DISCUSSION PAPER SERIES

DP17664
(v. 2)

THEORY-DRIVEN STRATEGIC MANAGEMENT DECISIONS

Arnaldo Camuffo, Alfonso Gambardella and Andrea Pignataro

ORGANIZATIONAL ECONOMICS

CEPR

ISSN 0265-8003

THEORY-DRIVEN STRATEGIC MANAGEMENT DECISIONS

Arnaldo Camuffo, Alfonso Gambardella and Andrea Pignataro

Discussion Paper DP17664
First Published 13 November 2022
This Revision 14 February 2023

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

THEORY-DRIVEN STRATEGIC MANAGEMENT DECISIONS

Abstract

This paper studies managerial decisions under uncertainty for which past data are not available. It shows that choosing a good theory and state space might make a larger difference in performance than solving optimally the ``wrong'' decision problem. Therefore, developing and testing alternative theories can be a key source of competitive advantage. The paper also shows that decision makers should experiment with more uncertain theories, because testing the "boundary" models of these theories elicit more informative signals that generates larger belief updating. The paper provides microfoundations to the theory-based view of the firm and explains why performance heterogeneity and competitive advantage is created by exploring theories with higher variance.

JEL Classification: L21, L26, M13, M21

Keywords: Decision problem

Arnaldo Camuffo - arnaldo.camuffo@unibocconi.it
Department of Management & Technology, Bocconi University

Alfonso Gambardella - alfonso.gambardella@unibocconi.it
Department of Management & Technology, Bocconi University and CEPR

Andrea Pignataro - ap@iongroup.com
ION Group

1 Introduction

This paper develops a framework that applies theory-based decision making under uncertainty to strategic management decisions in which executives cannot rely on past data. These are central and most often impactful strategic decisions in firms. They are "low-frequency high-impact" decisions in the sense that they are not ordinary decisions, but foundational decisions, such us investments in new markets and technologies (innovation); the choice of a firm's capital structure, ownership and governance; M&As; the recruiting and selection of CEOs, top management team members or other key human capital. The more innovative are these decisions, the less they can rely on past data or insights from analogous situations or conditions. For these decisions, before choosing actions executives need to develop theories to identify the states that define their problems and learn through experiments which theory to choose.

Executives often lack established practices or structured frameworks to make these decisions. We address this gap by proposing that they could ground strategic management decisions on theories and experiments, similar to the structured and disciplined approach used by scientists. This hinges on a combination of intuition, imagination, creativity, logic and beliefs about the key elements of the problem and their causal links, and a rigorous scientific approach for moving from antecedent to consequences and for testing these theories (Felin & Zenger, 2017; Camuffo et al., 2020; Zellweger & Zenger, 2021; Ehrig & Schmidt, 2022). Furthermore, we contend that problem framing should be the central aspect of the strategic decision-making process as it plays a crucial role in defining future state spaces, shaping the choices among alternative actions and ultimately impacting the performance of the firm (Nickerson & Zenger, 2004; Nickerson & Argyres, 2018).

We recognize that "low-frequency high-impact" decisions are characterized by "Knightian" uncertainty and we build on insights from three distinct areas of research in the strategy literature. First, decision-makers define decision problems "subjectively" and with "agency" so that they themselves represent the ultimate "un-caused cause" of uncertainty generation and reduction (Alvarez & Porac, 2020; Rindova & Courtney, 2020; Packard & Clark, 2020). Second, mental representations are critical for envisioning strategies and strategic problems (Csazar & Levinthal, 2016; Csazar, 2018). Third, strategic decision makers learn to make choices through experimentation (Gans et al., 2019; Gans, 2022). Moreover, the burgeoning literature on the economics of management has shown that managerial practices have important effects on firm productivity (Bertrand & Schoar, 2003; Bloom & Van Reenen, 2007; Bloom et al., 2013; Gosnell et al., 2020; Bandiera et al., 2020). When dealing with uncertainty, this literature has focused on prediction and the ability to use past data to foresee future performance (Altig et al., 2022), but has not addressed how to make decisions in non-ergodic contexts (Alvarez & Porac, 2020).

We also build upon new contributions in decision science, which has advanced significantly the development of theories that allow for structured and informed decision-making under uncertainty (Klibanoff et al., 2005; Denti & Pomatto, 2022; Cerreia-Vioglio et al., 2022; Karni, 2022). However, these models tend to be highly theoretical and distant from the practical reality of decision-making in firms. Furthermore, they do not provide applications or prescriptions for strategic management.

This paper shows how and why, in the absence of readily available decision problems or past data, executives should develop alternative theories, experiment with them, and choose the most promising and valuable one. We model this process as a Bayesian approach, in which executives form, test, and update beliefs about theories. A key insight is that executives should experiment with more uncertain beliefs or theories that exhibit higher variance of potential outcomes, even when they show lower expected returns or more generally seem less plausible, because they learn more. We show that seeking confirmation of most likely beliefs or theories is instead sub-optimal. The framework also provides a practical guidance on how executives should explore options when decision problems are undefined and past data is not available.

Theory-based decision making under uncertainty can be particularly powerful in the digital economy. As data and algorithms improve the ability to make decisions when the state space is given and data are plentiful, the source of value creation and competitive advantage shifts from choosing the best action for known decision problems to generating and choosing novel decision problems (new state spaces) where there is little or no data (Choi & Levinthal, 2022). In the digital economy, outperforming competitors by being better at solving known decision problems will be harder. Envisioning new, theory-driven state spaces can have a greater impact on value creation and performance (Baer et al., 2013; Nickerson & Argyres, 2018) as it reduces the risk of solving the wrong strategic problem (Mitroff & Silvers, 2010) and potentially provides a hardly imitable source of competitive advantage.

Section 2 provides a motivating example that illustrates the concepts and principles of the framework which is presented in Section 3. In Section 4, the framework's versatility is demonstrated by applying it to a different firm example, PayPal. In Section 5, we relate to existing strategy research streams, highlighting potential developments. Section 6 concludes by summarizing key takeaways, discussing the implications, and identifying opportunities for future research.

2 A Motivating Example

Founded by Leonardo Del Vecchio in 1961, EssilorLuxottica is today the world leader in the design, manufacturing and distribution of ophthalmic lenses, frames and sunglasses (Camuffo, 2003). Since the 1980s Luxottica rested on a successful business model that targeted individuals with eyesight defects. This represented a large, steadily increasing, global market. Spectacles correct eyesight defects, and across the world opticians are the key actors in the market as they deploy ophthalmologists' prescriptions and assemble lenses and eye frames.

Del Vecchio knew that by controlling opticians, he could control the market. The larger the market, the larger production volumes. The larger production volumes, the lower unit production costs because of learning curves and economies of scale. The lower the production cost, the lower the prices that can be charged. Given product quality, lower prices allowed further market penetration and market share.

However, in the late 1980s Del Vecchio observed that people with aesthetic needs represented an emerging and potentially large and global market, and that eyeglasses could become a design accessory that complements one's personal lifestyle. He thought that in this case the final customer is the key actor in the market since eyewear reflects personal style, identity and image. Designing, producing and marketing stylish eyeglasses, together with access to designer brands and retail, allow direct contact, profiling and control of the final customer. Customer control can generate high margins (premium prices) and global supply chain efficiency (low cost). This leads to business growth and profitability, and then to further investment for growth.

This eyewear theory was novel, unexplored, and largely unknown when Del Vecchio started considering it. Del Vecchio envisioned it as a future state space comprising multiple states to which he assigned subjective probabilities. These probabilities were subjective because Del Vecchio had no or few past data, information, or experience about the future states.

