DISCUSSION PAPER SERIES

DP17392-2

MICROFOUNDATIONS OF LOW-FREQUENCY HIGH-IMPACT DECISIONS

Alfonso Gambardella, Arnaldo Camuffo, Fabio Maccheroni, Massimo Marinacci and Andrea Pignataro

ORGANIZATIONAL ECONOMICS

CEPR

ISSN 0265-8003

MICROFOUNDATIONS OF LOW-FREQUENCY HIGH-IMPACT DECISIONS

Alfonso Gambardella, Arnaldo Camuffo, Fabio Maccheroni, Massimo Marinacci and Andrea Pignataro

Discussion Paper DP17392-2
Published 17 June 2022
Submitted 05 October 2022

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

MICROFOUNDATIONS OF LOW-FREQUENCY HIGH-IMPACT DECISIONS

Abstract

We show that grounding low-frequency high-impact strategic decisions under uncertainty on the development and test of theories is a foundational strategic choice of firms, and an important determinant of performance. Our normative Bayesian model shows that decision-makers benefit from comparing alternative theories and when theories are reliable (less uncertain) they should not experiment and they should focus on their most plausible theory (highest prior). Otherwise, they ought to experiment with their less reliable theories because they learn more. We also show that the variability of the optimal experiment matches the variability of the prior. In particular, large scale experiments may be overprecise and penalize radical new theories with more variable priors. Overall, our framework explains how structured exploration improves strategic decisions by reducing model mis-specification and why performance heterogeneity and competitive advantage is created by exploring theories with higher variance.

JEL Classification: L21, L26, M13, M21

Keywords: Decision problem, Experiments, Exploration, Framing, Strategy, Theory, Uncertainty

Alfonso Gambardella - alfonso.gambardella@unibocconi.it
Department of Management and Technology, ICRIOS & ION Management Science Lab, Bocconi University, Milan and CEPR

Arnaldo Camuffo - arnaldo.camuffo@unibocconi.it
Department of Management and Technology, ICRIOS & ION Management Science Lab, Bocconi University, Milan

Fabio Maccheroni - fabio.maccheroni@unibocconi.it
Department of Decision Sciences & ION Management Science Lab, Bocconi University, Milan

Massimo Marinacci - massimo.marinacci@unibocconi.it
Department of Decision Sciences & ION Management Science Lab, Bocconi University, Milan

Andrea Pignataro - ap@iongroup.com
ION Group & ION Management Science Lab, Bocconi University, Milan

1 Introduction

We study the microfoundations of low-frequency high-impact (LFHI) strategic decisions which are decisions under uncertainty that impact on firm performance and cannot be grounded on extensive past data and algorithms (Choi & Levinthal, 2022). Examples are decisions about product or service innovation (new markets or technologies), capital structure strategy (ownership and governance), M&As (markets for control and firm boundaries), top talent and management team hiring (Camuffo, Gambardella, & Pignataro, 2022).

LFHI decisions differ from high-frequency low-impact (HFLI) decisions which are routine decisions within firms for which decision makers (DMs) can rely on past data. In LFHI decisions, DMs have neither pre-defined, readily available decision problems to solve, nor context-specific past data to ground strategic decisions on. Under these conditions, DMs first aim at increasing their knowledge of the world by formulating and choosing decision problems, i.e. building theories about reality (framing decisions) before they can solve them by choosing actions conditional on the theory they have selected.

This is consistent with the body of research that emphasizes problem definition (Nickerson & Zenger, 2004), theorizing (Felin & Zenger, 2017), and mental representations (Csaszar & Levinthal, 2016) in strategic decision making, as well as with the definition of strategy as the smallest set of choices to optimally guide other choices (Van den Steen, 2017).

Since DMs do not observe the "true" states of the world, they will never know whether their theories are correct (type III error, Raiffa, 1968; Mitroff & Betz, 1972). They can only formulate and experiment with alternative theories and ultimately choose the theory that they believe is more plausible.

In our model of LFHI decisions, DMs have prior probabilities on theories, and they can be more or less confident about these probabilities depending on whether they expect that an experiment will significantly change them. We show that DMs benefit from comparing two theories with different priors. When the variability of both prior is low, DMs commit to the theory with the higher prior (more plausible) and do not experiment. Otherwise, they experiment with the theory with the more variable prior (less reliable). When this is the theory with the higher prior, DMs conduct a confirmation experiment aimed at testing whether the posterior on this theory is not lower than the prior of the other theory. Alternatively, DMs conduct a falsification test aimed at testing whether the posterior on the theory with the lower prior is higher than the prior of the other theory.

We posit that theory exploration and knowledge accumulation are the foundations of strategic decision-making under uncertainty. Our emphasis on theories and experiments, and the "methodic doubt" that encourages DMs to explore and test alternative theories, evokes the application of a scientific approach to strategic decisions (Camuffo, Cordova, Gambardella, & Spina, 2020; Zellweger & Zenger, 2022). It also redirects strategy research back to its original mission of scientifically grounding top management's decision-making under uncertainty and reconnects strategy to decision science (Schlaifer & Raiffa, 1961; Ansoff, 1965; Mason, 1969, Ansoff & Brandenburg, 1971; Ackoff, 1974).

Section 2 grounds the study in the strategy literature. Section 3 provides the microfoundations of LFHI decisions. Section 4 presents our model. Section 5 provides an example (with two additional examples in the Appendix). Section 6 discusses limitations and future research directions.

2 Motivating literature

This paper is theoretically grounded on the burgeoning literature that has recently questioned the foundations of cross-firm strategic heterogeneity and competitive advantage, suggesting that they reside on how strategists navigate uncertainty (Alvarez & Porac, 2020).

A stream of strategy research parallels strategies to "theories" (Ehrig & Schmidt, 2022; Felin & Zenger, 2017), i.e. abstract models of reality strategists envision to ground their business decisions. Another stream suggests that strategists use mental representations, i.e. models of reality, to generate predictions, shape profitable strategies (Csaszar & Levinthal, 2016) and affect strategic performance (Csaszar, 2018). Another stream has underlined the importance of strategic problem definition as key antecedent of strategic decisions (Nickerson & Zenger, 2004). According to this perspective, the core of strategic decision-making is the imaginative act through which strategists recognize, discover or create strategic problems (Nickerson & Argyres, 2018).

While this emergent literature has shifted the focus of strategy research from resources and capabilities to strategists' cognition and agency, we still lack an understanding of why different models of reality ("theories", "mental representations" or "strategic problems") make a difference in terms of strategic decision making, how strategists can develop them under conditions of fundamental uncertainty and how these models of reality shape classic strategic decisions in terms of competitive positioning. This paper represents a step in this direction.

At the same time. this paper is motivated by the literature on strategic experimentation, which sees experiments as ways to elicit signals of an underlying state variable (Agrawal, Gans & Stern, 2021) and, more broadly, as methods to collect information regarding future contingencies. Experiments inform and redirect strategic decisions (Kerr, Nanda, & Rhodes-Kropf, 2014; Thomke, 2020; Koning, Hasan, & Chatterji, 2022) and, conditional on their cost, can be thought of as real options (Trigeorgis & Reuer, 2017), which strategists buy and exercise to learn about optimal strategies (Adner & Levinthal, 2004; Posen, Leiblein & Chen, 2018).

While this literature provides a rationale for flexible strategies (McGrath, 2001), with experiments enabling strategic adaptation to uncertain environments, it does not address how strategists can leverage experimentation to cope with more fundamental uncertainty. To this aim, this paper explores three developmental avenues. The first avenue connects strategic experimentation with the streams of strategy research on theories, mental representations and strategic problems. If strategists choose theories of reality to shape strategic actions, they need to define and experiment with them in order to learn which theory to use. The second avenue points to the conditions under which strategists experiment with their theories and with which theories they experiment. Recent studies show that DMs learn more by experimenting when their priors are moderate and the experimentation cost is sufficiently low (Che & Mierendorff, 2019). Furthermore, they show that DMs should experiment with alternatives business ideas before committing (Gans, Stern & Wu, 2019). The third avenue calls for distinguishing between uncertainty regarding the quality of DMs' theories and uncertainty regarding DMs' strategies (Agrawal, Gans, & Stern, 2021; Gans, 2022).

