Miscellanea

Modeling ignorance is at the heart of economic modeling. This chapter attempts to explain how current modeling techniques and interpretations have been shaped by the desire to quantify ignorance, tracing back the various intellectual influences behind these techniques. This book has proposed an alternative way to think of, and model, ignorance that does not require agents to precisely know the contours of their ignorance. We discuss some of the recurrent themes in decision theory and game theory, explaining how our perspective complements or competes with the more traditional views.

1. *Making precise what is not known.*

To many economists, the Pavlovian response to ignorance consists of *making precise what is not known*. If one is ignorant about the number of balls in an urn, then one should think of the range of possible numbers and put a probability distribution over them. This distribution is a probabilistic belief, meant to be a precise representation and quantification of one's ignorance. Probabilistic beliefs play a central role in modeling ignorance – they are viewed as primitives in many models.

Much of this book is about avoiding probabilistic beliefs in decision making, asking instead that one focus on the perceptions that agents might plausibly get, and on what they can plausibly make of them. This does not mean that we dismiss beliefs altogether. Our view is that beliefs are often more casual than the way most theory portrays them to be, and that by avoiding a fine and detailed representation of what is not known, models may not only be more realistic but also more parsimonious.

2. *Risk and uncertainty.*

The precise quantification of ignorance is related to the old risk and uncertainty debate. Following Frank Knight,[1] one usually makes a distinction between

[1] Knight (1921).

risk, akin to a throw of a die, and *uncertainty*, which cannot be measured. In a literature beginning with Ramsey (1926) and culminating with Savage (1953), it became legitimate to think of any uncertainty as quantifiable. The logic is, essentially, that choices and bets can be made even in situations of ignorance, thus these choices must reflect or reveal some personal measure of likelihoods.

Imagine that you face a magician's hat with an unknown amount of money under it. You are offered a choice between getting \$5 or getting the amount under the hat. Choosing the second option might reveal something about your perception of what's hidden (possibly along with the extent to which the contents of a magician's hat piques your curiosity.)

Of course, this single choice cannot reveal anything precise about your perception. The precise quantification comes about when one assumes that you are (most often fictitiously) confronted with other choice situations that are variations of the original one – getting $X = \$5$, or getting whatever amount under the hat that is in excess of $Y = \$50$. Varying X and Y could, in principle, permit one to elicit a precise probabilistic belief that you might hold about the hidden amount. Or at least, if your answers are consistent in a way defined by Savage, your choices can be understood as though you had a precise probability distribution in mind.

This book is not about trying to determine or elicit what is your mind when you see a particular hat. The logic of our exercise is to try to find patterns of behavior or systematic effects that apply *across* choice or strategic situations. For example, many choice problems are akin to the hat experiment: one option is easy to evaluate, while the other has uncertain consequences. When the stakes are high (higher than \$5), the perception that there is a qualitative asymmetry between the two options might trigger some systematic cautiousness in behavior that applies across similar problems where that perception is triggered.

3. *Isolated versus representative problems.*

Savage has also shaped the way many theorists think about decision making in situations of ignorance, allowing us to transport consumer theory to such situations. In essence, Savage provided conditions under which analysts may approach decisions under ignorance in the same way they do for decisions with known (and easy to evaluate) consequences. One typically represents decision making as if each alternative was evaluated *in isolation*: that is, *independently of other qualitatively similar problems faced*, with the evaluations then compared.

This book takes a different perspective: decision problems are never considered in isolation. Rather, we emphasize that decisions be viewed as byproducts of a rule of behavior or strategy that applies to a pool of similar decision problems.

Why?

Our view is that many of our choices build on broadly defined perceptions that are neither alternative-specific nor problem-specific.

One sometimes feels that a strange deal is being offered, or that the deal seems too good to be true. We recognize that feeling because we have been confronted with it many times, and we have learned (or been told) to exert greater caution in such situations.

Similarly, we sometimes face choice problems in which one alternative seems easier to evaluate than the other. If one also perceives that the stakes are high and that the decision is not trivial, the perception that one alternative is simpler to evaluate may play a key role, prompting us to opt for the simpler to evaluate alternative.

Perceptions are cues that guide our choices. They acquire relevance (and value) from the numerous instances in which they come to mind, and the benefit that we draw from their use, given the set of instances in which they surface, and our various responses to them.

