertain landscapes). Thus, any findings of support for the theory's robustness may be misleading even though they appear legitimized by appearing in a top specialty journal.
When translating from reality to theory there are simplifications needed to capture the main elements of the phenomena, but many of the simplifications of an NK simulation are ill-fitting to managerial–strategic phenomena. Feeding the results of such a simulation to *another* simplified model of reality (i.e., Cournot competition) may actually *amplify* those simplifications (i.e., about capturing reality with sufficient accuracy) without proper understanding of those interactions.	The legitimacy of the NK model methodology—as alluded to through citations—was leveraged to justify its use in their application to theory-testing in this instance. The reviewer pool—perhaps thin in terms of the overlap of expertise in both NK modelling and effectuation required—failed to pick up on important issues even when the explicit coding provided clearly revealed what was modelled in what way and what was not.
Consider some of their NK model simplifications: N is constant (but in the real world the list of product characteristics and process steps usually increases over time); search is mostly local (but can involve the limited imitation of the best rival) and is costless (whereas in the real world, no search is costless, and imitation can violate intellectual property protections); search is constrained in artificial ways; changes are done gene-by-gene and are costless (whereas in the real world, changes often affect more than one element and the costs reflect that); firms only participate in the industry when profitable (whereas in the real world, especially for new firms, this is unusual at least in the short-term); competition is solely cost-based (whereas in reality, most products are not commodities); firms face no entry or exit costs (but such costs exist in reality, even in Cournot models through fixed costs); firms can alter scale instantaneously and without cost (which is unrealistic other than for digital goods); and, searches are perfectly accurate (whereas in the real world, firms spend resources to spread disinformation, especially about profitability and imitability, to generate causal ambiguity). The applicability of their model specification hinges on the assumption that the interaction dynamics do not alter the shape of the production function; instead, it only shifts the function in terms of costs (Ganco & Hoetker, 2009)—this appears unrealistic in most industries of interest past the short-term.	The simulation did indicate one thing clearly—i.e., that a more flexible decision rule (i.e., one held for fewer periods) outperforms a less flexible one when confronting a changing landscape (unless the firm—through its rule—can predict that landscape's peaks with high accuracy). The authors did not model effectuation (or causation) as defined multi-dimensionally in that stream. They did not model planning. They did not model risk. They did not model uncertainty. So, the translation from theory to reality through the NK simulation was faulty, as neither the theory nor the reality was captured correctly in the model coded in their paper. ***Unaddressed Questions***: What does a stable landscape have to do with risk? What does an uncorrelated change of landscape have to do with uncertainty when that change relies on a random draw from a uniform distribution? Why aren't the real intermediate benefits associated with planning (e.g., improved accuracy and efficiency of future actions) and the real intermediate costs associated with flexibility (e.g., retraining expenses, and penalties for being caught under-capacity) captured in the (intendedly realistic) simulated testing of the alternative logics?

Table 2 (continued)

Reality-to-Theory Mediation Example using NK Simulation	Theory-to-Reality Mediation Example using NK Simulation
The patterns generated by their two-step process model were *ex ante* predictable, making the actual simulation and its description redundant. The patterns included: (1) continued but declining improvements in efficiency over time—that is what evolution, in general, promises and that is an artifact of an NK model; (2) industry output increasing at a decreasing rate—with $dq/dc < 0$—but this must happen as c decreasing implies q increasing, that following from (1) and so is another artifact of the NK model when feeding the Cournot model; (3) prices steadily declining at a decreasing rate—this follows from (2) and a downward-sloping inverse-demand curve; (4) an industry participation pattern of rapid entry followed by mass exit, leading to a shakeout and a stable number of competitors—with that stable competition resulting from the imposed constraint on being profitable to enter and from the way entrants are seeded, with the rapid over-entry being due to initial inefficiencies and homogeneous search skills and random initial assignments, and with exit due to stable demand and an imposed relative profitability condition [all following from evolutionary processes that allow quick entry and exit and limited capacity]; and, (5) these patterns are solely related to the interconnectedness of the *technological* solution—to the *K* in the NK, but this is *not* necessarily true, because the patterns are connected to the complexity of landscape, not to K itself. This is an issue, because the landscape can be affected by K, but also by N (when K > 1), by the allowable levels of elements in N, and by the forms of the K-functions.	When there is no real post-publication correction process in the journals that publish such NK simulation-based theory-testing studies, even when editors are made aware of the problems, what can be done to correct any misleading results? If such journals do not provide a dialogue outlet, and would not publish a replication of such a study (given the issues deal with a problem that is not about the data-gathering but instead about accurate coding), what is to be done in a management field to improve the role of theory and our role as diligent scholars to correct issues when we find them?