Our framework elaborates on these three avenues. Particularly, we suggest that DMs should experiment with alternative theories, and then follow the strategies (actions) that fit their most plausible one. We suggest that experimenting with theories enhances learning to a greater extent than experimenting with strategies.

Our study also refers to behavioral strategy (Lovallo & Sibony, 2018) and more specifically to the literature about cognitive biases in strategic decision-making (Kahneman, Lovallo, & Sibony, 2019). LFHI decisions are grounded on evaluative judgements and characterized by well-documented errors. These errors can be systematic, when they derive from cognitive biases (Kahneman, 2011), or random ("noise") (Kahneman, Sibony, & Sunstein, 2021).

While this literature aims at de-biasing and reducing "noise" raising DMs' awareness about their heuristics and biases, we lack a more general approach through which strategists can navigate fundamental uncertainty and improve the overall framing of their decisions. Our framework provides a rationale to form, test and update beliefs about theories so that strategists can be less exposed to decision making biases and judgement noise.

Finally, recent research has emphasized that strategists cope with uncertainty selecting what to pay attention to. For example Hanna, Mullainathan, & Schwartzstein (2014) model DMs' behavior as "learning through noticing". DMs choose which input dimensions to attend to and subsequently learn about them from available data. Similarly, "rational inattention" theory suggests that DMs, characterized by memory/beliefs and limited attention resources, have to choose which information to attend to and which information to ignore, optimally solving the trade-off between learning about the current optimal action and the best predictors of future optimal actions (Maćkowiak, Matějka, & Wiederholt, 2018). Furthermore, Ocasio's (2011) attentional perspective and attentional engagement concepts point to the criterion through which DMs select what to pay attention to and how much intentional effort DMs make in doing that.

While this research has highlighted the importance of the cognitive processes through which strategists focus on some variables and discard others, we still lack an understanding of how strategists ought to make such choices in order to better frame their strategic decisions.

3 Microfoundations

3.1 Attributes and exploration domains

Since the seminal work of Savage and Wald, decision-making under uncertainty is modelled as the resolution of a decision problem in which DMs have to choose among alternative courses of actions whose consequences depend on contingencies outside the DMs control.[1]

Definition 1 (classical decision problem). *A decision problem is a quartet* (A, S, C, ξ) *in which:*

- A *is a collection of available actions*
- S *is the space of all payoff-relevant contingencies called states of nature*
- C *is a collection of consequences*
- $\xi : A \times S \to C$ *is a consequence function that details the consequence* $c = \xi(a, s)$ *of action* a *when state* $s \in S$ *obtains*

DMs confront (state) uncertainty when they ignore the true state of nature. Let $p_\tau(s) \equiv p(s \mid \tau)$, or p_τ for short, be the "true" probability distribution with which states occur, indexed by τ. If DMs know this probability distribution, they choose the action $a \in A$ that maximizes the expected utility $V_\tau(a) = \sum_{s\in S} u(\xi(a,s))p_\tau(s)$, where $u(\xi(a,s))$ is the utility of the consequence of action a in state s, and the subscript τ denotes that V also depend on characteristics of the correct probability distribution.[2] The optimal action is a function of τ, i.e. $a^*(\tau)$. We can then write the optimal expected utility as

$$V_\tau = \sum_{s\in S} u(\xi(\tau, s))p_\tau(s) \tag{1}$$

[1]See Gilboa (2009) and, for a strategy perspective, Hofer and Schendel (1978) and Mintzberg (1994).

[2]Formally, $p_\tau \in \Delta(S) = \left\{p \in \mathbb{R}^S_+ : \sum_s p(s) = 1\right\}$ where $\Delta(S)$ is the set of all probability distributions on S.

A classical decision problem presupposes that DMs know the relevant contingencies that form the state space S and it abstracts from unforeseen contingencies. This is a demanding assumption especially in the case of LFHI decisions in which DMs do not have sufficient data to know the states $s \in S$ and their probability distribution $p_\tau(s)$ (Hansen, 2014; Hansen & Marinacci, 2016).

For this reason, in LFHI decisions, DMs start their exploration by identifying, through conjectures and mental experiments, what they believe are relevant attributes for the decision.

Definition 2 (**attributes**). *An attribute X_j is the set of all the alternative outcomes of one element of a problem. A space of attributes $X = \prod_{j \in J} X_j$ is the space of all the attributes $j \in J$ that DMs believe are relevant for their problem. An attribute X_j and a space of attributes X are maps $D \to X$ that reduce the dimensionality of the exploration problem within the DMs' domain D.*

Definition 3 (**domain**). *A domain is the set of all attributes $\mathcal{D} = \prod_{j \in J} X_j$, $J = \{1, 2, ...N\}$ that DMs know or could be aware of at any given moment in time given their knowledge, experience or reference points. It does not necessarily contain S. This depends on whether DMs are aware of all the relevant contingencies for the consequences of their problem. Domains change as DMs learn.*

Examples of attributes are the level of demand for a product, the quality of a technology, or the degree of complementarity of an acquisition target. For instance, demand can be high or low, with $X_d = \{high, low\}$, and similarly for quality of technology or complementarity. Suppose that DMs focus on demand and technology to assess the potential of a new product. If the technology attribute is $X_t = \{good, bad\}$, the space of all attributes is $X = X_d \times X_t = \{(high, good), (high, bad), (low, good), (low, bad)\}$. If the space of each attribute was a continuous measure over the non-negative real numbers (e.g. number of customers and an index of the quality of the technology), then $X_d = \{x : x \in \mathbb{R}_0^+\}$, $X_t = \{x : x \in \mathbb{R}_0^+\}$, and $X = X_d \times X_t = (\mathbb{R}_0^+)^2$.

3.2 Models and theories

In using models and theories within a decision problem we adopt a recent approach in decision theory.[3] DMs explore the space X through logical links ($\wedge$, $\vee$, $\oplus$, $\to$, $\leftrightarrow$). Logical links focus DMs on a subset $x \in X = \{x, \overline{x}\}$, where $\overline{x}$ is the complement of x, and they identify a *family* of probability distributions $P_\Theta = \{p_\theta\}_{\theta \in \Theta}$ of the probability p_θ of this subset of X, where p_θ, parameterized by $\theta \in \Theta$, is a shorthand notation for this probability, or $p_\theta(x) \equiv p(x \mid \theta)$ and $\mu(\theta)$ is the likelihood of θ.[4] Thus, θ identifies a specific probability distribution of x, $p_\theta(x)$, within the family P_Θ. The set Θ identifies the family of probability distributions, and it is defined by the DMs's theory. DMs entertain a prior π on Θ, which is the probability with which they believe that the "true" θ belongs to Θ, or simply the probability with which they believe that the theory is true.

Definition 4 (**model**). *A model θ is the realization of a parameter (or vector of parameters) that identifies a specific probability distribution p_θ within a family of probability distributions $P_\Theta = \{p_\theta\}_{\theta \in \Theta}$ on the space of attributes.*

Definition 5 (**theory**). *A theory is a family of probability distributions $P_\Theta = \{p_\theta\}_{\theta \in \Theta}$. We call theory P_Θ or Θ interchangeably.*

This definition epitomizes the notion of theory used by strategy scholars (e.g. Felin and Zenger, 2017).

[3] See Cerreia-Vioglio et al. (2013a, 2013b), Marinacci (2015), and Cerreia-Vioglio et al. (2020).

[4] Of course, DMs could use more logical links at the same time and focus on more subsets. We stay for simplicity with the simpler case of a partition made of two elements of X.