Said differently, a choice problem offered to us is accompanied by perceptions. But what we make of these perceptions cannot be independent of the various choice problems in which similar perceptions are triggered. Among other things, perceptions create correlations in the way decisions are made across the problems in which given perceptions crop up. What we make of perceptions is unlikely to be problem-specific.

So, whereas standard theory often portrays each decision problem in isolation, independently of other decision problems, our perspective is that perceptions create links across decision problems, common responses, and that one cannot analyze a choice problem without considering: (i) the perceptions that are triggered and (ii) the various choice problems in which similar perceptions would be triggered. This book is an attempt to follow this path, *taking as primitive the class of situations considered*, and modeling it as a (sufficiently rich) joint distribution over problems and perceptions.

Of course, the classic perspective has an advantage. Each perception or signal defines a distinct problem that can be analyzed separately from other problems without questioning how other choice problems might affect the current one. The drawback, however, is that we miss the big picture. We may get buried in second-order details and fail to identify the systematic effect that perceptions have on choices, missing their role in helping us simplify and handle the extraordinarily vast array of problems that we face and that never exactly repeat themselves.

4. *Mistaken or cue-driven behavior.*

Out of numerous candidates who participated in a contest, you have been selected as the winner, with a cash prize of 1.5 million dollars. Upon collecting your check, you are offered the choice of getting instead S million dollars, where $S = \sqrt[4]{8}$. That number might look attractive, yet the second number looks a bit like those that you see on those insurance contracts that you don't read, and that often come with unpleasant surprises. Not seeing an obvious comparison between the two numbers, you opt for the simpler one.

We used this example in Chapter 21 to illustrate how perceptions of complexity may guide our choices. Whatever impression one gets, if any, concerning the value ranking of the two options, other impressions compete and plausibly dominate: the inability to make a definitive comparison, and the feeling that stakes are high and that one option is simpler.

Complexity perception is a cue that one uses to guide choices. For a particular choice problem taken in isolation we might indeed realize that we could have been better off not using the cue – in our story, $\sqrt[4]{8}$ is actually larger than 1.5. Yet a cue may remain valuable in spite of the adverse consequence that it has for one specific choice problem. A cue is like a signal, and you cannot ask that decisions based on signals be perfect every time. A cue/signal generates decisions that are on average better than if you did not use it; and instances in which the cue fails to produce a good outcome do not necessarily constitute a reason to abandon it.

Finding the "right" lens. Analysts often differ in the lens through which they approach problems. Much of decision theory focuses on very precise and detailed examples, in which the consequence of each of the available alternatives can be objectively measured by the analyst. Analysts often presume that since all the characteristics of the choice problem are equally available to both the agent and the analyst, one can assume that the evaluation of each alternative poses no difficulty to the agent. Consequently, choices that are inconsistent with the objectively measurable characteristics of the choice problem are interpreted as mistakes, errors, or biases. In this respect, if the analyst focuses on one particular choice problem (1.5 or S), the choice of 1.5 over S looks like a mistake.

From a broader perspective, however, one can take the position that choices result from perceptions-based strategies which apply *across a broad set of problems*. From this perspective, the choice of 1.5 over S is not an error, despite the fact that, from the analyst's perspective, S is objectively larger than 1.5. It does not reveal a preference for 1.5 million dollars over S million dollars. It might simply reveal that forgoing a large sum for an unknown amount is not worth the risk.

5. *Subjective and objective views.*

To many theorists, beliefs are taken to be a subjective or personal representation of what an agent is ignorant about. If a seller is ignorant about the value of the transaction to a given buyer, one may represent his decision process as though he had in mind a probability distribution over possible values.

A potential problem with this approach is that in general, there is nothing that necessarily ties the agent's beliefs to the real-world consequences of the choices he makes, and possibly little or no discipline on what could be "reasonable" beliefs. The seller's belief about the buyer's valuation could bear no relationship to the actual valuation of the buyer. Trade might be missed, but only because the seller holds implausible beliefs.

In practice, papers often proceed by assuming that:

(i) There is a distribution ω that determines the buyer's valuation v and a signal z received by the seller.
(ii) The seller's belief is then determined for each signal z, as if ω was known and the seller was making an inference from the signal z.