Once Del Vecchio articulated his theory, he conducted experiments to learn about it. For example, he noted that, since 1970 Optyl, a small Austrian producer, was thriving thanks to a licensing agreement with Christian Dior. It was an early and isolated case, but strongly signalled that eyeglasses and fashion could be coupled in meaningful and valuable ways. Similarly, Del Vecchio turned his attention to his competitor Safilo not to learn about its ordinary actions, but intrigued by a somewhat minor action: the 1984 acquisition of Optifashion, a small Italian producer who pioneered the idea to connect eyeglasses and fashion becoming a portfolio of licensing agreements with fashion brands like Missoni, Laura Biagiotti, Ferrè and Gucci. This provided a signal about the fact that eyewear was becoming increasingly likely and potentially valuable.

Del Vecchio conducted other experiments. For example, he entered his first licensing agreement with Armani in 1988, and carried out a few acquisitions of small sunglasses producers (Briko and Persol). These experiments gradually updated his beliefs about the potential of the eyewear theory. Since then Luxottica's strategies changed significantly into accelerated external growth through licensing agreements with iconic fashion brands such as Armani, Bulgari, Chanel, Dolce & Gabbana, Prada, and Versace; the acquisition of large retail chains such as Lenscrafters and Sunglass Hut; and the acquisition of iconic brands such as Ray Ban and, later, Oakley. Today Luxottica is the undisputed world leader in the luxury eyewear market that Del Vecchio originally envisioned and explored.

3 Building and testing theories

3.1 Background

In a classical decision problem, decision-makers do not know which state will realize among a set of potential future states of the world and evaluate alternative actions under this set of possible states. Since decision-makers know the set of possible states and their probabilities, they compute the expected values of the different actions, and pick the action with the highest expected value.

In strategic management, decision makers rarely know states and probabilities objectively. In some occasions they can refer to decision problems they have solved in the past to map actions onto consequences so that they consider states as known. In that case, they act as if they are faced with state uncertainty (Marinacci, 2015) and can use past data, information, or experience to predict the probability of occurrence of the states.

For example, Luxottica's long-standing experience with spectacles for eyesight defect correction implied that, when the company started exporting to other countries, it knew that the key was to understand whether they could control opticians. Moreover, based on the experience with the earlier countries in which it operated, Luxottica could get a good estimate of the probability about whether it was able to control a new national market, and then trigger the chain of factors – from control of opticians to large scale volumes, lower production costs, lower prices, high demand – that led to optimal decisions.

In other strategic decisions, however, both states and probabilities are unknown. States themselves are realizations of underlying random variables that are part of a generative mechanism that represents some natural or social phenomenon, or decision makers' hypothetical acts deliberately geared toward generating potentially valuable states. Decision makers confront uncertainty through the lens of the models they build, but may not know the mechanism underlying the phenomenon they model (ambiguity or model uncertainty, Marinacci, 2015). Furthermore, they might acknowledge the approximate nature of the models under consideration, view them as simplifications, and hence, as mis-specified (Hansen & Marinacci, 2016).

In the case of strategic management decisions characterized by unknown states and probabilities, decision-makers seek to define models on future, potentially valuable, state spaces that are novel and untapped by other firms, on which they can therefore build a competitive advantage. They resort to a "constructivist" approach (Alvarez & Porac, 2020) in which they acknowledge "fundamental" uncertainty, see it as not necessarily aversive, and embrace it instead of trying to avoid or mitigate it (Griffin & Grote , 2020). Instead of adopting a predictive approach to ill-defined decision problems, adapting to contingencies as they occur, or resorting to non-rational heuristics, decision-makers proactively build their decision problems envisioning future states (Arikan et al., 2020).

For example, when Del Vecchio envisioned the new eyewear market, he could not coast on existing states, nor had any specific data to build probability distributions. He logically built the future state space and developed beliefs about it.

3.2 Attributes, Causal Links, and Beliefs

When decision makers can neither coast on clearly analogous decision problems nor rely on past data, they identify *attributes* which are elements of the future state space set that have uncertain realizations and that decision makers consider as key for the strategic decision problem they face. When Del Vecchio started to imagine the new eyewear market, he focused on two attributes, whether it could successfully design, produce and market fashion eyeglasses, and whether there will be a significant market, i.e. high demand for fashion glasses. Let these two attributes and their spaces be

$$X_f = \{Y, N\} \qquad X_d = \{H, L\}$$

where subscripts f and d stand for "fashion" and "demand", Y and N for *yes* and *no*, and H and L for *high* and *low* demand. A simple way of thinking about attributes is that they are random variables whose realizations are answers to questions. In our example, the questions are: "will Luxottica be able to design, produce and market fashion eyeglasses?" and "what will demand for high fashion glasses be?".

Decision makers then generate a series of *causal links* moving from antecedent attributes to consequent attributes that can be represented as Bayesian networks (Pearl, 2009). The end state of the network is the state of interest for the decision maker. We call the selection of attributes and causal links framing decisions. In Luxottica's case, Del Vecchio was interested in understanding whether there was going to be a new market for eyewear with high demand. The other attributes represent elements of the future state space that decision makers believe can casually generate the state of interest (i.e. that are causally linked to the end attribute). In Luxottica's simple case, the connection is

$$X_f \longrightarrow X_d$$

The attributes and causal links identified by decision makers are the generative mechanism of the future state space and of its probability distribution. They can be grounded on regularities in empirical observations or other rules, principles and laws distilled by decision makers (Karni, 2022). In our case, the two attributes define a state space comprising four possible states: $(H, Y), (H, N), (L, Y), (L, N)$.

Decision makers are aware that this space is subjective. They conceptualize it by explicitly and deliberately selecting certain attributes and identifying the causal relationships between them and by excluding attributes or causal links that they believe will not have a significant impact on the end attribute of interest (i.e., in our example, the belief about whether eyewear demand will be high or low). They are aware that that the state space they have envisioned might not realize (their theory can be wrong), and that any decision they make (for example investing in product design or marketing) would be conditional on the assumption that their model specification is correct.

In the Luxottica's case, the four probability distributions of the four states are

$$
\begin{aligned}
&P(H, Y) = P(H \mid Y)P(Y) \\
&P(H, N) = P(H \mid N)\,(1 - P(Y)) \\
&P(L, Y) = (1 - P(H \mid Y))\,P(Y) \\
&P(L, N) = (1 - P(H \mid Y))\,(1 - P(Y))
\end{aligned}
$$

Decision makers are interested in the probability distribution of the relevant subset of states for their decision, in Luxottica's case $P(H)$. This probability distribution is the sum of two paths that lead to high demand: $P(H, Y) = P(H \mid Y)P(Y)$ and $P(H, N) = P(H \mid N)\,(1 - P(Y))$. Therefore, the probability distribution that Luxottica is interested in is

$$
P(H) = P(H, Y) + P(H, N) = P(H \mid Y)P(Y) + P(H \mid N)\,(1 - P(Y)) \tag{1}
$$

Luxottica's theory is that the ability to design, produce, and market fashion eyeglasses will generate high demand for fashion glasses – that is, Del Vecchio's theory sets $P(H \mid Y) > P(H \mid N)$.

The term $P(Y)$, instead, is an "un-caused cause" at the top of the causal structure. As we will see, these un-caused causes play an important role. There are three types of un-caused causes. First, decision makers can conceptually commit to a feasible, hypothetical action that represents the antecedent. In our example, Luxottica thinks that it is possible to design, produce and market fashion eyeglasses with probability $P(Y)$. Second, decision makers make an assumption about a future, exogenously determined future state. In this case, the theory is conditional on this assumption. These attributes correspond to elements of the strategic problem held as true or characterized by strong beliefs. Third, decision makers have another theory, nested into the one under consideration, that explains $P(Y)$. In this case, the attribute encapsulates a micro-theory that justifies $P(Y)$.