As an example, suppose that attributes are "competence" of CEO and future "growth" of the company, with $X_c = \{high, low\}$ and $X_g = \{\gamma : \gamma \in \mathbb{R}\}$. In this case, $X = X_c \times X_g$ is a $2 \times \mathbb{R}$ space with generic element $x = (i, \gamma)$, where $i = (high, low)$ and $\gamma \in \mathbb{R}$. Suppose that DMs develop the theory that a competent CEO implies non-negative growth. This logical link is an implication ($\rightarrow$) that reduces the probability of the subset $x = \{(i, \gamma) : i = high \text{ and } \gamma < 0\} \in X$.

DMs then focus on the coarser space $X = \{x, \overline{x}\}$, where $\overline{x}$ is the complement of x, and they identify a *family* of probabilistic models $P_\Theta = \{p_\theta(x)\}_{\theta \in \Theta}$, where $p_\theta(x)$ is the probability of x, and $\mu(\theta)$, $\theta \in \Theta$, is the likelihood of θ. DMs predict low p_θ, and if the probablity of high p_θ increases with θ (e.g. θ is the expected value of p_θ), the theory suggests that DMs only consider likelihoods $\mu(\theta)$ such that θ is "small". For example, they pick $\Theta = \{\theta : \theta < \theta^*\}$ where θ^* is a threshold.

DMs may entertain different theories, which may stem from different logical links within the same space of attributes or from different spaces of attributes. For example, DMs may predict that some factors raise the competence of CEOs and their ability to sustain company growth. In this case, the logical link is a conjunction ($\wedge$) and DMs focus on $x = \{(i, \gamma) : i = high \text{ and } \gamma \geq 0\}$. In this case, Θ is a set of relatively "high" θ. The case of different attributes is straightforward.

Table 1 is a "truth table" in which the five basic logical links identify which element $x \in X$ is true or false when $x_c = high$ or $x_g = \{\gamma : \gamma \geq 0\}$ are individually true or false. A belief on anyone of these logical links will increase the joint (subjective) probability of all the true elements and lower the joint (subjective) probability of all the false elements.

Table 1: *Basic Logical Links*

($x_c = high$; $x_g = \{\gamma : \gamma \geq 0\}$); T, F = true, false

x_c	x_g	conjunction $x_c \wedge x_g$	inclusive disjunction $x_c \vee x_g$	exclusive disjunction $x_c \oplus x_g$	implication $x_c \Rightarrow x_g$	implication $x_c \Leftarrow x_g$	biconditional $x_c \Leftrightarrow x_g$
T	T	T	T	F	T	T	T
T	F	F	T	T	F	T	F
F	T	F	T	T	T	F	F
F	F	F	F	F	T	T	T
		(and)	(or)	(not both)	(if)	(if)	(iff)

3.3 Value of theories

Once DMs fix Θ and $\mu(\theta)$, they pick the action a that maximizes the classical subjective expected utility criterion $V_\Theta(a) = \sum_{\theta \in \Theta} v(a, \theta)\mu(\theta)$, where $v(a, \theta)$ is the expected utility of the consequence of action a and the subscript Θ denotes that V depends on the DMs' theory, which sets the likelihoods $\mu(\theta)$ that they use to compute this expected value. The optimal action is a function of the theory, and we write it as $a^*(\Theta)$. In turn, we can write the optimal expected utility of theory Θ as

$$V_\Theta = \sum_{\theta \in \Theta} v(\Theta, \theta)\mu(\theta) \tag{2}$$

To help intuition, take again our example where $X = \{x, \overline{x}\}$, $x = \{(i, \gamma) : i = high \text{ and } \gamma < 0\}$, DMs predict that $p_\theta(x)$ is small, and set for simplicity $\mathbb{E}p_\theta(x) = \theta$. In this case the expected utility $v(\Theta, \theta)$ is the average of the utilities that DMs enjoy under x and $\overline{x}$ weighted by the expected probabilities θ and $1 - \theta$. The expected utility of theory Θ, V_Θ, is the expected value of $v(\Theta, \theta)$ across all models

$\theta \in \Theta$ weighted by the likelihood of each model $\mu(\theta)$. A different logical link, possibly on a different space of attributes, will focus on different x and $\overline{x}$, a different probability p_θ and $\theta = \mathbb{E}p_\theta$, a different Θ and therefore a different $\mu(\theta)$, producing a different V_Θ.

In HFLI decisions, DMs can better approximate S and p_τ. They have enough past information to realize the relevant states, and to know p_τ, or even if they do not know it, they can predict the expected value of an action because any error that they make is just a random variation around p_τ (e.g., they do not know the variance of the distribution). In these cases, they do not need to develop a theory because the distribution p_τ is either known, or known up to a random error that does not affect the predicted expected value.

In LFHI decisions, instead, past information is not available, and DMs have to perform a preliminary step. They have to identify a theory, and then optimize actions. As we will see, this requires that they experiment with alternative theories before they settle on one.[5]

3.4 Experiments

Through exploration of new attributes or different logical links, DMs identify partitions of the space of attributes they focus upon, and identify different models θ and theories Θ. As they cannot conduct experiments directly on their theories (that is on the prior π on the theory Θ), they run experiments (mental or real) to update the probabilities $\mu(\theta)$.

They change their prior π on Θ depending on their estimates of θ. In our example we set $\Theta = \{\theta : \theta < \theta^*\}$. Suppose that DMs find θ close to θ^* or even higher. They may start doubting about their theory, and reduce their prior π. Alternatively, they may be confident about their theory, and reduce their prior marginally or not at all. Similarly, they may lower their prior, or change it marginally, if they find evidence of θ close to 0. In what follows we speak interchangeably of testing models or theories. However, what we mean is that DMs test models but then use the information outcome of the experiment to draw implications about their theories.

Changes in the prior π of the theory Θ from experimental evidence about models θ depend on behavioral traits of DMs. DMs with greater "methodic doubt", or more generally who believe that their theories are less reliable, are more likely to significantly update their priors on theory depending on experimental outcomes.

Definition 6 (**experiment**). *An experiment is a map $f : \Theta \to \Delta(Y)$ from models θ to probability distributions $f(y \mid \theta)$, $y \in Y$, where $f(y \mid \theta)$ is the probability of receiving signal y under model θ.*

An experiment determines a joint probability Pr on $Y \times \Theta$, where $\Pr(y, \theta) = f(y \mid \theta)\,\mu(\theta)$ is the probability of observing y and that θ is the true parameter. The likelihood $\mu(\theta)$ is the marginal, on models, of this joint distribution, that is

$$\Pr(\theta) = \sum_{y \in Y} \Pr(y, \theta) = \sum_{y \in Y} f(y \mid \theta)\,\mu(\theta) = \mu(\theta)$$

[5] DMs may be sensitive to the fact that they may not be working with the right theory. Their decision problem about actions, which occurs after the decision to commit to a theory Θ, may then in any case consider the more general decision criterion $V_\mu^\phi(a) = \sum_{\theta \in \Theta} \phi(v(a, \theta))\,\mu(\theta)$, where now θ, Θ and $\mu(\theta)$ play the same role as s, S, and p_τ in the classical decision problem, and a is an action, again similar to the classical decision problem. Here ϕ is a strictly increasing and continuous function (Klibanoff et al. 2005). If concave, ϕ accounts for a negative attitude toward the uncertainty over the correct model θ. In what follows, we work for simplicity with the linear criterion. However, all our analyses below apply under these more general specification as well.