Papers are often ambiguous as to whether one should interpret ω as a "subjective" or an "objective" distribution, as even subjective but "reasonable" distributions may generate interesting insights. For economic problems in which one wishes to quantify real consequences, the "objective" interpretation tends to be favored. In that case, the seller is in the same position as that of a poker player who tries to assess the strength of his opponent's hand (five cards), based on seeing only three of his opponent's cards. Of course, buyers and sellers are not playing a card game, and nobody is distributing cards according to a well-defined process. But we represent the buyer-seller relationship as if there were an objective lottery selecting a valuation v to the buyer, and a signal z to the seller.

The first step, with the definition of an "objective" distribution ω, is a somewhat artificial theoretical construct. Yet we do not fundamentally object to it, as it provides a convenient way to tie perceptions with real consequences. Hunger is correlated with depleting reserves, and fear is correlated with danger. It seems fine to take an omniscient outsider's perspective to define these joint distributions. Our interpretation is that ω represents the typical situation (z, v) that the seller faces.

The second step is more problematic. It presumes that this distribution is known and it allows agents to take full advantage of this knowledge. This book aims to amend that second step, by reducing the agent's ability to take advantage of the structure of the model. He does not know ω, nor can he think about all the consequences of that knowledge. An agent compares strategies given their average performance across the problems that he confronts. His knowledge is limited to these comparisons, and the strategies that he compares are a primitive of the model. His knowledge of ω is implicit, not explicit.

Our perspective is that linking perceptions and real consequences is a useful discipline. We suggest keeping this link as a primitive assumption, without necessarily assuming that agents know precisely this link, or behave as if they knew them precisely. Said differently, this book has been an attempt to separate the question of tying perceptions to real consequences, and the question of what people know of these ties. What agents end up knowing is limited to their experience in using that the perceptions they get – no less and no more.

6. *Subjective and objective views in games.*

Harsanyi made a fundamental contribution to the analysis of games. He provided a relatively parsimonious way to think about beliefs about what others know in a game situation, proposing that we represent the ignorance of each player by a belief derived from a joint distribution over types. For example, to model an auction, we often define a joint distribution over values $f(v_1, \ldots, v_n)$. Each possible value v_i is a possible type for agent i, and to each agent i with type v_i, one may associate the conditional probability distribution $f(\cdot \mid v_i)$. This conditional distribution is a belief that represents the agent's ignorance when his type is v_i. The construction is particularly flexible, as a type can include any thought process that the agent might have. Calling z_i this thought process, one may consider a more elaborate model in which we define a joint distribution $f(z_1, v_1, \ldots, z_n, v_n)$. The pair $t_i = (z_i, v_i)$ is a type, and to each agent i and type t_i, one may associate the conditional probability distribution $f(. \mid t_i)$ (over other agents' types). This distribution represents the agent's ignorance when his type is t_i.

How should one interpret the distribution f? One could hold the subjective view that these distributions and beliefs are only meant to be a representation of the agents' ignorance, and not necessarily tied to any "objective" auction situation. Another view is that the distribution is meant to reflect a typical auction situation, as if values v and perceptions/signals z were drawn from an objective lottery.

This book favors the second interpretation. Essentially, if we want to talk about efficiency in auctions, we should tie what agents perceive or know to real consequences. In contrast to the literature however, we have taken these ties to be a primitive of the model *without necessarily assuming that agents are able to fully exploit them*.

7. *Belief hierarchies and common knowledge.*

When playing poker, I may try to assess the strength of the other player's hand, But this assessment seems insufficient. The other player's decision to fold or not probably depends on his own assessment of my strength. So my assessment of his assessment of my own strength may thus be relevant.

In the spirit of making precise what is not known, theorists are often tempted to model these assessments as probabilistic beliefs. One thus defines for each player a belief about the other's strength, and a belief over the belief that the other holds over my strength – and so on indefinitely. The collection of these beliefs is a rather abstract object, which can be defined for each player. Such a collection is called a belief hierarchy.

One virtue of Harsanyi's construction mentioned above is that it avoids entering these complex constructions. The analyst defines a set of types for each player (with each type possibly representing one admissible belief hierarchy), and a joint distribution over types. This joint distribution over types is difficult to motivate, especially if one holds the view that a type should be interpreted as a belief hierarchy. In addition, it is assumed that each player knows the distribution, knows that others know, and so on indefinitely. In brief, that this joint distribution is commonly known, or common knowledge.