3.3 Models and Theories

To illustrate our framework, we use a specific form of the probability distribution. Suppose that the probabilities of the four states that represent the state space defined in the previous section are jointly distributed as Dirichlet with parameters $n_{IJ} > 0$, $I = H, L$; $J = Y, N$:

$$\propto P(H,Y)^{n_{HY}-1} \cdot P(H,N)^{n_{HN}-1} \cdot P(L,Y)^{n_{LY}-1} \cdot P(L,N)^{n_{LN}-1} \tag{2}$$

In a Dirichlet distribution

$$\mathbb{E}P(I,J) = \frac{n_{IJ}}{n_{HY} + n_{HN} + n_{LY} + n_{LN}}$$

which we re-parameterize as

$$\begin{aligned} \mathbb{E}P(H,Y) &= \theta_{HY}\theta_Y & \qquad \mathbb{E}P(H,N) &= \theta_{HN}\left(1-\theta_Y\right) \\ \mathbb{E}P(L,Y) &= \left(1-\theta_{HY}\right)\theta_Y & \qquad \mathbb{E}P(L,N) &= \left(1-\theta_{HN}\right)\left(1-\theta_Y\right) \end{aligned}$$

where θ_{HY}, θ_{HN}, and θ_Y are parameters that represent the decision makers' expected values of $P(H \mid Y)$, $P(H \mid N)$, and $P(Y)$.

Our framework builds on Cerreia-Vioglio et al. (2013). We assume that states in the state space are generated by a set of probability models $\theta \in \Delta(S)$ belonging to a finite set $\Theta \subseteq \Delta(S)$, where $\Delta(S)$ is the set of all possible probability distributions on the state space S that in our example is made of the four states $(H, Y), (H, N), (L, Y), (L, N)$.

We call probability *models* the single probability measures θ in the collection Θ (Marinacci, 2015, p.1037). In our example, $\theta = \{\theta_{HY}, \theta_{HN}, \theta_Y\}$ is a vector composed of three specific values of the three parameters $\theta_{HY}, \theta_{HN}, \theta_Y$. They represent a probability model because they identify a specific probability distribution (2) over the space of states (the four states above in our example) defined by the attributes of the problem identified by the decision makers.

Note that the state space is subjectively defined by the specific choice of attributes that decision makers believe are relevant for their problem. Different decision makers may pick different attributes, and thus different state spaces. More generally the lack of past data, or the inability to link the problem to patterns that are sufficiently regular to create "objective" states that all decision makers agree upon, makes the choice of the space idiosyncratic and subjective. For example, Del Vecchio, or other decision makers, could have chosen different attributes to address their decision about eyewear. In turn, this means that the probability distributions on this state space are subjective because they are defined over a state space that is subjective.

Differently from a model, a *theory* is the set Θ of probability models compatible with the causal links and beliefs of the decision maker. Thus, a theory restricts the choice of parameters from the set of all the values that these parameters can take to the subset of this space compatible with the theory. The distinction between models and theories is important. A model is a specific probability distribution over the space of states defined by the decision makers' attributes; a theory is a family of probability distributions on this space of states associated with each parametric models θ that belong to the set Θ. Because Θ is a subset of all potential realizations of these parameters, the family of probability distributions identified by the theory is a subset of all possible probability distributions on the state space. For example, Luxottica's theory is that the ability to design, manufacture and market fashion glasses raises the probability of high demand. Thus, θ_{HN} and θ_Y can take any value, but θ_{HY} must be such that $\theta_{HY} > \theta_{HN}$.[1] The parameters θ are themselves random variables.

Let $\mu(\theta \mid \Theta)$ be the joint probability distribution (likelihood) of the parameter set $\theta = \{\theta_{HN}, \theta_{HY}, \theta_Y\}$ under Luxottica's theory $\Theta = \{\theta_{HN}, \theta_{HY}, \theta_Y : \theta_{HY} > \theta_{HN}\}$. Under this theory a model of (1) is

$$v(\theta) = \theta_{HY}\theta_Y + \theta_{HN}(1 - \theta_Y) = \theta_{HN} + \theta_Y(\theta_{HY} - \theta_{HN}) \tag{3}$$

where θ_{HN}, θ_{HY}, and θ_Y are specific values that can take any value constrained only by the condition $\theta_{HY} > \theta_{HN}$.

The *conditional expected value of the theory* is instead

$$\mathbb{E}\left(v(\theta) \mid \Theta\right) \equiv V_\Theta = \int_{\theta\in\Theta} v(\theta)\mu(\theta \mid \Theta)d\theta \tag{4}$$

As an example, suppose that the three parameters are independent and their probability

[1]Our definition is complementary to Karni's (2022) definition of theories as mappings of feasible acts onto consequences (state space). Karni's definition includes a stochastic component that accounts for factors not included in the theory that may play a role in determining the state space. We emphasize that the attributes and causal links posed by decision makers (their theories) reduce uncertainty by concentrating probabilities.

distributions are uniform between 0 and 1. The expected values of θ_{HN} and θ_Y are then equal to $\frac{1}{2}$. According to theory Θ, the expected value of θ_{HY} is, instead, conditional on $\theta_{HY} > \theta_{HN}$. With a uniform distribution, this expected value is $\frac{1+\theta_{HN}}{2} = \frac{3}{4}$, where the second equality stems from the fact that the expected value of θ_{HN} is $\frac{1}{2}$. Replacing these expected values in (3), and recalling that the underlying parameters are independent, the expected value of the theory, conditional on the theory to be true, is $V_\Theta = \frac{1}{2} + \frac{1}{2}\left(\frac{3}{4} - \frac{1}{2}\right) = \frac{5}{8}$.

Because expected value (4) is conditional on the theory to be true, decision makers have beliefs on whether the theory is true, which we represent by a probability $\omega \in (0,1)$. Clearly, the higher ω the more the decision makers are confident about their theory and vice versa.

However, we cannot think of a belief about a theory without a null hypothesis. This is a different configuration of the parameters against which decision makers compare their theory. In Luxottica's example, a natural null hypothesis is that $\theta_{HY} = \theta_{HN}$. This yields a different set of parameters $\tilde{\Theta} = \{\theta_{HN}, \theta_{HY}, \theta_Y : \theta_{HY} = \theta_{HN}\}$, where θ_{HN} and θ_Y can take any values, and $\theta_{HY} = \theta_{HN}$. It is easy to see from (3) that, under the null hypothesis, models will be $v(\theta) = \theta_Y$. Thus, different values of the parameter θ_Y will represent different models under the null hypothesis. The expected value of the theory under the null hypothesis is then be $V_{\tilde{\Theta}} = \displaystyle\int_{\theta\in\tilde{\Theta}} v(\theta)\mu(\theta \mid \tilde{\Theta})d\theta$

The unconditional *expected value of the theory* is then

$$V = \omega V_\Theta + (1-\omega)V_{\tilde{\Theta}} \tag{5}$$

3.4 Experiments

3.4.1 Experiment to Test Theories

Decision makers learn about their theories by conducting experiments. Experiments are deliberate attempts to collect evidence about a phenomenon to update the belief ω about a theory. Therefore, an experiment yields a new updated belief ω' about the theory that can be higher or smaller than ω depending on positive or negative update from the experiment.

An experiment generates observations about the phenomenon, and can be conceptual or real. Conceptual experiments involve using reasoning and hypothetical observations to update and refine their understanding, while real experiments involve collecting real data or observations.[2] Decision makers can produce these observations either from quantitative data analyzed using statistical tools or by drawing qualitative information from phenomena.

For example, in Luxottica's case, most likely Del Vecchio started with a series of conceptual experiments in which he explored different scenarios in his head to form an initial belief about his theory and about the opportunity to move into fashion glasses. Then, as discussed in Section 2, he started with a series of real experiments. He first observed early attempts to link fashion to glasses by competitors, and then conducted early acquisitions and alliances with fashion brand companies. They were all initial attempts to learn about the phenomenon by acquiring information before deciding whether to scale up.