The *predictive distribution* on signals is given by the marginal of Pr on signals,

$$\Pr(y) = \sum_{\theta\in\Theta} \Pr(y,\theta) = \sum_{\theta\in\Theta} f(y \mid \theta)\,\mu(\theta) = f(y)$$

Finally, the *posterior distribution* on models, denoted by μ_y, gives the probability of model θ upon receiving signal y, or

$$\mu_y(\theta) = \Pr(\theta \mid y) = \Pr(y \mid \theta)\frac{\Pr(\theta)}{\Pr(y)} = \frac{f(y \mid \theta)\,\mu(\theta)}{\sum_{\theta'} f(y \mid \theta')\,\mu(\theta')}$$

3.5 Choice of experiments

After the experiment the value of a theory upon receiving signal y is

$$V_{f,\Theta}(y) = \sum_{\theta\in\Theta} v(\Theta,\theta)\,\mu_y(\theta)$$

where subscripts f denotes that V now also depends on characteristics of the experiment f.

Before performing it, the expected value of experiment f on theory Θ is

$$V_{f,\Theta} = \sum_{y\in Y} V_{f,\Theta}(y)\,f(y) = \sum_{\theta\in\Theta} v(\Theta,\theta)\underbrace{\sum_{y\in Y} f(y)\,\mu_y(\theta)}_{\sum_{y\in Y} f(y|\theta)\mu(\theta)=\mu(\theta)} = V_\Theta$$

where the last step stems from the application of the Bayes theorem. This is important in that it says that before running the experiment on a given theory DMs expect that the experiment yields the same value of the theory expected before the experiment.

Therefore, experiments are worth only when DMs run them to compare more than one theory. If $\tilde{\Theta}$ is the set of all theories identified by DMs, after the experiment the value of the optimal theory upon receiving signal y is

$$\mathbb{V}_{f,\Theta}(y) = \max_{\Theta\in\tilde{\Theta}} \sum_{\theta\in\Theta} v(\Theta,\theta)\,\mu_y(\theta)$$

Before performing it, the expected value of experiment f is

$$\mathbb{V}_{f,\Theta} = \sum_{y\in Y} f(y)\,\mathbb{V}_{f,\Theta}(y) = \sum_{y\in Y} f(y)\max_{\Theta\in\tilde{\Theta}} \sum_{\theta\in\Theta} v(\Theta,\theta)\,\mu_y(\theta)$$

This value is greater than or equal to the value of the optimal theory without experimentation[6]

$$\mathbb{V}_\Theta = \max_{\Theta\in\tilde{\Theta}} \sum_{\theta} v(\Theta,\theta)\,\mu(\theta)$$

When a set of experiments F is considered, all with the same cost, the optimal choice of experiments amounts to solve the (constrained) optimization problem

$$\max_{f\in F} \mathbb{V}_{f,\Theta}$$

[6] It is enough to observe that, for each $\theta\in\Theta$, $\sum_{y\in Y} f(y)\,\mu_y(\theta) = \mu(\theta)$.

Therefore DMs *choose attributes, theories and experiments.* Theory of framing decisions deals with the ability to select attributes and logical links that lead to concentration of probabilities in a domain (Cai & Lim 2022).Theory of the experiment deals with the ability to design experiments that yield more precise outcomes at given cost. Better theories or experiments can produce quicker or greater concentration of probabilities at lower cost. Better framing depends on both better theories of attributes and their logical links and more informative or precise designs of experiments. DMs then make decisions about actions on a set of states in which they have reduced uncertainty, and therefore learned, to a greater extent.

4 A model of framing LFHI decisions

We develop a model of framing LFHI decisions that builds on the microfoundations discussed in the previous section.

Without loss of generality, we assume that DMs consider two attributes $X_0 = \{x_0, \overline{x}_0\}$ and $X_1 = \{x_1, \overline{x}_1\}$, where x_i and $\overline{x}_i$, $i = 0, 1$, represent two elements of the attributes.[7]

DMs posit a family of models or theory $P_\Theta = \{p_\theta\}_{\theta \in \Theta}$ over $X_0 \times X_1$, and we set for simplicity

$$\mathbb{E}p_\theta(x_0) = \theta_0 \quad \text{and} \quad \mathbb{E}p_\theta(x_1) = \theta_1$$

where, with a convenient abuse of notation, $x_0 \equiv \{(x_0, x_1), (x_0, \overline{x}_1)\}$ and $x_1 \equiv \{(x_0, x_1), (\overline{x}_0, , x_1)\}$. The parameter space is $\Theta = \Theta_0 \times \Theta_1$, each parameter $\theta = (\theta_0, \theta_1)$ is a bidimensional vector, and the joint likelihood is $\mu(\theta_0, \theta_1)$.

We make three assumptions to simplify the analysis.

Assumption 1. *DMs can only run one experiment*

Assumption 2. *Attributes are independent under each θ_i. Therefore, the likelihood is a product $\mu(\theta_0, \theta_1) = \mu^0(\theta_0)\mu^1(\theta_1)$*

Assumption 3. $v_i \equiv v(\Theta_i, \theta_i) = ln\frac{\theta_i}{1-\theta_i} \sim \mathcal{N}(\overline{\mu}_i, \sigma_i^2)$, $\theta_i \in \Theta_i$, $i = 0, 1$

Assumption 1 avoids the complications of a continuation value. This assumption is not unrealistic. On many occasions DMs only have one opportunity to test an important theory about their business model, as in the example in the next section.

Assumption 2 implies that attributes are more likely to come from different spaces.[8] The Bayesian values of the two frameworks can thus be separated. DMs can take advantage of this feature by experimenting separately on each θ_i. In particular, an experiment that depends only on θ_i can be written as $f_i(\cdot \mid \theta_i)$ and is easily seen to update only μ^i through $\mu_y^i(\theta_i)$.[9] As a result, an experiment on θ_i changes only the value of $V_{f,\Theta_i}(y)$.

Assumption 3 specifies an expected utility over the probabilities θ_i of the elements x_i of the two attributes $X_0 = \{x_0, \overline{x}_0\}$ and $X_1 = \{x_1, \overline{x}_1\}$, given theory Θ_i. Higher θ_i raises expected utility because, given theory Θ_i, DMs envisage actions that generate a higher expected value v_i when the expected probability of x_i is higher. The normal distribution enables us to work with a specific functional form,

[7] Section 3.2 shows that these two elements could be the outcome of a partition of a space of attributes made of more elements that logical links combine and separate in two elements.

[8] Formally, this assumption implies $V_{\Theta_i} = \sum_{\theta_i,\theta_j} v(\Theta_i, \theta_i)\mu(\theta_i, \theta_j) = \sum_{\theta_i,\theta_j} v(\Theta_i, \theta_i)\mu^i(\theta_i)\mu^j(\theta_j) = \sum_{\theta_j} \mu^j(\theta_j) \sum_{\theta_i} v(\Theta_i, \theta_i)\mu^i(\theta_i) = \sum_{\theta_i} v(\Theta_i, \theta_i)\mu^i(\theta_i)$, where $i, j = 0, 1$.

[9] That is $\mu_y(\theta_i, \theta_{1-i}) = \mu_y^i(\theta_i)\mu^{1-i}(\theta_{1-i})$.

so that we draw the implications of our model in a simple and intuitive way. The normal distribution is, however, flexible depending on the parameters $\overline{\mu}_i$ and σ_i, where $\overline{\mu}_i \equiv V_\Theta$ in (2), and σ_i accounts for the variability of v_i. DMs with high σ_i believe that their utility v_i can be very different from the expected value. This occurs, typically, when theories are "novel", "ambitious" or "ground-breaking".

DMs commit to the theory with the highest expected utility. To do so, they can run an experiment to update $\overline{\mu}_i$.[10] The experiment employs n_i observations $\{y_{i1}, y_{i2}, ..., y_{in_i}\}$. For example, $y_{ij} = ln\frac{s_{ij}}{1-s_{ij}}$, $j = \{1, 2, ..., n_i\}$, are observations on v_i in n_i randomly generated groups in which DMs observe x_i or $\bar{x}_i$. Depending on whether the experiment is mental or real, s_{ij} is the share of pseudo- or real observations of x_i (vs $\overline{x}_i$) in each group j. In our example in the previous section, the groups could be different clusters of firms with similar characteristics, and DMs look for the share of cases in which a competent CEO is associated with negative growth.