This common knowledge assumption has been criticized by many. Nevertheless, the general view is that Harsanyi's construction is a useful shortcut to analyze games, and that this shortcut provides a useful and tractable approximation to the "real" problem that a strategic player faces. For example, Morris and Shin write:[2]

"*In principle, optimal strategic behavior* ***should***[3] *be analyzed in the space of all infinite belief hierarchies.*"

In other words, if we were not constrained by our mathematical abilities, it would seem like a good idea to capture more accurately strategic behavior and understand how each infinite belief hierarchy affects behavior.

We hold a different view. Our perspective is that one should focus on: (i) perceptions that agents can plausibly get; and (ii) perceptions for which players can understand or learn what *use* they can make of them. Belief hierarchy is problematic on both grounds. Thoughts about others may be relevant in some games, but are often too vague to be represented as a precise belief hierarchy. Furthermore, one suspects that these thoughts seldom provide good guidance on how to behave, and may be no more than a recipe for generating random decisions.[4]

8. *Vagueness.*

Savage himself acknowledged a difficulty with postulates that imply a precise representation of one's ignorance, a difficulty stemming from the vagueness typically associated with judgments:[5]

"*The postulates of personal probability imply that I can determine, to any degree of accuracy whatsoever, the probability [for me] that the next president will be a Democrat. Now it is manifest that I cannot determine that number with great accuracy, but only roughly.*"

Judgments are typically vague, and a theory based on precise probabilistic beliefs seems, at the very least, descriptively inaccurate.

The statement also reveals a common shift in interpretation, reinforced by the terminology used. Through the agent's numerous answers to our hat experiment, an outsider could in principle elicit a probability distribution over the monetary rewards hidden under the hat. Theorists often view this distribution as a *representation* (made by the analyst) of the uncertainty faced by the agent, rather than a personal assessment (made by the agent) of that

[2] Morris and Shin (2003, Chapter 3, page 56).
[3] Our emphasis.
[4] This does not mean that we want to exclude sophisticated thinking from models. On the contrary, as suggested in Chapter 4, our view is that strategic thinking potentially shapes the signals that agents pay attention to, as well as the family of behavioral responses that are considered by the agent (see also Chapters 15 and 20).
[5] Savage (1953, page 59).

uncertainty. However, the statement above and the terminology "personal," "belief," "subjective" suggest the latter interpretation, something attached to the agent, rather than the way an outsider might represent the agent's decisions.

A convenient way to circumvent the difficulty has been for theorists to say that agents behave *as if* they held such beliefs, maximizing expected utility given these "as-if beliefs." By and large, however, and despite the cautionary "as-if" qualification, much of the literature treats beliefs as real objects, as if they were actual ingredients of the agent's decision process. An agent's behavior is then often described as a complex mapping between an overly detailed specification of beliefs and a decision, with the hope that this mapping provides intuition – and a foundation, for the behavioral predictions of the model.

This book has been an attempt to avoid endowing agents with detailed representations of uncertainty, unless we have reasons to believe that the representation has relevance in actual decision making. In modeling a repeated interaction, we endowed agents with only two belief states regarding the current state of the relationship (good or bad). Of course, the entire history of play and observations made by the other player influences his current state, hence his current behavior. One player could thus be tempted to calculate more precisely a probability distribution over the other's history, based on his own history of play and observations. We deliberately ignored such elaborate thoughts and focused on comparing simple behavioral rules, each characterized (for example) by one's inclination to spontaneously change belief state from bad to good.

9. *Qualitative insights.*

Modeling vagueness, however, may not be the main challenge, and possibly not even a modeling objective. In the repeated interaction mentioned above, the modeling challenge is to clarify the strategic issue faced by each individual. A theory that results in a complex mathematical object that defines a mapping between the set of possible histories and decisions does not help develop intuition. A theory that focuses on one aspect of behavior, say one's inclination to forgive in a relationship, and that explains the tradeoffs associated with that inclination, is less ambitious, but it opens the door to qualitative statements which may be useful in shaping one's intuition.