effectuation are missing, such as the concept of affordable loss). In addition, what characterizes the various landscapes is *not* what is stated (e.g., risk is not captured by a static landscape but rather by a set of possible known outcome states and their known probabilities of occurrence). As such, the theory is *not* actually tested. Thus, any apparent support for it is spurious. In addition, we understand why this could have occurred. It is *very* challenging to code a multi-faceted theory or an informationally complex context (e.g., involving a particular type of uncertainty) when the basic toolkit of this method was *never* intended to capture either. Such limitations perhaps should have been more thoroughly discussed. Furthermore, what was *not* accurately captured in the model needed to be made more explicit, and any support from the model made more conditional.[6]

In the second example—a less recent paper on process innovation that has received mostly positive reaction in the NK model review pieces—Lenox et al. (2007) use the NK model to generate data patterns that mimic real observations of industry evolution. Those patterns are based on a then-proposed-as-new-to-the-literature set of driving forces, some of which are coded in an NK model. The relationships emerging among the

---

[6] That raises the potential for a new form of flawed study, one that cannot be traditionally challenged by a replication because the theory claimed to be tested is not actually tested. Editors struggle to handle such cases when there are not dialogue outlets provided in journals that help to identify and discuss such issues. When such papers with these kinds of flaws are not tagged and discussed post-publication then such studies are effectively condoned, cited and even repeated, potentially harming the research integrity of our fields.

forces are leveraged to build a new partial theory. Their NK simulation produces a coded output indicating a firm's per-period production cost level, which is then fed into a one-period Cournot-competitive industry model, to provide a measure of fitness for the NK simulation. The two-part process is then repeated so that the population evolves. This generates patterns over time of each firm's costs, of the market's prices, and of the number of participating firms. That complete data-generation exercise is then repeated for cost functions entailing differing levels of production factor-interdependency captured by the landscape-defining dimension *K*. The analysis of the array of patterns is described and arguments are made that the industry evolutions depicted emerged from a new set of drivers.

 Unfortunately, there are several issues with their application of the NK-based approach. First, the most appropriate form of the NK model (i.e., the NK*C* form—a form that directly accounts for rival firms affecting the focal firm's landscape) was *not* used. Furthermore, the choice not to use a NKC model was *not* actually justified by them (Ganco et al., 2020). Second, running the simulation was *not* even required, given the standard NK simulation evolution output pattern was known, as were the reaction functions in Cournot models to changes in variable costs and in the number of competing firms (and, especially where the assumption of the kind of cost shift possible was unrealistically restricted—Ganco & Hoetker, 2009). By simply stating all of the assumptions of the NK and the Cournot models and how they were linked, it would have been straightforward to logically, deductively explain and predict the outcome patterns produced in the simulation.[7] Third, there was a failure to review the literature available at that time that had already established the NK model as an alternative explanation for observed industry evolution patterns (e.g., Huygens et al., 2001; Martin & Mitchell, 1998). Inducing a proposed new theory from an NK simulation's data appeared forced in this case (because the best model was not applied and the reasons why the modelling assumptions didn't obviously and directly drive the outcomes were not provided). Furthermore, the interpretation of the simulation's data from the visual patterns produced scrubbed away the majority of variance in firm-level outcomes, such that the real-world factor interdependencies underlying co-evolution among suppliers, and between supply and demand, were largely lost. The point arising from this critique is that the NK model methodology can be applied *un*necessarily or improperly, and that can lead to questionable results or under-justified partial theorizing.[8]