Decision makers can run experiments on any subset of the parameters of their theory. They can be joint experiments, when they focus on more parameters at the same time, or they can be experiments on a specific parameter. In Luxottica's example, Del Vecchio

[2] For the purpose of our discussion, we do not distinguish between these different types of experiments. In what follows, we speak broadly about experiments, which can be any of these types of experiments.

collected information on the demand of fashion glasses. He tested whether producing high quality fashion glasses generated a higher probability of high demand than if the company was unable to produce high quality fashion glasses. This amounts to testing his belief about $\theta_{HY} > \theta_{HN}$ vis-à-vis the null hypothesis that $\theta_{HY} = \theta_{HN}$, or equivalently he aimed at updating to ω' his belief ω about his theory against the null hypothesis.

Note that in our framework, when they make framing decisions, decision makers do not use experiments to make predictions about the value of the parameters, but to test theories. In other words, they do not experiment to update the value of θ_{HY}, or of any other parameters, as the specific value of a parameter is not the focus of the theory. The focus of the theory is to test specific implications of the theory, for example whether a parameter such as θ_{HY} takes values greater than θ_{HN}.

3.4.2 Experiment against Alternative Theories

If we assume that decision makers use all the information they have to form their belief about their theory, then before running an experiment on the theory, their expected update of the belief is equal to the current belief – that is $\mathbb{E}\omega' = \omega$. Of course, after the experiment, when they observe what the experiment produces, they make a positive or negative update. But before the experiment, they weigh positive or negative updates in such a way that the expected future belief is still ω.

If this was not the case, they must have, before the experiment, information that put greater weight on favorable or unfavorable information about the theory. However, in this case, their current belief must also be different from ω and equal to what they expect to obtain after the experiment. In other words, any information before the experiment, will be incorporated in the current belief about the theory. In turn, this implies that, using (5), the expected update of V is V. Therefore, before running the experiment, if it is even minimally costly, it is not worth running it because it does not provide new information.

In this section, we show that experiments are valuable only if decision makers test a theory against an alternative theory. An alternative theory comprises a different state space characterized by different attributes or causal links. For example, Luxottica's alternative theory was about its existing successful business on spectacles for eyesight defect correction, rather than eyewear. Del Vecchio realized that eye surgery was picking up, and might have jeopardized its traditional business. He set attributes $X_s = \{Y, N\}$ and $X_g = \{H, L\}$ for whether the technology of eye surgery develops or not (X_s) and whether demand for traditional glasses is high or low (X_g), where Y and N stand for *yes* and *no* and H and L for *high* and *low*. The causal link is $X_s \rightarrow X_g$. Let parameters γ, instead of θ, represent expected values of the probability distribution of the underlying probabilities of demand for traditional glasses.

Del Vecchio's theory is that even if the technology of eye surgery develops, the business of standard glasses for eyesight defects would not be affected because eye surgery is more costly and inconvenient than wearing a pair of glasses. Thus, under this alternative theory, a model for the probability of high demand for standard glasses, that is $P(H)$, is

$$q(\gamma) = \gamma_{HY}\gamma_Y + \gamma_{HN}(1 - \gamma_Y) = \gamma_{HN} + \gamma_Y(\gamma_{HY} - \gamma_{HN}) \tag{6}$$

where we interpret the set of parameters $\gamma = \{\gamma_{HN}, \gamma_{HY}, \gamma_Y\}$ analogously to the θ in the previous section. Now the theory is that γ_{HN} and γ_Y can take any value, but $\gamma_{HY} = \gamma_{HN}$. The theory is then the set $\Gamma = \{\gamma_{HN}, \gamma_{HY}, \gamma_Y : \gamma_{HY} = \gamma_{HN}\}$. If $\phi(\gamma \mid \Gamma)$ is the likelihood of these parameters conditional on the theory to be true, the conditional expected value of the theory is $Q_\Gamma = \displaystyle\int_{\gamma \in \Gamma} q(\gamma)\phi(\gamma \mid \Gamma)d\gamma$, analogous to (4). Since it is reasonable to believe that the technological development of eye surgery cannot increase the probability of high demand for standard glasses, a natural null hypothesis for this theory is that $\gamma_{HY} < \gamma_{HN}$. Assuming beliefs $\xi \in (0,1)$ about the theory against its null hypothesis, the expected value of the alternative theory is $Q = \xi Q_\Gamma + (1-\xi) Q_{\tilde{\Gamma}}$, where $Q_{\tilde{\Gamma}}$ is the expected value conditional on the null hypothesis. For the same reasons discussed earlier, decision makers cannot test this theory against itself.

As in the case of ω', the updated belief ξ' is such that $\mathbb{E}\xi' = \xi$, that is the expected update before the experiment is equal to the current belief. Therefore, before running the experiment, the expected update of Q is equal to Q, which makes it not worth running the experiment if it is minimally costly.

The two theories, Θ and Γ, with expected values V and Q, then have to be tested against each other. In particular, decision makers could run an experiment on theory Θ, that is on the parametric restrictions implied by this set, against the expected value of the alternative theory Γ. This experiment updates the belief ω, and leaves the belief ξ unaltered. Conversely, an experiment on theory Γ tests the parametric restrictions of this set against the expected value of theory Θ. It updates belief ξ leaving ω unaltered. The following proposition sets a necessary condition for running these experiments.

Proposition 1. *A necessary condition to experiment with theory Θ against theory Γ is that $V_{\tilde{\Theta}} < Q < V_\Theta$. A necessary condition to experiment with theory Γ against theory Θ is that $Q_{\tilde{\Gamma}} < V < Q_\Gamma$.*

Proposition 1 states that, in order to learn through experimentation about the value of theories, the two theories tested against each other cannot have very distant expected values. The reason why the conditions of Proposition 1 are necessary is a direct consequence of the fact that, because before the experiment the expected update of V after the experiment is equal to V, and similarly for Q, decision makers can only test one theory against the other. However, to do so, there must be a marginal belief $\omega^* \in (0,1)$ such that $V^* \equiv \omega^* V_\Theta + (1-\omega^*)V_{\tilde{\Theta}} = Q$, which is possible only if $V_{\tilde{\Theta}} < Q < V_\Theta$. If Q falls outside of these boundaries, decision makers commit to V or Q without experimenting depending on whether $Q < V_{\tilde{\Theta}}$ or $Q > V_\Theta$.

If instead the necessary condition of Proposition 1 is satisfied, decision makers know that if they run an experiment and observe update $\omega' > \omega^*$, the update of V will be $V' \equiv \omega' V_\Theta + (1-\omega')V_{\tilde{\Theta}} > Q$. If, instead, they observe $\omega' < \omega^*$ they update V to $V' < Q$. Therefore, before running the experiment they know that if they observe $\omega' > \omega^*$, they will stick to theory Θ with its updated expected value V', otherwise they switch to theory Γ with expected value Q. Without the alternative theory, they would have to stay with option V' even if $\omega' < \omega^*$. In this case, the expected value of the experiment would be V, which makes it not worth running an experiment if it is even minimally costly. In contrast, the lower bound Q makes the expected value of the experiment higher than V.

The logic is symmetric for experiments on Q against V. Let $\xi^* \in (0,1)$ be the marginal belief such that $Q^* \equiv \xi^* Q_\Gamma + (1-\xi^*)Q_{\tilde{\Gamma}} = V$, which is possible only if $Q_{\tilde{\Gamma}} < V < Q_\Gamma$. Then, decision makers experiment on Q and stick to this theory if they observe $\xi' > \xi^*$, and switch to theory Θ and expected value V if they observe $\xi' < \xi^*$. The expected value of this experiment is higher than Q, and may be worth running even if costly.

Let $c > 0$ be a fixed cost of running an experiment, and $\Omega(\omega')$ and $\Xi(\xi')$ the cumulative probability distributions of the experiments on, respectively, Θ and Γ – that is, Ω and Ξ are the probabilities of generating updates smaller than their arguments ω' and ξ'. The following proposition establishes sufficient conditions for experimentation.