In a criticism of quantitative economics, Herbert Simon wrote:[6]

"By 'anticipating the future' I do not mean estimating joint probability distributions, for the most important kind of futurology is to anticipate, qualitatively more than quantitatively, changes in the important dimensions of the space in which the firm will operate."

The claim is that most often, firms do not care about getting the precise details of what may be coming, but rather, the big picture – a broad perspective

[6] Herbert Simon (1993, page 135).

on the nature and significant changes to expect. The same applies to many readers of economic models. Readers care for the punch line, the first order effect, the broad statements that capture the essence of a strategic phenomenon.

There is a difficulty with mathematical precision in that it often comes with detailed descriptions, which make it difficult to make qualitative statements and to disentangle the many forces that may affect behavior. What this book has proposed is to cut through the vast array of possible behavioral responses, and attempt to identify the strength of *a priori* selected responses. Our view is that often there is no need to look at large strategy sets to get the punch line; Worse without restriction, we may end up being unable to grasp or appreciate the relevant strategic forces.

We consider this restriction to a few dimensions a useful modeling challenge, a disciplining device that forces one to think *a priori* of the relevant strategic dimensions, before determining whether, indeed, that restriction has the relevance that we anticipate.

10. *Faith in beliefs.*

Another implication of the standard representation exercise is that it portrays agents who behave as if they had complete faith in their beliefs. So long as this remains an "as-if" statement, one is not assuming that agents *do* have faith or confidence in their beliefs. But once one slips into a more literal understanding (in which one would attach actual beliefs to agents), the implication (that agents have faith and confidence in their beliefs) is more questionable. Paraphrasing Savage, I could attach a probability 61 percent that the next president will be a democrat, and a probability 32 percent that he wins by a 3 percent margin at least – but the degree to which I would have confidence in these numbers would seem limited.

As a response to this discomfort, there have been attempts to weaken Savage's postulates, leading to representations of uncertainty that would not have the agent behave as though he had complete faith in a particular distribution. One such attempt leads to the maxmin path, or representations of behavior in which the agent behaves as if he had a set of beliefs in mind.

These more elaborate representations are similarly problematic: (i) they push even further the need to describe in detail what is not known, now relying on the precise description of a *set* of probability distributions (rather than a single probability distribution); (ii) it is still not clear whether the representation is meant to be the analyst's construction (with agents behaving *as if* they had multiple beliefs in mind), or as an actual description of the agent's perception (a description that would attempt to take seriously the idea one cannot be certain of his own belief); (iii) they remain silent on the connection between the agent's perception and the actual situation faced.

11. *The Bayesian route.*

Economic models are mathematical objects that come with precisely defined model parameters. For example, if one wishes to model the relationship between observations and underlying preferences, one defines a joint probability distribution ω over observations and states (see Part I). That distribution ω *is* a model parameter. In solving models, we presume that agents behave as if they knew the model parameters perfectly (including ω). Part II of this book has explained why sometimes we derive insights that hinge (too much, in our view) on the agents' ability to tailor one's behavior to the model parameters.

We have proposed in Part III a way out of this difficulty. We argued that with direct strategy restrictions, one could prevent an overly fine-tuned adjustment to model parameters.

An alternative path consists of taking the classic Bayesian route. If we think that agents are better modeled as being unaware of model parameters, then one should model that ignorance, and make precise, or quantify, the ignorance that agents face regarding model parameters.

For example, in modeling a (common-value) auction, one could assume that the n estimates of the common value v are drawn from a joint distribution f. For example:

$$z_i = v\theta_i$$

where v and θ_i are independent lognormal distributions. In specifying the basic model, we proposed a particular distribution with $\log \theta_i$ distributed according to $\mathcal{N}(0,\sigma^2)$ for some fixed σ.

If one objects, as we did, to the idea that many agents can simultaneously adjust to a particular specification of σ and n, one may assume that σ and n are random variables. This is the route that we took in Chapter 16 on information aggregation. In that problem, fixed σ and n implies that the estimation error of the k^{th} most optimistic estimate is almost entirely determined by k/n. This lack of variability of the k^{th} most optimistic estimate is an artifact of the basic modeling assumption (in which n and σ are fixed), and adding noise to the dispersion of estimates avoids this implausible lack of variability. The Bayesian route can thus sometimes be a useful complement to ours.