Discussion and suggestions for improvement

This commentary complements and counter-balances the works by Csaszar (2018), Fioretti (2013) and others that have optimistically described how NK models can aid management, entrepreneurship and innovation research. It does so by describing the

---

[7] If the coded assumptions of the simulation predetermine the outcome then there is no need to run it (but when that is not so, as with *chaotic* systems, the analysis should precede the theorizing – Gleick, 2011).

[8] Lenox et al. (2006)'s conclusions were considered by Lenox et al. (2010) and then by Lee and Alnahedh (2016). Lenox et al.'s (2010) support of their own paper's results referred more to its empirical verification of the well-known U-shape pattern involved and not to their own NK modelling itself. This is despite the fact that Lee and Alnahedh (2016: p286) state that Lenox et al. (2010) did not sufficiently provide support even to the U-shape conclusion. That confusion aside, this set of papers also reinforces the concern that the NK methodology's circle of scholars is much too tight, as Lenox was the senior editor at the journal who guided and accepted the Lee and Alnahedh (2016) paper. Such a fact raises a conflict of interest issue that needlessly puts their support of the Lenox et al. (2006) paper in question.

limitations of those models and the downsides from their mis-applications. At a high level, our general and specific critiques of the use of the NK model as a standard mediation-type approach in theorizing point to the possibility that it may not be *the* answer for theory-testing or theory-building. In fact, the examples point to the likelihood that using the NK model alone may, in fact, be *in*appropriate and produce poor results. On reflection, it seems overly optimistic to have expected that a model based on the mechanics of biological evolution would accurately capture the strategic challenges that entrepreneurial managers face in their real and idiosyncratic organizations.

Fioretti (2013, p. 233) implies three basic guidelines for when the NK model is more appropriate: (i) when the structure of interactions between social actors matters; (ii) when overall organizational behaviors arising the bottom-up out of interacting actors matters; and, (iii) when out-of-equilibrium dynamics matter. We are more restrictive in our updated guidelines for when the NK model is more appropriate: (i) when the focus is on the difference between local and global searches; (ii) when the focus is on population-level (and not individual-level) outcomes that evolve only from indirect interactions among routinized firms/ actors; and, (iii) when the focus is on the differences in patterns of a forced evolution towards greater fitness induced by the variation of specified factors.

Fioretti (2013, p. 235-6) also suggests further guidelines for NK model application regarding *validation*, specifically: (i) for theory-testing—that the model be faithful to the original theory; and, (ii) for the imitation of observations—that the model be authenticated at both the individual and aggregate levels of behavior. In our critical analyses of the two examples, those guidelines were *not* adhered to. In the first example, the original theory was not captured correctly. As such, that paper's suggestion that it was properly tested was inappropriate, as was the suggestion that it was tested under specific conditions (given those conditions were not accurately captured either). In the second example, observational authentication was a concern, because the paper's overall simulation forced together two incomplete models at different levels of behavior. Specifically, the NK model does not capture firm-level production scale, and the one-shot Cournot game model does not capture aggregate-level dynamics. As such, the second guideline for validation appears missed for the NK part of the simulation. In terms of agent-based models, it would have been more legitimate to start with the application of the most suitable model (i.e., an NKC variant), or at least one where the agents evolve *and* compete together on the same landscape. Having evolution occur on one landscape that involved low-rationality decision-makers while competing on a different one that involved high-rationality decision-makers was problematic for validation in that second example.