Let $\overline{y}_i \sim \mathcal{N}(\overline{\mu}_i, \frac{\sigma^2_{y_i}}{n_i})$, where $\overline{y}_i = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}$ is the sample mean, $\sigma^2_{\overline{y}_i} \equiv \frac{\sigma^2_{y_i}}{n_i}$, and $\sigma^2_{y_i}$ is the variance of the experiment. It is natural to assume that the sample mean has the same expected value $\overline{\mu}_i$ of the process. Normal distributions are conjugate priors, and after the experiment, when DMs observe $\overline{y}_i$, the update is

$$\mathbb{E}(v_i \mid \overline{y}_i) = (1 - \omega_i)\overline{\mu}_i + \omega_i \overline{y}_i \tag{3}$$

where $\omega_i \equiv \frac{\sigma^2_i}{\sigma^2_{\overline{y}_i} + \sigma^2_i}$. Before the experiment, DMs expect that the experiment yields $\overline{\mu}_i$ because, using (3), the expected value of $\overline{y}_i$ before the experiment is $\overline{\mu}_i$.

Without loss of generality we set $\overline{\mu}_0 > \overline{\mu}_1$. Thus, before running the experiment, Θ_0 is the most plausible theory, that is the theory with the highest prior. However, DMs benefit from questioning their theories and doubt about the strength of their priors they hold on them. This occurs by testing alternative theories. Specifically, DMs can either run an experiment on theory Θ_1 to estimate whether, updated by the experiment, its prior becomes higher than the prior on Θ_0, or an experiment on Θ_0 to estimate whether, updated by the experiment, its prior becomes smaller than the prior on Θ_1.

Whatever experiment DMs conduct, before the experiment they choose an optimal threshold $\overline{y}^*_i$ such that if they observe $\overline{y}_i > \overline{y}^*_i$, they focus on theory Θ_i, otherwise they focus on theory Θ_j, $j = 1, 2$, $j \neq i$. Experiments command a cost $c > 0$ that, for simplicity, we assume to be fixed and the same for the two experiments. The net expected values of the experiments on Θ_i, $i = 0, 1$, are then

$$\mathbb{V}_i = \overline{\mu}_i \left(1 - \Phi^*_i\right) + \omega_i \sigma_{\overline{y}_i} \phi^*_i + \overline{\mu}_j \Phi^*_i - c \tag{4}$$

where Φ^*_i and ϕ^*_i are the cumulative and density standard normal distributions evaluated at $\frac{\overline{y}^*_i - \overline{\mu}_i}{\sigma_{\overline{y}_i}}$.

The first two terms of this expression are the expected values $\mathbb{E}(v_i, \overline{y}_i \geq \overline{y}^*_i)$, which we obtain by integrating (3) over $\overline{y}_i \geq \overline{y}^*_i$ and by using the property of the truncated normal distribution that $\mathbb{E}(\overline{y}_i, \overline{y}_i \geq \overline{y}^*_i) = \sigma_{\overline{y}_i} \phi^*_i$. The third term captures the fact that, if the experiment yields $\overline{y}_i < \overline{y}^*_i$ DMs focus on the alternative theory which yields expected utility $\overline{\mu}_j$. After the experiment DMs choose $\overline{y}^*_i$ to maximize $\mathbb{V}_i$. The first order condition is

$$-\omega_i(\overline{y}^*_i - \overline{\mu}_i) + (\overline{\mu}_j - \overline{\mu}_i) = 0 \tag{5}$$

It is easy to see that the second order condition for a maximum is satisfied.

DMs choose one of three available experiments:

[10] As noted in Section 3.4, DMs can only run experiments on models: higher $\overline{\mu}_i$ indicates higher expected probability θ_i of the subspace x_i identified by the logical links of DMs' theory, which is associated with a higher prior on the theory. Similarly, DMs associate higher variability of models σ_i to higher variability of their priors on the theory.

- $f_\varnothing$, *no draw from* X_0 *or* X_1 - that is, no experiment, and DMs stick to the most plausible theory Θ_0
- f_0, *a single draw from* X_0 - that is, an experiment with attribute 0 or theory Θ_0
- f_1, *a single draw from* X_1 - that is, an experiment with attribute 1 or theory Θ_1

Proposition 1 (***choice of experiment***)**.** *Let* $\tilde{\sigma}_i^2$ *be the value of* σ_i^2 *that yields* $\mathbb{V}_i = \overline{\mu}_i$ *and* $g : \sigma_0^2 \to \sigma_1^2$ *the function* $\sigma_1^2 = g(\sigma_0^2)$ *defined implicitly by* $\mathbb{V}_1 = \mathbb{V}_0$*. DMs choose experiments as follows:*

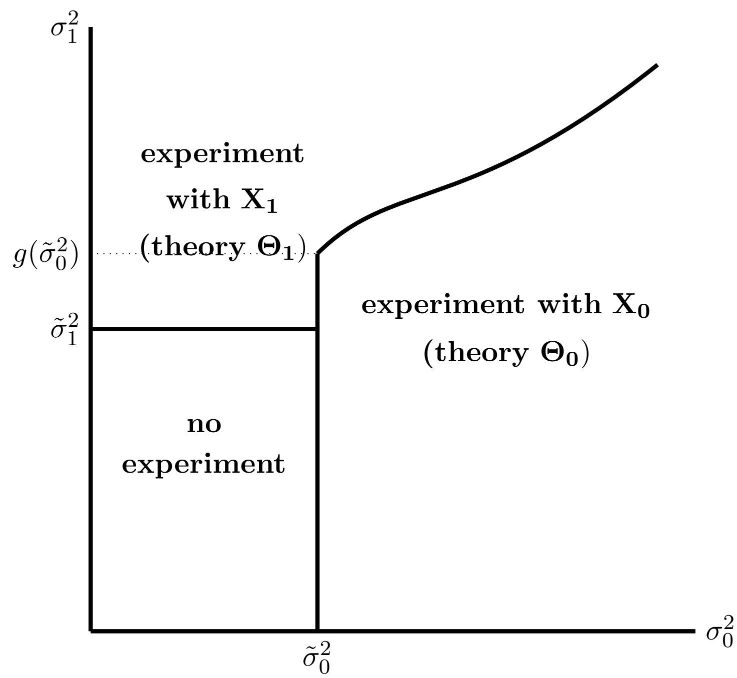
- $\left(\sigma_1^2 < \tilde{\sigma}_1^2 \text{ and } \sigma_0^2 < \tilde{\sigma}_0^2\right) \Rightarrow f_\varnothing$, *that is* ***DMs do not experiment*** *and pick theory* Θ_0
- $\left(\sigma_1^2 < g(\sigma_0^2) \text{ and } \sigma_0^2 \geq \tilde{\sigma}_0^2\right) \Rightarrow f_0$, *that is* ***DMs experiment with attribute 0 or theory*** $\boldsymbol{\Theta_0}$, *and pick theory* Θ_0 *or* Θ_1 *depending on whether the experiment yields* $\overline{y}_0 \geq \overline{y}_0^*$
- $\left(\sigma_0^2 < \tilde{\sigma}_0^2 \text{ and } \sigma_1^2 \geq \tilde{\sigma}_1^2\right)$ *or* $\left(\sigma_0^2 \geq \tilde{\sigma}_0^2 \text{ and } \sigma_1^2 \geq g(\sigma_0^2)\right) \Rightarrow f_1$, *that is* ***DMs experiment with attribute 1 or theory*** $\boldsymbol{\Theta_1}$, *and pick theory* Θ_1 *or* Θ_0 *depending on whether the experiment yields* $\overline{y}_1 \geq \overline{y}_1^*$