Part III of this book, however, suggests that it is often not necessary to introduce artificial randomness in the modeling parameters to prevent players from exploiting the structure of the model. Direct strategy restrictions can do the job in a more parsimonious way.[7]

[7] In addition, the Bayesian route is a device that only produces a modification in ω. In the auction example, randomness in the dispersion of values σ only produces a modification of the joint distribution over estimates. This modification may be helpful to check the robustness of initial conclusions (as in information aggregation problems), but it does not modify the basic

12. *Wilson's critique and the classic "robustness" route.*

The idea that the conclusions of our models hinge on players' assumed ability to exploit the structure of the model is related to Wilson's critique.[8]

To address that issue, many advocate the robustness route, which can be portrayed as an enrichment of the original model. In essence, it consists of: (i) allowing the original model parameter to vary (this is the Bayesian route mentioned earlier); and (ii) assuming that players are possibly differentially informed about these variations. In the language of the above auction model, this might mean, for example, that; (i) the dispersion parameter σ becomes a random variable; and (ii) each agent gets a signal σ_i correlated with σ.

There are several difficulties with this route. First, it often leads to intractable models. In the context of the above enriched auction model, this leads to strategies that are two dimensional (bids are functions of value v_i and signal σ_i), with little hope of obtaining a characterization of equilibrium strategies.

Second, the more elaborate model is subject to the same critique as the original, with strategies that potentially exploit finely the structure of the new model. One could obtain an apparent lack of robustness in the original model, simply because one introduces a peculiar signal structure that, when properly exploited, would lead to some unraveling that destroys the original equilibrium. Which setup should one consider as lack robustness? The more complex construction or the original one?[9]

Third, our view is that the main role of the additional signal structure is to introduce artificial randomness in behavior, and we see no reason to endogenize this randomness through such an elaborate mechanism, whereby we simultaneously add an artificial signal structure and request that players behave optimally conditional on each signal realization. There are many plausible sources of randomness and it isn't clear why one should necessarily endogenize them. One can instead circumvent the elaborate construction and directly assume that players inevitably make errors when taking decisions, and evaluate how the general shape or magnitude of these errors affects the conclusions of the original model (see Chapters 19 and 20).

assumption that is typically made – namely, that agents are able to tailor their behavior (i.e., the bid function $b_i(v_i)$) to that new joint distribution.

[8] Wilson (1987, page 34) writes, "Game theory has a great advantage in explicitly analyzing the consequences of trading rules that presumably are really common knowledge; it is deficient to the extent it assumes other features to be common knowledge, such as one agent's probability assessment about another's preferences or information. I foresee the progress of game theory as depending on successive reductions in the base of common knowledge required to conduct useful analyses of practical problems."

[9] Many of the constructions inspired by the e-mail game (Rubinstein 1989) rely on a special signal structure. Our view is that, sometimes, it is more the agent's fine ability to exploit that special signal structure that one should question, rather than the initial equilibrium prediction.

13. *Information.*

Economic theory has accustomed us to think of signals as information, with additional signals meaning more information. Throughout this book, we have avoided equating these. Our view is that a signal per se is not helpful. An agent has to determine what use he can make of it, and to be valuable, he must also determine whether he is better off utilizing the signal, rather than ignoring it.

In standard models, this ability comes without cost because one assumes that agents behave as if they can freely exploit the structure of the model. A consequence is that "information/signals" can never hurt in standard decision problems. If an agent is provided with an irrelevant signal, he understands that it is irrelevant, and that he should ignore it. More generally, an agent provided with several signals also understands the relative weight he should attribute to each signal. Information aggregation is never an issue nor an obstacle in these models. Both claims seem at odds with common sense and experiments (Gigerenzer, 2007, page 37).[10]

Our perspective is that one often uses "information" to refer to "raw data," and that raw data seldom comes with a recipe how the data might be used. In real life, richer data comes with questions as to how it might be used or aggregated. Getting richer data is potentially harmful because it creates opportunities to use it in inappropriate ways, hence more opportunities for mistakes. In models, more data cannot hurt because one can always ignore the additional data. But the difficulty is precisely in determining which part of the data should be ignored.

Rather than referring to information as "raw data," we prefer to think of information as the recognition by the agent that some ways of using the data are better than others. Using the language of Gigerenzer (page 60), agents have tools, i.e., a number of plausible behavioral responses to signals or perceptions, each of which can be thought of as an instrument adapted to some (and only some) of the problems he faces.