Proposition 2. *Given the necessary condition in Proposition 1, a sufficient condition for* *running an experiment on theory Θ against theory Γ is that*

$$V^E \equiv \int_{\omega^*}^{1} V' d\Omega(\omega') + Q\Omega(\omega^*) - c > max(V, Q) \tag{7}$$

and analogously for theory Γ against Θ, $Q^E \equiv \int_{\xi^*}^{1} Q' d\Xi(\xi') + V\Xi(\xi^*) - c > max(Q, V)$.

In this proposition, V^E is the net expected value of the experiment. The first term is the expected value of the theory for different levels of the update V' weighted by their probabilities when $\omega' > \omega^*$. When, instead, $\omega' < \omega^*$, which occurs with probability $\Omega(\omega^*)$, decision makers earn Q, which is the second term of V^E. When $V > Q$, decision makers consider whether to run the experiment against the current expected value of the theory V. If $V^E > V$, it is worth running it; otherwise, they do not run the experiment and commit to $V > Q$. If $Q > V$, they check V^E against their current best option Q. The logic is analogous for Q^E and the experiment on Γ.

3.4.3 Experimenting with More Uncertain Theories

Finally, we discuss the choice of which theory to experiment with. To streamline our discussion, we assume that decision makers have the resources to run only one experiment. With this simplification, we neglect that, when choosing one experiment, decision makers may also take into account other experiments that they can run after they observe the results of this experiment. A dynamic perspective would require a more elaborate framework and discussion without affecting the main points below.

Without loss of generality, we focus on the Θ-theory. We can rewrite condition (7) for running an experiment as

$$V^E \equiv V_\Theta - \Delta V \int_{\omega^*}^{1} \Omega(\omega') d\omega' - c > max(V, Q) \tag{8}$$

where we obtain this expression for V^E after integrating (7) by parts, and $\Delta V \equiv V_{\Theta} - V_{\tilde{\Theta}} > 0$. This yields the following proposition.

Proposition 3. *Decision makers are more likely to experiment with theories on which they hold more uncertain beliefs*

If decision makers are confident about their beliefs on the theory, experiments are not "surprising" in the sense that they will not generate updates of the belief distant from the currently held belief. Confidence in the initial beliefs stems from the fact that decision makers have much information and experience with the problem, possibly because of past data or expertise. The new information is then not going to weigh much compared to the information held before the experiment, and therefore it is not going to change the initial belief by much. The greater concentration in potential updates around the expected update ω of ω' is equivalent to second-order stochastic dominance in Ω, which raises the integral in (8) (Rothschild & Stiglitz, 1970). This makes it harder to satisfy condition (8). Conversely, a greater spread in the potential update of the original belief reduces the integral on the left-hand side. This makes it more likely that condition (8) is satisfied, and therefore that decision makers experiment with the theory.

This is an intriguing result. Theories with a higher wedge of potential updates from the experiment are theories on which decision makers hold less information, which makes the experiment more informative. They are real "low-frequency high-impact" decisions that do not rely much on past data and for which past information or analogies about known situations provide little guidance on what to expect. They are also less conventional theories because they cannot be extrapolated from observing the past, but have to be "invented" using new forward-looking perspectives.

Clearly, the mechanism of this process depends on the "benchmark" produced by the alternative theory Q that shelters decision makers from the downside of the experiment on the more uncertain theory. However, this simply highlights the importance of developing alternative theories to enable decision makers to test unconventional theories that they are less confident about. These theories are not plausible before running the experiment, but can become far more plausible after the experiment.

Our framework is compatible with testing theories that are less likely than the alternative theory. Condition (8) could be satisfied by tests of the Θ-theory when $V < Q$, and not just vice versa. This adds to the fact that not only can it be optimal to test more uncertain, but also less plausible theories. Overall, this captures well our emphasis on the idea that innovative decision makers test novel and unconventional theories.

We call experiments on less plausible theories *falsification experiments*. In this case, the company has a dominant theory. Most likely these are theories associated with what the company is currently doing. The company then sees a new opportunity with lower expected returns, but that could become more valuable after an experiment. The experiment can falsify the existing theory and highlight the new innovative opportunity.

We call experiments on more plausible theories *confirmation experiments*. In this case, external shocks may challenge the value of the current business. Firms develop and test theories to assess whether the value of the current business is actually higher than an alternative less valuable theory. The experiment could confirm that the current theory is more valuable, or it could suggest that it has become less valuable, and decision makers switch to the alternative theory.

In the Luxottica example, Del Vecchio had a pretty good understanding of his traditional business, which made him fairly confident about his relatively high belief about ξ. As discussed earlier, the probability distribution Ξ was pretty concentrated around its mean.

This made the experiment less informative in the sense that the wedge of future updates ξ' of the belief around ξ was not high. In turn, this raises the integral in (8), which lowers the expected value of the experiment Q^E. This made Q^E smaller than Q or V^E. In either case, an experiment on Γ is not worth running, and Del Vecchio did not test the impact of eye surgery. At the same time, eyewear was a largely novel and unconventional idea. Del Vecchio had little experience and information about the business in practice. He was less knowledgeable about it than his old business, held weaker beliefs, and was highly uncertain about ω'.

He then experimented with this theory. As discussed in Section 2, he looked for information to test it, such as observing its competitor Safilo's seemingly minor acquisition of Optifashion, a pioneering company in high-fashion glasses, or signing initial deals with fashion brand companies to experiment with the new market.

Luxottica ran a falsification experiment. The new theory was less valuable than the status quo, but had potential. Del Vecchio tested it and discovered the new opportunity. However, had eye surgery been a more uncertain theory, he could have tested whether it was going to pick up and disrupt his current business, running, instead, a confirmation experiment. In the former case, he would pick eyewear because it proved to be a better opportunity than his current business. In the latter case, he would have picked eyewear because his current business was becoming a worse opportunity.

4 PayPal

4.1 Background

We re-interpret the PayPal case showing that the founders developed alternative theories, experimented with the one with higher variance and eventually committed to it as the most plausible and promising.

Born in Ukraine in Kiev and moved to Chicago in 1991, Max Levchin began to take an interest in security and encryption technologies as a young computer scientist. With Luke Nosek and Scott Bannister, he developed an application on PalmPilot that improved upon existing password generation solutions. In 1998, he moved to Silicon Valley where he met Peter Thiel and together they founded Confinity Inc. with the belief that the growth of the internet and the digital economy would revolutionize financial transactions. The company developed software for data security, digital wallets, and peer-to-peer money transfers. By merging with X.com, an online bank founded by Elon Musk, they created PayPal, a secure digital payment system for peer-to-peer online transactions which was then adopted as company name due to the popularity of the product. The company in 2002 was acquired by eBay for $1.5 billion and continued to expand, as a subsidiary of eBay, its services and user base, becoming one of the most widely used online payment systems in the world. In 2015, PayPal was spun off from eBay as an independent company. As of 2021, PayPal has over 300 million active users and is one of the most widely used and trusted online payment systems in the world.

4.2 The Original ("Plausible") Theory

In the mid 1990s, Levchin's, Nosek's and Bannister's initial theory was focused on data security on handheld devices and comprised two "un-caused causes". The first was an assumption: handheld devices would spread out. Conditional on such diffusion, data security would become a widespread need for all handheld device users. The second was an action: the development of a new technology (an encryption software) that would address the problem of password protection and more generally of data security. The rise of data security needs on handheld devices and the availability of the new encryption technology would jointly generate a new market for data protection on handheld devices. Levchin's and partners' theory rested on the following causal structure.

$$\begin{tikzcd}
 & X_e \arrow[dr] & \\
X_h \arrow[dr] & & X_d \\
 & X_s \arrow[ur] &
\end{tikzcd}$$

In the previous section we worked with discrete attributes and a Dirichlet distribution, which sets dichotomous realizations of the attributes. Here we generalize the framework by working with continuous variables.