Those guidelines aside, we do understand that modelling involves a paradoxical challenge of balance—*capturing reality while abstracting away from it.* At their best, models focus on a few key factors for a specific research question to provide new insights and generalization. At their worst, models misinform and lead to worse decisions. We believe, however, for *new* modelling methods, such as NK simulation, that there is a greater onus on researchers in terms of proving that such balancing is being pursued properly. Without that onus, there are two big dangers: the first is over-extension of this method, perhaps akin to where someone with a hammer sees too many things as nails. The application of this new method calls for restraint and careful choices. The second is that alternative models—those written specifically

for managerial or entrepreneurial or innovation problems—will be crowded out by applications of this method that are under-modified away from contexts that are not biological (Baumann, 2015). That being said, we do recognize that some recent NK modelling has improved to address some of its past limitations. For example, Gavetti et al. (2017) model allows landscape shaping by firms; but such work appears more the exception than the rule so far. Thus, at this stage, we believe that this method remains a better complement to traditional methods than a substitute.

Conclusions

We have explored the limitations and critiqued recent example applications of the NK model methodology to justify our conclusion. Our conclusion is that, as a stand-alone method, the NK model lacks evidentiary substance, but it remains an effective supplement to the more traditional methods of empirical analysis and mathematical–logical analysis. We hope that our analysis of the generic NK model has provided a balance to the mostly positive and uncritical descriptions of what that method involves, so that audiences who are not experienced with coding it can better understand its limitations and the premises upon which it is based. We also hope that our commentary will help those advocating that method to update and improve its minimal model specifications in the future. We hope that the detailed examination of two example applications of the method highlights important concerns, and that that leads to more careful use of the method and more scrupulous pre-publication reviewing of such research. The conclusion that this new—arguably *third* main evidentiary method of research—has severe potential downsides (that have not previously been fully listed and exemplified) is worth repeating because of its relevance to our business fields. The phenomena we often study are not always easy to gather data for and, so, the attraction of using the NK model methodology to provide pseudo-empirical results may be high. This commentary provides a way to assess that option, a caution for what issues may arise, and some advice about what modifications to the base model should be considered (i.e., where the onus is on the researchers to prove the method is both necessary and suitable to their specific application). We hope that this commentary leads to clearer appreciations and uses of all newer pseudo-data-generating methods in the future, and better understandings of our entrepreneurial and innovative phenomena of interest.

Declarations

References

Adner, R., Csaszar, F. A., & Zemsky, P. B. (2014). Positioning on a multi-attribute landscape. *Management Science, 60*(11), 2794–2815.

Albert, D., & Siggelkow, N. (2022). Architectural search and innovation. *Organization Science, 33*(1), 275–292.

Baumann, O. (2015). Models of complex adaptive systems in strategy and organization research. *Mind & Society, 14*(2), 169–183.

Baumann, O., Schmidt, J., & Stieglitz, N. (2019). Effective search in rugged performance landscapes: A review and outlook. *Journal of Management, 45*(1), 285–318.

Conlisk, J. (1996). Why bounded rationality? *Journal of Economic Literature, 34*(2), 669–700.

Csaszar, F. A. (2018). A note on how NK landscapes work. *Journal of Organization Design, 7*(1), 1–6.

Csaszar, F. A., & Levinthal, D. A. (2016). Mental representation and the discovery of new strategies. *Strategic Management Journal, 37*(10), 2031–2049.

Csaszar, F. A., & Siggelkow, N. (2010). How much to copy? Determinants of effective imitation breadth. *Organization Science, 21*(3), 661–676.

Fioretti, G. (2013). Agent-based simulation models in organization science. *Organizational Research Methods, 16*(2), 227–242.

Ganco, M. & Hoetker, G. (2009). NK modelling methodology in the strategy literature: Bounded search on a rugged landscape. In: *Research methodology in strategy and management*. Emerald Group Publishing Limited. https://doi.org/10.1108/S1479-8387(2009)0000005010

Ganco, M. (2017). NK model as a representation of innovative search. *Research Policy, 46*(10), 1783–1800.