Information, thus, relates to both with the toolbox (i.e., the set of tools that one is endowed with) and the ability to identify or recognize the most appropriate tool given the environment faced. Our concern is that in many models, the toolbox grows without bound, to a degree that the recognition requirement becomes farfetched. As a consequence, much of the theoretical work in this book involves the analyst being able to: (i) define a plausible toolbox for each broad category of problem; and (ii) highlight interesting connections between properties of the environment faced by individuals and the appropriate tools used within the toolbox.

[10] Gigerenzer for example suggests that having less time to decide forces you to use intuition, while having more time enables you to include more data in your decision process, at the risk of being dragged into decisions driven by irrelevant details.

14. *Rationality.*

How does (or should) a rational agent behave? What does rationality mean? There is little reference to these questions in this book. Our aim has been to describe models that portray the behavior of agents who attempt to behave according to their best interest, and to provide qualitative insights concerning the consequences that this self-interest generates.

For agents, behaving according to one's best interest is a challenge. Situations never exactly repeat themselves. This generates gaps between the situation at hand and the agent's perception of it, and these gaps create discrepancies between the decisions made and the optimal decisions. Whatever behavior that looks attractive beforehand may look less so ex post. We think of rationality as an economic force that tends to reduce these gaps and discrepancies.

For the analyst, the challenge as we see it is to provide a tractable model that seems rich enough to make the optimal decision non-obvious, and yet not so complex that the economic forces that we attempt to characterize cannot be spelled out or captured in a parsimonious way. The goal of modeling is finding the "right" balance between introducing inadequate perceptions (too coarse, too noisy) on one side, and means of countervailing these inadequate perceptions (through the comparison of various instruments or strategies, that is, various ways of handling these perceptions).

In the models we consider, as in standard models, all players are "rational," in the sense that they use the instruments available (i.e., the strategies that they are endowed with) as best as they can. Whether we model a baby getting internal signals correlated to the states of his reserves and "transforming" these signals with more or less delay into a cry for help, or whether we model a chess player getting a reasoned perception about the strength of his position and "transforming" this signal into a safe or a bold move, we think of agents doing the best they can with the tools they have. Across models, players differ only in the degree to which they manage/are allowed to adjust to the specific environment they face.

Throughout the book, our thrust has been:

(i) Sometimes the balance is not right (Part II). We may be giving too many instruments to players, enabling them to undo the mistakes/noise/errors that we introduce in the first place, through strategies that one cannot expect them to play.

(ii) Sometimes the balance is easier to achieve by directly restricting the set of instruments or by adding noise to their instruments (Part III).

Note that this leaves aside the question of how people manage to learn which instrument works best for them. Rather than embedding the analysis into a more complex model, we prefer our shortcut approach. If the analyst believes that learning is unrealistically difficult (either because the set of instrument is too rich, or because the model implies behavior that depends too much on the particular model specification – as in Chapter 12), one can either reduce further the set of instruments or add noise to reduce further the power of these instruments, thereby implicitly limiting further the agent's ability to adjust to the environment.

15. *Focus.*

What do we wish to explain? What do we wish to endogenize? What aspect of behavior do we want to focus on? Focus is at the heart of any modeling enterprise, and any modeling exercise must define what is exogenous, and what is to be endogenized.

In this, the analyst faces two challenges. The first stems from the fact that a simpler environment facilitates focus, yet a simpler environment also facilitates the agent's ability to adjust to the environment and to one another, sometimes implausibly. The consequence is that some models may teach us more about the inner working of our modeling tools than about the economic forces at work. This book (in particular, Chapters 19 and 20) illustrates how one can limit players' ability to exploit the structure of the model, directly introducing noise in behavior that reduces the scope for unraveling and for coordination. It also illustrates that a more complex environment may sometimes be more enlightening (as in Chapter 16 on information aggregation).

Another challenge is to resist the temptation to endogenize many dimensions of behavior simultaneously. This temptation is a natural one, as this seems to be a path to greater generality. What we often get, however, is complex behavior in which first and second-order effects are entangled. What we also often get is an illusion of generality, with players' behavior finely tuned to one another, or to special and artificial features of the model. This book has taken the perspective that limiting the dimensions of strategic behavior under scrutiny is a useful disciplining device.