Let the relevant outcome of interest be the attribute $X_d = \{x_d\}$ where x_d is a continuous measure of the demand for handheld devices encryption software that depends on two attributes: $X_e = \{x_e\}$, a continuous index of the efficiency of the encryption technology that will be available, and $X_s = \{x_s\}$, which captures the extent to which data security will become a widespread need for all handheld device users. In turn, $X_s = \{x_s\}$, is determined by $X_h = \{x_h\}$, which captures the spread of handheld devices.

Levchin's and partners' theory and causal structure generates the following chain of probabilities or beliefs

$$p(x_d, x_e, x_s, x_h \mid \theta) = p(x_d \mid \theta_{des}, x_e, x_s)\ p(x_e \mid \theta_e)\ p(x_s \mid \theta_{sh}, x_h)\ p(x_h \mid \theta_h) \tag{9}$$

where $\theta = \{\theta_{des}, \theta_e, \theta_{sh}, \theta_h\}$ is the parameter set of the distributions.

In the Appendix we show that a causal structure such as (9) generates a sequence of expected values that in the end produces the following linear approximation of the expected value $v(\theta)$ of the state (x_d, x_e, x_s, x_h) conditional on the parameter set θ of the underlying probability distribution:

$$v(\theta) = \theta_{de}\theta_e + \theta_{ds}\theta_{sh}\theta_h \tag{10}$$

where:

- θ_{de} and θ_{ds} denote the link between demand and an efficient encryption technology or data security needs on handheld devices
- θ_{sh} denotes the link between the diffusion of handheld devices and data security needs about them
- θ_e captures the expected quality of the encryption technology
- θ_h captures the expected spread of handheld devices

The theory is the set Θ with θ_{de}, θ_{ds}, $\theta_{sh} > 0$, and relatively high values of θ_e and θ_s. The expected value of this theory is V_Θ, defined by (4), with a probability distribution $\mu(\theta \mid \Theta)$ of these parameters. Given a null hypothesis on the parameters and beliefs, (5) defines the unconditional expected value of this theory.

This theory looked highly plausible to Levchin and partners. In the mid 1990s, handheld device adoption was seen as a clear emerging trend. Today, with hindsight, we know they were soon replaced by smartphones. But at the time there was general consensus that handheld devices like PDAs would diffuse and develop. Levchin and partners shared this strong belief so that they expected $p(x_h \mid \theta_h)$ to be high. At the same time, they already had a good technology and were confident that it could be made to good use so that they expected $p(x_e \mid \theta_e)$ also to be high. The conditional probabilities associated to the causal links $p(x_d \mid \theta_{des}, x_e, x_s)$ and $p(x_s \mid \theta_{sh}, x_h)$ were also high because of the large amount of information existing on the handheld device market. In our framework language, this was a "low-variance" theory as the difference in the likelihoods of the two "causal chains"/models is small.

4.3 The "New" Theory

When in 1998 Levchin moved to Silicon Valley and met Thiel, they developed an alternative theory. Based on Thiel's experience as securities lawyer, derivatives trader and venture capitalist, they identified a new set of attributes and connected them through novel causal links, thus envisioning a new state space.

The assumption of their theory (denoted by the attribute $X_i = \{x_i\}$) was that the online commerce and P2P transactions would rise with the internet economy. This would generate the need for efficient and secure digital payments for consumers and businesses (denoted by the attribute $X_s = \{x_s\}$) as the use of checks and money orders via the U.S. Postal Service would be too slow and unreliable to support online commerce. At the same time, they thought that they could develop a technology (a platform with encryption) to host secure, fast and low-cost digital payments. We denote this attribute by $X_d = \{x_d\}$. The platform would enable consumers to pay merchants quickly and easily without sharing sensitive information but only disclosing their email addresses and hence reducing and preventing frauds. Their theory rested on the following causal structure.

$$\begin{tikzcd}
 & X_d \arrow[dr] & \\
X_i \arrow[dr] & & X_p \\
 & X_s \arrow[ur] &
\end{tikzcd}$$

The relevant outcome of interest is attribute $X_p = \{x_p\}$ which denotes the extent to which people will use digital payments for their online purchases. This depends on two attributes, which jointly and causally determine it: $X_d = \{x_d\}$, which measures the extent to which a secure and efficient digital payment platform will be available, and

$X_s = \{x_s\}$, which denotes whether consumers and merchants will need faster and secure digital payments to support online commerce and P2P transactions. In turn, $X_s = \{x_s\}$, is determined by $X_i = \{x_i\}$, which captures the extent to which e-commerce and the internet economy will rise. Their theory belief structure is

$$p(x_p, x_d, x_s, x_i \mid \theta) = p(x_p \mid \theta_{pds}, x_d, x_s)\ p(x_i \mid \theta_i)\ p(x_s \mid \theta_{si}, x_i)\ p(x_d \mid \theta_d)$$

As per previous section, we can consider for simplicity the linear approximation $v(\theta) = \theta_{pd}\theta_d + \theta_{ps}\theta_{si}\theta_i$, where

- θ_{pd} and θ_{ps} are the two elements of the vector θ_{pds} that represent the links between a market of digital payments for P2P online transactions and a secure and efficient digital payment platform or the need for fast and secure payments for online commerce
- θ_{si} denotes the link between the rise of the internet economy and the need for efficient and secure digital payments
- θ_d captures the efficiency and security of digital payment platform
- θ_i captures the rise of the internet economy

The theory is the set Θ with θ_{pd}, θ_{ps}, $\theta_{si} > 0$, and relatively high values of θ_d and θ_i. The conditional expected value of this theory is again analogous to V_Θ, defined by (4), with a probability distribution $\mu(\theta \mid \Theta)$ of these parameters, while (5) represents the unconditional expected value of the theory given a null hypothesis and beliefs.

This was a more complex theory, less plausible than their original theory because the attributes and logical links were more uncertain. The digital payment platform was harder to conceive and less likely to be feasible (lower $p(x_d \mid \theta_d)$). The causal links were not obvious. For example, will consumers really feel the need to speed up the online transactions? Why can't they continue to use traditional means of payments like checks for online transactions? This implies greater uncertainty about $p(x_p \mid \theta_{pds}, x_d, x_s)$. In the language of our framework, this theory had a larger variance, as the likelihoods of the "causal chains" and models were more dispersed.

Levchin, Thiel, Nosek and Musk explored this theory and in 1999 they ran an experiment to test it. The experiment consisted in the launch of their digital payment system on a specific platform, Power Seller on E-bay, comprising 20,000 customers, known for being conservative, difficult and demanding customers. In our framework, this corresponds to testing the parameters θ_{pd} and θ_{ps}. The experiment was falsificatory. The founders thought that, if the digital payment system worked for Power Seller, it would have worked everywhere else. The experiment showed that the digital payment system worked and, hence, updated the founders' beliefs so that the "new" theory became more plausible and valuable than the "original" one.

At that point, Levchin, Thiel, Nosek and Musk had no doubt that their theory of a digital payment platform was more plausible and promising than the original theory of providing encryption software for handheld devices. It was under this theory that Confinity Inc. became Paypal, got adopted by E-bay and became a worldwide diffused payment system (Jackson, 2004, Furr & Dyer, 2014; Soni, 2022).

5 Research Developments

Theory-based decision making under uncertainty, not only re-aligns strategic decision-making with its roots in decision science (Schlaifer & Raiffa, 1961; Ansoff, 1965; Mason, 1969; Mitroff & Betz, 1972) but also opens up opportunities for incorporating recent developments in the economics and management of decision making under uncertainty, experimentation and information acquisition.