Ganco, M., Kapoor, R., & Lee, G. K. (2020). From rugged landscapes to rugged ecosystems: Structure of interdependencies and firms' innovative search. *Academy of Management Review, 45*(3), 646–674.

Gavetti, G., Helfat, C. E., & Marengo, L. (2017). Searching, shaping, and the quest for superior performance. *Strategy Science, 2*(3), 194–209.

Gavetti, G., & Levinthal, D. A. (2000). Looking forward and looking backward: Cognitive and experiential search. *Administrative Science Quarterly, 45*(1), 113–137.

Gleick, J. (2011). *Chaos: Making a New Science* (Enhanced Edition). Open Road Media.

Hannan, M. T., & Freeman, J. (1977). The population ecology of organizations. *American Journal of Sociology, 82*(5), 929–964.

Huygens, M., Van Den Bosch, F. A., Volberda, H. W., & Baden-Fuller, C. (2001). Co-evolution of firm capabilities and industry competition: Investigating the music industry, 1877–1997. *Organization Studies, 22*(6), 971–1011.

Jain, A., & Kogut, B. (2014). Memory and organizational evolvability in a neutral landscape. *Organization Science, 25*(2), 479–493.

Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge University Press.

Kauffman, S., & Weinberger, E. (1989). The NK Model of rugged fitness landscapes and its application to the maturation of the immune response. *Journal of Theoretical Biology, 141*(2), 211–245.

Lee, G. K., & Alnahedh, M. A. (2016). Industries' potential for interdependency and profitability: A panel of 135 industries, 1988–1996. *Strategy Science, 1*(4), 285–308.

Lenox, M. J., Rockart, S. F., & Lewin, A. Y. (2007). Interdependency, competition, and industry dynamics. *Management Science, 53*, 599–615.

Lenox, M. J., Rockart, S. F., & Lewin, A. Y. (2010). Does interdependency affect firm and industry profitability? An Empirical Test. *Strategic Management Journal, 31*(2), 121–139.

Levinthal, D. A. (1997). Adaptation on rugged landscapes. *Management Science, 43*(7), 934–950.

Li, C., & Csaszar, F. A. (2019). Government as landscape designer: A behavioral view of industrial policy. *Strategy Science, 4*(3), 175–192.

Martin, X., & Mitchell, W. (1998). The influence of local search and performance heuristics on new design introduction in a new product market. *Research Policy, 26*(7–8), 753–771.

McGrath, J. E. (1981). Dilemmatics: The study of research choice and dilemmas. *American Behavioral Scientist, 25*(2), 179–210.

Nelson, R., & Winter, S. (1982). *An evolutionary theory of economic change*. Belkap Press.

Platt, J. R. (1964). Strong inference. *Science, 146*(3642), 347–353.

Posen, H. E., & Martignoni, D. (2018). Revisiting the imitation assumption: Why imitation may increase, rather than decrease, performance heterogeneity. *Strategic Management Journal, 39*(5), 1350–1369.

Puranam, P., Stieglitz, N., Osman, M., & Pillutla, M. M. (2015). Modelling bounded rationality in organizations: Progress and prospects. *Academy of Management Annals, 9*(1), 337–392.

Rahmandad, H. (2019). Interdependence, complementarity, and ruggedness of performance landscapes. *Strategy Science, 4*(3), 234–249.

Rivkin, J. W., & Siggelkow, N. (2003). Balancing search and stability: Interdependencies among elements of organizational design. *Management Science, 49*(3), 290–311.

Sarasvathy, S. D. (2001). Causation and effectuation: Toward a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review, 26*(2), 243–263.

Wall, F. (2016). Agent-based modelling in managerial science: An illustrative survey and study. *Review of Managerial Science, 10*(1), 135–193.

Welter, C., & Kim, S. (2018). Effectuation under risk and uncertainty: A simulation model. *Journal of Business Venturing, 33*(1), 100–116.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.