The function $g(\sigma_0^2)$ *increases with* σ_0^2 *and* $\sigma_0^2 \geq \tilde{\sigma}_0^2 \Rightarrow g(\sigma_0^2) > \tilde{\sigma}_1^2$

Proof. Since $\overline{y}_i^*$ is chosen optimally, σ_i^2 affects (4) only by increasing ω_i. Thus, $\sigma_i^2 < \tilde{\sigma}_i^2 \Rightarrow \mathbb{V}_i < \overline{\mu}_i$. This also implies that increases in σ_1^2 require decreases in σ_0^2 to keep $\mathbb{V}_1 = \mathbb{V}_0$, which implies that g is an increasing function of σ_0^2. To complete the proof we show that $g(\tilde{\sigma}_0^2) > \tilde{\sigma}_1^2$, The value of $g(\tilde{\sigma}_0^2)$ is σ_1^2 when $\mathbb{V}_1 - \mathbb{V}_0 = 0$ and $\mathbb{V}_0 = \overline{\mu}_0$ because it is evaluated at $\tilde{\sigma}_0^2$. Using (4), obtain $\omega_1 = \frac{c+(\overline{\mu}_0-\overline{\mu}_1)(1-\Phi_1^*)}{\sigma_{\overline{y}_1}\phi_1^*} > \tilde{\omega}_1 = \frac{c-(\overline{\mu}_0-\overline{\mu}_1)\Phi_1^*}{\sigma_{\overline{y}_1}\phi_1^*}$, where $\tilde{\omega}_1$ is ω_1 evaluated at $\mathbb{V}_1 = \overline{\mu}_1$, that is when $\sigma_1^2 = \tilde{\sigma}_1^2$, and the inequality stems $\overline{\mu}_0 > \overline{\mu}_1$. Since ω_i increases with σ_i^2 then $g(\tilde{\sigma}_0^2) > \tilde{\sigma}_1^2$. ■

Figure 1 provides a graphical representation of Proposition 1.

Figure 1: Experimenting with X_0 vs X_1



According to this Proposition, DMs conduct an experiment if $\mathbb{V}_i > \overline{\mu}_i$, i.e. a necessary condition is that the net expected value of the experiment $\mathbb{V}_i$, defined by (4), is higher than the expected value of the theory $\overline{\mu}_i$ before the experiment. If no experiment yields a higher net expected value, DMs do not run any experiment and focus on Θ_0. If only one experiment (Θ_0 or Θ_1) yields a higher net expected value, DMs will run that experiment. If both experiments yield a higher net expected value, DMs will run the experiment on the attribute or theory i such that $\mathbb{V}_i > \mathbb{V}_j$.

Proposition 1 shows that DMs are more likely to experiment with theories with higher variability. The intuition is they generate more sizable updates of their priors. In our case, the experiment with theory Θ_0 can update $\overline{\mu}_0$ below $\overline{\mu}_1$, or the experiment with theory Θ_1 can update $\overline{\mu}_1$ above $\overline{\mu}_0$.

We call experiments on Θ_0 *confirmation experiment* because DMs want to make sure that the more plausible theory is indeed more plausible. We call experiments on Θ_1 *falsification experiment* because DMs conjecture that the alternative less plausible theory can be more plausible. In confirmation experiments DMs doubt about their more plausible theory. In falsification experiments they test les plausible innovative theories that can become more plausible. Moreover, they are more likely to run these tests the more radical is the innovation (higher variability of theory).

Proposition 2 (*optimal experiment*). *The optimal experiment implies* $\sigma_{\overline{y}_i} = \sigma_i$.

Proof. From (4) $\frac{\partial \mathbb{V}_i}{\partial \sigma_{\overline{y}_i}} = \left[-(\overline{\mu}_j - \overline{\mu}_i) \frac{\overline{y}_i^* - \overline{\mu}_i}{\sigma_{\overline{y}_i}^2} + \omega_i \left(\frac{\overline{y}_i^* - \overline{\mu}_i}{\sigma_{\overline{y}_i}} \right)^2 + \frac{\sigma_i^2 (\sigma_i^2 - \sigma_{\overline{y}_i}^2)}{(\sigma_{\overline{y}_i}^2 + \sigma_i^2)^2} \right] \phi_i^*$. By replacing $\overline{y}_i^* - \overline{\pi}_i$ from (5) the first two terms of this expression cancel out, which implies $\frac{\partial \mathbb{V}_i}{\partial \sigma_{\overline{y}_i}} = 0$ *iff* $\sigma_{\overline{y}_i} = \sigma_i$. It is easy to see that the second order condition for a maximum is satisfied. ■

Proposition 2 shows that the variability of the optimal experiment ought to match the variability of the prior. The insight of this Proposition is that the design of experiments ought to reproduce the conditions postulated by the prior. Thus, if the theory has highly variable outcomes, the experimental design (its outcomes) ought to reproduce such a wide range. In practice, this corresponds to designing "biased" experiments that enable DMs to detect "extreme" or "surprising" heterogenous effects (Cao, Koning & Nanda, 2021; Gans, 2022)

Proposition 2 warns against the overprecision of experiments. If $\sigma_{y_i} > \sigma_i$, a larger n_i sets $\sigma_{\overline{y}_i} = \sigma_i$. However, further increases in the scale of the experiment will reduce the value of the experiment, making the experiment overprecise. Since overprecision implies $\sigma_{\overline{y}_i} < \sigma_i$, which raises ω_i, it is easy to see from (5) that in a confirmation test this reduces too much the threshold $\overline{y}_0^*$. DMs become too lenient towards accepting the more plausible theory. In a falsification test, the threshold $\overline{y}_1^*$ becomes too high, making DMs overly inclined to accept the more plausible theory.

The issue is particularly relevant when DMs develop theories with highly variable outcomes (high σ_1) that they test against a more plausible theory. In these cases, DMs should choose large $\sigma_{\overline{y}_1}$ to mimic the variability of the theory, and then possibly increase n_i if $\sigma_{\overline{y}_1} > \sigma_1$. Otherwise, they penalize radical new theories.

5 Luxottica

Founded by Leonardo Del Vecchio in 1961, today EssilorLuxottica is the world leader in the design, manufacturing and distribution of ophthalmic lenses, frames and sunglasses. By the mid 1980s, Del Vecchio had already managed to make the company a European leader in the design and manufacturing of eyeframes for spectacles (Camuffo, 2003).

At the time, Del Vecchio had two alternative theories of how the industry would work and Luxottica should look like in the future. We can think of these two theories as sets of attributes, connected by logical links and characterized by a prior.

The first theory was that people with eyesight defects 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. If you control opticians, you 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 allow further market penetration and market share.

This theory rested on a few attributes that Del Vecchio believed to be relevant for the market of *eyeglasses as medical devices*, most notably "opticians are key actors in this market" ({*yes, no*}), "scale-based production generates efficiency" ({*yes, no*}), and "growth and profitability" ({*high, low*}). He connected them logically, for example positing that "access to opticians and large scale in-house production *implies* growth and profitability." In the framework of Section 3.2, he partitioned $X = \{x, \overline{x}\}$, with $x = \{yes, yes, low\}$, and theorized a set of $\theta \in \Theta$ such that, if higher θ lowers the expected $p_\theta(x)$, expected utility $v(\Theta, \theta)$ increases with θ as in Assumption 3 of Section 4.

The second theory was 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. The final customer is the key actor in the market since eyewear reflects personal style, identity and image. Access to designer brands and retail allow direct contact, profiling and control of the final customer. Customers' control allows high margins (premium prices) and global supply chain efficiency (low cost). This leads to business growth and profitability which create the conditions for a dominant competitive position and further investment.