Our framework incorporates the recent developments of management research as regards managerial decision-making under fundamental or “knightian” uncertainty. It incorporates the fact that, when decision problems are not defined (“unknown-unknowns”) and data are not available, decision-makers should define decision problems “subjectively” and with “agency” (Packard & Clark, 2020; Rindova & Courtney, 2020). Consequently, our framework views decision-makers as the ultimate “un-caused cause” of uncertainty generation and reduction. Furthermore, it is consistent with a “constructivist” approach to fundamental uncertainty (Alvarez & Porac, 2020). Decision makers acknowledge that uncertainty about the future is not necessarily aversive, and should be embraced instead of merely avoided or mitigated (Griffin & Grote , 2020).

To summarize, we offer a protocol to structure the deliberate, cognitive “generative process” which allows decision makers to “construct uncertainty” and, hence, decision problems (Arikan et al.; 2020). Instead of adopting predictive approaches to ill-defined decision problems, adapting to contingencies as they occur, or resort to non-rational heuristics, we posit that decision-makers should better frame decision problems envisioning promising and valuable state spaces through theories and experimenting with them to choose the one they believe most plausible. Our framework shares with the “problem-finding problem-solving” perspective (Nickerson & Zenger, 2004; Baer et al., 2013) the idea that the core of strategic decision-making is to find and solve new problems that generate knowledge.

We also bridge recent developments in management research on decision-making under condition of “knightian” uncertainty with the ongoing debate in decision science. The increasing relative importance of ”low-frequency/high impact” strategic decisions makes Ellsberg’s 1961 paradox more salient for executives, which might tend to sub-optimally focus on "high-frequency/low impact" strategic decisions – for which there are increasingly large amount of data available – and disregard more important "low-frequency/high impact" strategic decisions which are harder to frame. This potential distortion requires an explicit consideration of ambiguity and ambiguity aversion in our framework (Maccheroni et al., 2016).

Our framework incorporates uncertainty in management theories – as suggested in many recent articles (Alvarez & Porac, 2020), editorials (Alvarez et al., 2018), and special issues (Alvarez et al., 2020). It explicitly considers strategic decision making under uncertainty not only across given/known states, but also across unknown state spaces, i.e. across models and theories. At the same time, we acknowledge that our framework only partially accounts for priors about theories and model mis-specification. We need to understand better how executives form these priors and more generally doubt about their theories. Incorporating these elements would represent a critical development in strategic management decision making under uncertainty (Klibanoff et al., 2005; Cerreia-Vioglio et al., 2013 and 2022; Hansen & Marinacci, 2016; Hansen & Sargent, 2022).

While we offer a template for how executives should craft their theories in a disciplined and rigorous way, we do not explain why and how executives decide what to include and leave out of their theories (choice of attributes and causal links). To this aim, the conceptual tools of two approaches might be usefully incorporated in our framework. The first is "learning through noticing" (Hanna et al., 2014) – which analyzes how decision makers choose which input dimensions to attend and subsequently learn about from available data. The second is "rational inattention" – which analyzes how decision makers, characterized by memory/beliefs and limited attention resources, choose which information to attend to and which information to ignore (Maćkowiak et al., 2018).

Furthermore, our framework could incorporate whether and to what extent executives are aware of being unaware of other states they do not consider and how they could include this awareness in theory definition and the corresponding process of belief formation and testing. Reverse Bayesianism (Karni & Vierø, 2013) and decision making under growing awareness (Karni & Vierø, 2017; Dominiak & Tserenjigmid, 2022) might help to understand how executives should include awareness of unawareness into their strategic decision making processes and understand how this affects the beliefs and priors of the theories and models they envision.

By incorporating recent literature on experimentation (Zhong, 2022) and information acquisition (Frankel & Kamenica, 2019), we could bolster the idea that executives should experiment with theories they are less confident about ("methodic doubt") and that they should choose experiments whose outcomes are more potentially "surprising" (Che & Mierendorff, 2019).

Finally, we leave several open questions for future research. We do not provide a theory of how decision makers identify, select, and combine attributes and logical links to formulate and rank theories. Relatedly, it would be important to investigate the determinants of learning speed. How quickly strategists explore alternative theories can represent a source of competitive advantage. Our analysis assumes identity between decision makers and firms. Future research should investigate how firms develop routines, design organizations and build management systems that allow this framework to be effectively, consistently and efficiently deployed across complex organizations.

6 Conclusion

This paper argues that strategic management decision making involves not only choosing the best action within a given decision problem, but also choosing the best theory or state space. It suggests that choosing optimally theories and state spaces can have a greater impact on performance than solving a sub-optimal decision problem. It recommends that decision makers should experiment with more uncertain theories because testing these theories can provide more informative signals and lead to greater belief updating. The development of alternative theories should be structured and disciplined by selecting attributes, linking them causally, and assigning likelihoods to the resulting causal chains.

Our framework pulls together multiple streams of strategy research, unifies its language, and provides a common ground to boost its rigor and impact. Decision makers hardly use economic and management theories in practice because they focus on decisions about actions rather than on framing decisions that relies on theories. Scientific knowledge provides the basis to make strategists domains richer and allow them to better craft their theories.

References

Ackoff, R. L. (1974). *Redesigning the Future.* Wiley, New York.

Altig, D., Barrero, M.J., Bloom, N., Davis, S., Meyer, B. & Parker, N. (2022). Surveying Business Uncertainty. *Journal of Econometrics*, 231(1), 282-303.

Alvarez, S. A. & Porac, J. (2020). Imagination, Indeterminacy, and Managerial Choice at the Limit of Knowledge. *Academy of Management Review*, 45(4), 735-744.

Alvarez, S., Afuah, A., & Gibson, C. (2018). Editors' Comments: Should Management Theories Take Uncertainty Seriously?. *Academy of Management Review*, 43(2), 169-172.

Alvarez, S. A., Zander, U., Barney, J. B., & Afuah, A. (2020). Developing a Theory of the Firm for the 21st Century. *Academy of Management Review*, 45(4), 711-716.

Ansoff, H. I. (1965). *Corporate Strategy: An Analytic Approach to Business Policy for Growth and Expansion.* McGraw-Hill Companies, New York.

Arikan, A. M., Arikan, I., & Koparan, I. (2020). Creation opportunities: Entrepreneurial curiosity, generative cognition, and Knightian uncertainty. *Academy of Management Review*, 45(4), 808-824.

Baer, M., Dirks, K. T., & Nickerson, J. A. (2013). Microfoundations of Strategic Problem Formulation. *Strategic Management Journal*, 34(2), 197-214.

Bandiera, O., Hansen, S., Prat, A. & Sadun, R. (2020). CEO and Firm Performance. *Journal of Political Economy*, 128(4), 1325-1369.

Bertrand, M. & Schoar, A. (2003). Managing with Style: The Effect of Managers on Firm Policies. *The Quarterly Journal of Economics*, 118(4), 1169-1208.

Bloom, N. & Van Reenen, J. (2007). Measuring and Explaining Management Practices across Firms and Countries. *The Quarterly Journal of Economics*, 122(4), 1351-1408.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D. & Roberts, J. (2013). Does Management Matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1), 1-51.

Camuffo, A. (2003). Transforming Industrial Districts: Large Firms and Small Business Networks in the Italian Eyewear Industry. *Industry and Innovation*, 10(4), 377-401.

Camuffo, A., Cordova, A., Gambardella, A. & Spina, C. (2020). A Scientific Approach to Entrepreneurial Decision-Making: Evidence from a Randomized Control Trial. *Management Science*, 66(2), 564-586.

Carroll, G. R. & Sorensen, J. B. (2021) Making Great Strategy: Arguing for Organizational Advantage. Columbia University Press, New York.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M. & Montrucchio, L. (2013). Classical Subjective Expected Utility. *Proceedings of the National Academy of Sciences*, 110, 6754â€“6759.

Cerreia-Vioglio, S., Hansen, L. P., Maccherioni, F. & Marinacci, M. (2022). Making Decision under Model Misspecification. Unpublished manuscript.