To illustrate, the standard repeated game framework provides a very flexible tool to analyze social interactions. The framework is an austere world in which any behavioral response to past histories is feasible. In this cold-blooded world, some highly unusual behavior may come to be optimal only because it fares well when played against others' highly unusual behavior. However, the model typically fails to explain how players might come to play these strategies in the first place. In addition, "equilibrium strategies" are abstract mappings that are often difficult to interpret because they lack the *a priori* structure that would facilitate the interpretation in common language. If we wish to endogenize the degree to which one forgives after being upset, we should define *a priori* how we capture that. If we wish to endogenize the extent to which feelings of injustice upset us, we should also translate this into a behavioral rule (and also consider whether the repeated game we analyze is the appropriate vehicle for endogenizing these feelings). Absent an *a priori* structure, we are in the same position as an econometrician who finds a clever but extraordinarily complex rule to describe the process that generates the data, but cannot interpret the process. Meaning arises from pooling histories and testing whether regularities emerge across them.

To summarize, our answer has been that it is useful to propose a class of candidate strategies. It forces one to think beforehand of a class that is plausible within the environment studied, or that reflects how agents plausibly comprehend the environment, even if this implies an ad hoc restriction on the set of histories of the relationship that makes one upset. It forces us to focus on one, or few, behavioral aspects of the interaction, whether one should react harshly to being upset, leaving aside aspects of behavior driven by other considerations (e.g., whether one should feel upset when betrayed).

16. *Bounded rationality or limited sophistication?*

Our restrictions may be viewed as stemming from bounded rationality considerations. Indeed, the strategy restrictions that we consider can sometimes be motivated by the way the agent thinks *a priori* of the problem he faces, and his way of thinking may not be well adapted to the actual problem faced.

In the repeated game, we modeled the agent as if he had in mind a learning situation, trying to determine from experience and experimentation which of two arms is currently best. This shaped the agent's strategy set in a particular way. In the sender-receiver game, we proposed two models. In the first model, the focus was on the fact that the expert might be biased, with the agent trying to determine the degree to which he should correct the expert's suggestion based on his perception of the expert's bias. In the second, the agent behaves as if the expert was either benevolent or stupid, not considering the possibility that he is biased, trying to determine whether the expert should be trusted. Each of these assumptions shapes the decision maker's strategy set in a particular way. Both aspects of behavior are likely relevant, and, through further endogenization, the degree to which one aspect prevails can be examined. The fact that the two aspects of behavior can be disentangled is, in our view, a plus.

In the end, the proposed strategy restrictions above are just another expression of the necessary gap between perceived and actual situations. One may want to view this gap as stemming from a bound on rationality, or one may think of our models as putting plausible limits on sophistication. Either interpretation is fine with us.

References

Gigerenzer, G. (2007). *Gut Feelings: The Intelligence of the Unconscious*, New York: Viking.

Knight, F. H. (1921). *Risk, Uncertainty, and Profit*, Boston, MA: Hart, Schaffner & Marx; Houghton Mifflin Co.

Morris, S. and Shin, H. S. (2003). Global games: theory and applications. In M. Dewatripont, L. Hansen and S. Turnovsky, eds., *Advances in Economics and*

Econometrics (Proceedings of the Eighth World Congress of the Econometric Society), Cambridge: Cambridge University Press, p 56.

Ramsey, F. P. (1926). Truth and probability. In R. B. Braithwaite, ed., Ramsey, 1931, *The Foundations of Mathematics and other Logical Essays*, pp. 156–198, London: Kegan, Paul, Trench, Trubner & Co., New York: Harcourt, Brace and Company.

Rubinstein, A. (1989). The electronic mail game: strategic behavior under "almost common knowledge." *The American Economic Review*, 79(3), 385–391.

Savage, L. J. (1954). *The Foundations of Statistics*, New York: John Wiley and Sons, 1954.

Simon, H. (1993). Strategy and organizational evolution. *Strategic Management Journal*, 14, 131–142.

Wilson, R. (1987). Game theoretic analysis of trading processes. In T. Bewley, ed., *Advances in Economic Theory*, Cambridge: Cambridge University Press, pp. 33–70.