In this case, Del Vecchio also focused on a few attributes that he believed to be relevant in the market for *eyeglasses as a style accessory*, most notably "end customers are key actors in the market" ({*yes, no*}) and "customers respond to fashion-based price differentiation" ({*yes, no*}). He connected them logically with the attribute "growth and profitability ({*high, low*}) , for example positing that "access to final customers and to designer brands *implies* growth and profitability," yielding the same logical structure of the previous theory.

Till the 1980s Del Vecchio had developed and consolidated his first theory, grounding Luxottica's strategy on it. In the late 1980s he had signals about the potential of the second theory. The first theory was still more plausible, in that it was based on deeper experience and information. However, he started to consider on which of the two theories Luxottica should ground its strategy.

He could run a confirmatory test on the first theory, prompted by the rise of eye surgery as a potential substitute for spectacles. This experiment, by providing information on future demand for spectacles, would allow to update the prior on the first theory, checking whether it remained more plausible than the second one. Alternatively, Del Vecchio could conduct a falsification experiment on whether people would actually use eyeglasses not as a medical or functional device, but as a style accessory. This experiment would allow to update his prior on the second theory, checking whether it would become more plausible than the first one.

Del Vecchio's first theory was already proven, dominant in the industry, and therefore reliable. He believed that eye surgery might represent a challenge, but unlikely to disrupt the market. In the language of our model, this theory Θ exhibited low σ^2.

In contrast, the second theory ("eyewear") was novel, unexplored, and contrarian to the dominant thinking in the industry (Felin and Zenger, 2017). Not only were the odds of creating a new market based on it less obvious, but for these reasons the theory was also less reliable. In the language of our model, this theory Θ was characterized by comparatively lower prior π and higher σ^2.

Consistently with our model, Del Vecchio experimented with the alternative theory (eyewear) to check whether it could become more plausible, rather than experimenting with the original theory (eye surgery) to check whether it could become less plausible.

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, Ferr‘e and Gucci. This provided a signal about the fact that states related to "eyewear" were becoming increasingly likely and potentially valuable.

Del Vecchio also conducted other relatively more precise (and costly) experiments, raising the bar of acceptance of the alternative theory. For example, the first licensing agreement with Armani (in 1988) and a few acquisitions of small sunglasses producers (Briko and Persol) were more stringent tests of the higher plausibility of the *eyewear* theory.

These experiments updated his prior on the alternative theory, raising it up to the point that it became more plausible.

Since then Luxottica's strategies changed significantly into accelerated external growth through licensing agreements with iconic fashion brands -like Armani, Bulgari, Chanel, Dolce & Gabbana, Prada and Versace-, the acquisition of large retail chains like Lenscrafters and Sunglass Hut and the acquisition of iconic brands like Persol, Ray Ban and, later, Oakley.

6 Conclusions

Our framework leaves several open questions for future research.

We do not provide a theory of how DMs 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 DMs and firms. Future research could develop the organizational implications of our framework, investigating how firms develop routines, design organizations and build management systems that allow the framework to be effectively, consistently and efficiently deployed across complex organizations. In addition, our framework has the potential to pull together multiple streams of strategy research, unify its language and provide a common ground to boost its rigor and impact.

Finally, our research program has the important implication for practitioners that they can lever many economic and management theories. DMs hardly use them 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, 1974.

Adner, R. & Levinthal, D. A. (2004). What Is Not a Real Option: Considering Boundaries for the Application of Real Options to Business Strategy. *Academy of Management Review*, 29(1), 74-85.

Agrawal, A., Gans, J. S., & Stern, S. (2021). Enabling Entrepreneurial Choice. *Management Science*, 67(9), 5510-5524.

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

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

Ansoff, H. I. & Brandenburg, R.G. (1971). A Language for Organization Design: Part I. *Management Science*, 17(12) B705–B716.

Arrow, K. J., & Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, 265-290.

Cai, Y. & Lim, L-H. (2022). Distances between Probability Distributions of Different Dimensions. *IEEE Transactions on Information Theory*, 68(6), 4020-4031

Cao, R., Koning, R. M., & Nanda, R. (2021). Biased sampling of early users and the direction of startup innovation (No. w28882). National Bureau of Economic Research.

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.

Camuffo, A., Gambardella, A., & Pignataro, A. (2022). Framing Strategic Decisions in the Digital World. *Strategic Management Review*, forthcoming

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., & Montrucchio, L. (2013a). Classical Subjective Expected Utility. *Proceedings of the National Academy of Sciences*, 110(17), 6754-6759.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., & Montrucchio, L. (2013b). Ambiguity and Robust Statistics. *Journal of Economic Theory*, 148(3), 974-1049.

Cerreia-Vioglio, S., Hansen, L. P., Maccheroni, F., & Marinacci, M. (2020). Making Decisions under Model Misspecification, preprint arXiv:2008.01071.

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.

Csaszar, 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.

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.

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

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

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

Gilboa, I. (2009). *Theory of Decision Under Uncertainty*, Cambridge University Press, Cambridge UK.

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. (2014). Nobel Lecture: Uncertainty Outside and Inside Economic Models. *Journal of Political Economy*, 122(5), 945–987.

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

Hofer, C. W. & Schendel, D. (1978). *Strategy Formulation: Analytical Concepts*. West Publ. Co., Chicago.

Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan, USA.

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A Flaw in Human Judgment*. Little Brown, New York.

Kahneman, D., Lovallo, D., & Sibony, O., (2019), A Structured Approach to Strategic Decisions, *MIT Sloan Management Review*, 60(3), 69-73.

Kerr, W. R., Nanda, R., & Rhodes-Kropf, M. (2014). Entrepreneurship as Experimentation. *Journal of Economic Perspectives*, 28(3), 25-48.

Koning, R., Hasan, S., & Chatterji, A. (2022). Experimentation and Start-up Performance: Evidence from A/B Testing. *Management Science*, forthcoming.

Lovallo, D. & Sibony, O. (2018). Broadening the Frame: How Behavioral Strategy Redefines Strategic Decisions. *Strategy Science*, 3(4), 658-667.

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) B403–B414.

McGrath, R. G. (2001). Exploratory learning, innovative capacity, and managerial oversight. *Academy of Management Journal*, 44(1), 118-131.

Milgrom, P., & Roberts, J. (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica: Journal of the Econometric Society*, 1255-1277.

Mintzberg, H. (1978). Patterns in Strategy Formation. *Management Science*, 24(9), 934-948.

Mintzberg, H. (1994). *Rise and Fall of Strategic Planning*. Simon and Schuster.

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

Nickerson, J. A. & 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.

Ocasio, W. (2011). Attention to Attention. *Organization Science*, 22(5), 1286-1296.

Posen, H. E., Leiblein, M. J., & Chen, J. S. (2018). Toward a behavioral theory of real options: Noisy signals, bias, and learning. *Strategic Management Journal*, 39(4), 1112-1138.

Raiffa, H. (1968). *Decision Analysis. Introductory Lectures on Choices under Uncertainty.* Reading, Mass.

Thomke, S. H. (2020). *Experimentation Works: The Surprising Power of Business Experiments.* Harvard Business School Press.

Trigeorgis, L., & Reuer, J. J. (2017). Real Options Theory in Strategic Management. *Strategic Management Journal*, 38(1), 42-63.

Van den Steen, E. (2017). A Formal Theory of Strategy. *Management Science*, 63(8), 2616-2636.

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

Appendix: Additional Cases

Mimoto

Mimoto is a start-up located in Milan that planned to launch a scooter-sharing service in early 2010s. (See also Spina, C. & Fronteddu, A., 2022. A Scientific Approach to Creating a New Business: MiMoto. INSEAD Case Study 6710.)

The default option was to launch the service and park the scooters in fairly populated areas of the city to capture demand. However, the founders thought they could do better. They figured that, unlike cars and bicycles, which are used by a vast share of the population, scooters face a more inelastic demand. Some people like them, while others will never use them. Thus, they conjectured that first they had to find the right target customers, and then identify locations where to park scooters in which they are more likely to find the target customers.