Che, Y. K. & Mierendorff, K. (2019). Optimal Dynamic Allocation of Attention. *American Economic Review*, 109(8), 2993-3029.

Choi, J. & Levinthal, D. (2022). Wisdom in the Wild: Generalization and Adaptive Dynamics. *Organization Science*, forthcoming.

Csazar, F. A. (2018). What Makes a Decision Strategic? Strategic Representations. *Strategy Science*, 3(4), 606-619.

Csaszar, F. A. & Levinthal, D. A. (2016). Mental Representation and the Discovery of New Strategies. *Strategic Management Journal*, 37(10), 2031-2049.

Denti, T. & Pomatto, L. (2022). Model and Predictive Uncertainty: A Foundation for Smooth Ambiguity Preferences. *Econometrica*, 90(2), 551-584.

Dominiak, A. & Tserenjigmid, G. (2022). Ambiguity under Growing Awareness. *Journal of Economic Theory*, 199, 105256.

Drucker, P. (1954). *The Practice of Management*. Harper & Row, New York.

Ehrig, T. & Schmidt, J. (2022). Theory-Based Learning and Experimentation: How Strategists Can Systematically Generate Knowledge at the Edge Between the Known and the Unknown. *Strategic Management Journal*, forthcoming.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4), 643-669.

Felin, T. & Zenger, T. R.. (2017). The Theory-Based View: Economic Actors as Theorists. *Strategy Science*, 2(4), 258-271.

Frankel, A. & Kamenica, E. (2019). Quantifying Information and Uncertainty. *American Economic Review*, 109(10), 3650-80.

Gans, J. S. (2022). Experimental Choice and Disruptive Technologies. *Management Science*, forthcoming.

Gans, J. S., Stern, S. & Wu, J. (2019). Foundations of Entrepreneurial Strategy. *Strategic Management Journal*, 40(5), 736-756.

Gosnell, G., List, J. & Metcalfe, R. (2020). The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains. *Journal of Political Economy*, 128(4), 1195-1233.

Griffin, M. A. & Grote, G. (2020). When is More Uncertainty Better? A Model of Uncertainty Regulation and Effectiveness. *Academy of Management Review*, 45(4), 745-765.

Hanna, R., Mullainathan, S. & Schwartzstein, J. (2014). Learning Through Noticing: Theory and Evidence from a Field Experiment. *The Quarterly Journal of Economics*, 129(3), 1311-1353.

Hansen, L. P. & Marinacci, M. (2016). Ambiguity Aversion and Model Misspecification: An Economic Perspective. *Statistical Science*, 31(4), 511-515.

Hansen, L. P. & Sargent, T. J. (2022). Structured Ambiguity and Model Misspecification. *Journal of Economic Theory*, 199, 105165.

Karni, E. (2022). A Theory-Based Decision Model. *Journal of Economic Theory*, 201, 105444.

Karni, E. & Vierø, M. L. (2013). "Reverse Bayesianism": A Choice-Based Theory of Growing Awareness. *American Economic Review*, 103(7), 2790-2810.

Karni, E. & Vierø, M. L. (2017). Awareness of Unawareness: A Theory of Decision Making in the Face of Ignorance *Journal of Economic Theory*, 168(C), 301-328

Klibanoff, P., Marinacci, M. & Mukerji, S. (2005). A Smooth Model of Decision Making under Ambiguity. *Econometrica*, 73(6), 1849-1892.

Maccheroni, F., Marinacci, M. & Rustichini, A. (2006). Ambiguity Aversion, Robustness, and the Variational Representation of Preferences. *Econometrica*, 74(6), 1447-1498.

Maćkowiak, B., Matějka, F. & Wiederholt, M. (2018). Dynamic Rational Inattention: Analytical Results. *Journal of Economic Theory*, 176, 650-692.

Marinacci, M. (2015). Model Uncertainty. *Journal of the European Economic Association*, 13(6), 1022-1100.

Mason, R. O. (1969). A Dialectical Approach to Strategic Planning. *Management Science*. 15(8) B403â€“B414.

Mitroff, I. I., & Betz, F. (1972). Dialectical Decision Theory: A Meta-Theory of Decision-Making. *Management Science*, 19(1), 11-24.

Mitroff, I. I., & Silvers, A. (2010). *Dirty Rotten Strategies: How We Trick Ourselves and Others into Solving the Wrong Problems Precisely*. Stanford University Press, Stanford CA.

Nickerson, J. & Zenger, T. R. (2004). A Knowledge-Based Theory of the Firm: The Problem-Solving Perspective. *Organization Science*, 15(6), 617-632.

Nickerson, J. & Argyres, N. (2018). Strategizing Before Strategic Decision Making. *Strategy Science*, 3(4), 592-605.

Packard, M. D. & Clark, B. B. (2020). On the Mitigability of Uncertainty and the Choice between Predictive and Nonpredictive Strategy. *Academy of Management Review*, 45(4), 766-786.

Pearl, J. (2009). *Causality*. Cambridge University Press, Cambirdge UK.

Rindova, V. & Courtney, H. (2020). To Shape or Adapt: Knowledge Problems, Epistemologies, and Strategic Postures under Knightian Uncertainty. *Academy of Management Review*, 45(4), 787-807.

Rothshild, M. & Stiglitz, J. (1970). Increasing Risk: I. A Definition. *Journal of Economic Theory*, 2(3), 225-243.

Schlaifer, R. O. & Raiffa, H. (1961). *Applied Statistical Decision Theory*. Division of Research, Harvard Business School, Cambridge, MA. Republished in Wiley Classic Library Series (2000).

Zellweger, T. M. & Zenger, T. R. (2021). Entrepreneurs as Scientists: A Pragmatist Approach to Producing Value Out of Uncertainty. *Academy of Management Review*, forthcoming.

Zhong, W. (2022). Optimal Dynamic Information Acquisition. *Econometrica*, 90(4), 1537-1582.

Appendix

A causal structure such as (9) generates the following sequence of expected values

$$\mathbb{E}(x_d \mid \theta_{des}, x_e, x_s) \equiv v_d(\theta_{des}, x_e, x_s) = \int_{X_d} x_d \; p(x_d \mid \theta_{des}, x_e, x_s) \; dx_d$$

$$\mathbb{E}\left[v_d(\theta_{des}, x_e, x_s) \mid \theta_e, \theta_{sh}, x_h)\right] \equiv v_{des}(\theta_{des}, \theta_e, \theta_{sh}, x_h) = \\ \int_{X_s} \int_{X_e} v_d(\theta_{des}, x_e, x_s) \; p(x_e \mid \theta_e) \; p(x_s \mid \theta_{sh}, x_h) \; dx_e \; dx_s$$

$$\mathbb{E}\left[v_{des}(\theta_{des}, \theta_e, \theta_{sh}, x_h) \mid \theta_h\right] \equiv v(\theta) = \int_{X_h} v_{des}(\theta_{des}, \theta_e, \theta_{sh}, x_h) \; p(x_h \mid \theta_h) \; dx_a$$

Consider for simplicity the following linear approximation

$$v_d(\theta_{des}, x_e, x_s) = \theta_{de} x_e + \theta_{ds} x_s, \qquad \text{with} \quad \mathbb{E}(x_e \mid \theta_e) = \theta_e, \quad \mathbb{E}(x_s \mid \theta_{sh}, x_h) = \theta_{sh} x_h$$

where we distinguish between the two elements θ_{de} and θ_{ds} of the vector of parameters θ_{des} that represent, respectively, the correlations between x_e and x_d and x_s and x_d. By replacing the two expected values in $v(\cdot)$, we obtain

$$v_{des}(\theta_{des}, \theta_e, \theta_{sh}, x_h) = \theta_{de}\theta_e + \theta_{ds}\theta_{sh} x_h$$

and

$$v(\theta) = \theta_{de}\theta_e + \theta_{ds}\theta_{sh}\theta_h$$

which is equation (10) in the text.