The focus on target customers provides the basis to identify the key attributes. For instance, Mimoto's founders excluded that they faced uncertainty about models of payment. Car- and bicycle-sharing had already tested payment apps that were easily adaptable to the scooter-sharing business.

Their theory was that ideal target customers have three features: they have to be *young*, with *ability to pay*, and *mobility needs*. This translates into the following space of attributes:

$$\mathcal{X}_y = \{yes, no\};\ \mathcal{X}_{atp} = \{yes, no\};\ \mathcal{X}_{mn} = \{yes, no\} \qquad \mathcal{X} = \mathcal{X}_y \times \mathcal{X}_{atp} \times \mathcal{X}_{mn}$$

Since all three conditions have to apply (*conjunction*), they partitioned $\mathcal{X}$ in *two subsets*:

$$x = \{(\text{y, y, y})\} \qquad \mathcal{X} \setminus x = \text{all other 7 subsets}$$

Based on this logic, they identified *college students* (CS) as their ideal target customer because they satisfy all three criteria.

To outline the decision problem that would follow this exploration phase, consider this linear demand

$$Q_z = \gamma_z - \gamma \cdot P + \varepsilon_z$$

where Q_z is demand in a given period of time (day) in a given location ("zone") z of the city, $\gamma_z = \gamma_{CS} CS_z + \gamma_{OT} OT_z$, CS_z and OT_z are, respectively, the number of college students and all other customers in location z, γ_{CS} and γ_{OT} are weights of demand by CS and OT, P is price, γ is the impact of price change, and ε_z is a random stochastic term.

The parameter γ_{OT} is the benchmark from the family of models P_0 (with prior π_0) about the use of scooters by any generic individual in the population. The alternative family of models P_1 (with prior π_1) is about γ_{CS}, and the theory of this alternative family of models is that $\gamma_{CS} > \gamma_{OT}$. Note that this is equivalent to theorizing that demand for the service is more likely when all three conditions apply, i.e. when $x = \{(y, y, y)\}$: in this case, college students represent the case of demand that corresponds to the subset $x = \{(y, y, y)\}$ of the space of attributes X.

Mimoto's founders run experiments to update their expected value of γ_{CS} which is the model that informs them about their prior π_1 on the theory that ideal target customers have the three features discussed above. The logic of the experiment is that customers have to find scooters nearby when they need them, which implies that in order to test hypothesis about γ_{CS} founders have to compare the use of scooters located near colleges, where they are more likely to find students, with the use of scooters parked randomly in the city. If $\gamma_{CS} > \gamma_{OT}$, there has to be higher usage of scooters near colleges.

The experiment failed to show that parking near colleges raise demand. This prompted Mimoto's founders to go back to theory to find a new model. They figured out that the key attribute is not *mobility per sè*. College students have regular class schedules that make their mobility predictable. This implies that they can plan their mobility in advance, and therefore the scooter service suffers from competition from private and public transportation. For instance, students can check public transportation schedules and plan accordingly, or they get experience about private parking near universities, such as knowing where to park at what time or subscribing to private parkings.

They revised their theory arguing that ideal target customers have to have these three features altogether: young, ability to pay, *unpredictable* mobility needs. This led to a revised space of attributes

$$\mathcal{X}_y = \{yes, not\};\ \mathcal{X}_{atp} = \{yes, not\};\ \mathcal{X}_{umn} = \{yes, not\} \qquad \mathcal{X} = \mathcal{X}_y \times \mathcal{X}_{atp} \times \mathcal{X}_{umn}$$

Since all three conditions have to apply, partition $\mathcal{X}$ in *two subsets*:

$$x = \{(y, y, y)\} \qquad \mathcal{X} \setminus x = \text{all other 7 subsets}$$

Now founders realize that the ideal target customers are *young professionals* (YP) who move unpredictably during working hours, which changes the model specification of demand by location to

$$\gamma_z = \gamma_{YP} YP_z + \gamma_{OT} OT_z$$

where YP_z and OT_z are the number of YP and "others" (OT) in $z \in Z$. The uncertainty is about γ_{YP}, and, in line with what we did for γ_{CS}, Mimoto's founders now want to collect evidence on whether $\gamma_{YP} > \gamma_{OT}$. The default theory is still that there is no ideal target customer, and scooters can be parked randomly wherever customers leave them.

The design of the new experiments on γ_{YP} was similar to the previous one. Mimoto's founders re-parked used scooters downtown, where it is more likely to find young professionals, and compared usage with a random allocation of scooters in the city. The experiment showed that parking downtown yields higher average usage than the standard option.

Mimoto set on this theory ($\gamma_{YP} > \gamma_{OT}$), which prompted them to solve a standard decision problem about actions. Specifically, the actions are now choices of the number of scooters x_z to be parked in different locations $z \in Z$ of the city, and the price of the service. Assuming monopolistically competitive conditions, Mimoto chose these actions by solving the following problem

$$\max_{\{P, x_z\}} \mathbb{E} \sum_{z \in Z} P \ min\left(Q_z, x_z\right) - c\left(x_z\right)$$

where $c(x_z)$ is the cost of parking x_z scooters in location z, Q_z is the linear demand defined earlier with $\gamma_z = \gamma_{YP} YP_z + \gamma_{OT} OT_z$, and apart from a random error the expectation is over $\gamma_{YP} > \gamma_{OT}$, which is the theory on which Mimoto's founders committed after the experiment.

In this program we assume that γ_{OT}, YP_z, OT_z, $z \in Z$, are given. Models $\gamma_{YP} \in \Gamma_{YP} = \{\gamma_{YP} : \gamma_{YP} > \gamma_{OT}\}$ are equivalent to $\theta_1 \in \Theta_1$ in the text, and the likelihood $\mu(\gamma_{YP})$ is equivalent to $\mu(\theta_1)$. The program yields optimal actions P and x_z as functions of the theory Γ_{YP} (and therefore likelihoods $\mu(\gamma_{YP})$), γ_{OT}, YP_z, and OT_z, $z \in Z$. Given γ_{OT}, YP_z, OT_z, if founders estimate a higher probability of higher γ_{YP} (that is, they find evidence about models with higher γ_{YP}), they increase their optimal choice of x_z in locations with higher YP_z, and expect higher returns to higher x_z in these locations.

Mimoto succesfully implemented the strategy that we described in this section. Over the past decade it succesfully expanded in Milan, and then in similar cities, Turin and Genoa. In 2020 Mimoto was succesfully acquired by Helbiz a larger international smart mobility company.

Li Jalantuùmene ("The Gentlemen")

Li Jalantuùmene ("The Gentlemen") is a high-end restaurant opened in the early 2000s by the chef Gegè Mangano in Monte Sant'Angelo in the Apulia region in Italy. Monte Sant'Angelo is a religious site that grew large scale religious tourism and related restaurant and hospitality services.

Gegè focused on an alternative strategy: high-end cuisine. He developed and tested the theory that exclusive dining experiences and, hence, perception of scarcity raise demand for high-end restoration. He worked with two attributes: *scarcity* $\mathcal{X}_s = \{yes, no\}$ and *demand* $\mathcal{X}_d = \{high, low\}$. The space of attributes is $\mathcal{X} = \mathcal{X}_s \times \mathcal{X}_d = \{(y,h), (y,l), (n,h), (n,l)\}$. A high probability of (y,l) falsifies the *implication* that scarcity raises demand.

His experiment was to reject, for some time, a sizable fraction of calls for reservations by claiming that he was fully booked. The logic of the test was that rejections raise the sense of scarcity, and therefore, if scarcity matters, rejections raise the number of future calls, which is what he observed. The outcome of the experiment allowed him to update downward the expected probability of (y, l) and raise his prior about his theory that scarcity raises demand. He then pursued his high-end cuisine project rather than the standard mass-tourism service. Today Li Jalantuùmene is a successful business.