1

Introduction and Background

1.1 ERRARE HUMANUM EST

That humans do not always optimize perfectly is not the least bit controversial, even among economists and game theorists. In games of skill, we often see experts making errors. Chess grandmasters blunder against weaker opponents. Downhill skiers fall. People get busy and forget to return important phone calls. Some of these imperfections may be rationalizable by incomplete information. The batter whiffs because of a sudden gust of wind; the skier falls after hitting an invisible patch of ice. But some of these errors are harder to explain, and even for these simple ones it is often difficult or impossible to identify explicitly the source. It turns out that this mundane observation, captured in a nutshell as a Latin proverb, has important implications for the theory of games, which brings us to the subject of this book: quantal response equilibrium (QRE).

The theory of noncooperative games, as embodied by the Nash equilibrium, is based on three central tenets, which combined, rule out errors in decisions and beliefs. First, choice behavior by players in the game depends on *expectations* about the choice behavior of other players in the game. Second, individual choice behavior is *optimal* given these expectations. Third, expectations about the behavior of other players are *correct*, in a probabilistic sense. When these three elements are in place, an internally consistent model of behavior emerges, in the form of an equilibrium distribution of action profiles, that is stable.

Quantal response equilibrium relaxes the second of these assumptions, that is, that individual choice behavior is always optimal, by allowing for players to make mistakes. In particular, the expected payoffs calculated from players' expectations may be subject to "noise" elements, as indicated by the "+ Error" component in the top part of figure 1.1, which may lead players to choose suboptimal decisions.

In turn, this has significant implications about what constitutes equilibrium play in the game. If we retain the other central tenets of game theory, then players in a game, who form correct expectations about how others behave, must take account of the fact that other players don't always perfectly optimize. But this implies that errors have *equilibrium effects*: one player's expectations about the

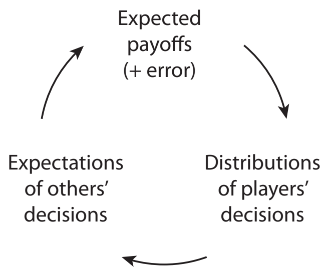

FIGURE 1.1. Quantal response equilibrium as a fixed point in beliefs and error-prone actions.

possibility of errors by other players alters the strategic calculus in ways that fundamentally change behavior in the game.

Examples abound. One might be advised to drive more cautiously late on New Year's Eve because of the presence of error-prone drunk drivers. Expert chess players may lay traps in hopes that their opponents will fall into them. While such moves are disadvantageous and can even lead to certain defeat when the opponent sees them coming, they are advantageous if the probability that the opponent will not respond optimally is high enough. In sports (and other cases), players even take into account the possibility of their own errors as well in adjusting their behavior. Tennis players may let up on their first serve because of the risk of also missing their second serve. A quarterback may take a sack rather than risking an errant throw. Some batters in baseball strike out often because they are waiting for the pitcher to make a mistake and throw a pitch that can be slugged out of the park.

There are countless other examples where players' awareness of their own or others' imperfect optimization alters strategic incentives. But the story does not stop there. In game theory, there are never-ending levels of strategic calculation. If one's tennis opponent is likely to ease up on the first serve, one should take a more aggressive position in returning the serve by moving in a few steps, but this tempts the server to serve harder, which increases the probability of error. A quantal response equilibrium takes account of all of these—potentially infinite—levels of strategic play. It recognizes that players make errors, that players understand and take account of the fact that errors are made, that they understand that other players recognize the possibility of error, and so forth ad infinitum. A quantal response equilibrium is a stable point, where everyone optimizes as best as they can subject to their behavioral limitations (in the sense that they may make errors), in full recognition that everyone else optimizes as best as they can subject to their behavioral limitations.

Quantal response equilibrium theory was developed largely in response to data from laboratory experiments that regularly show systematic and sometimes large departures from the basic predictions of standard game theory. This is true even in the simplest imaginable games, such as two-person bimatrix zero-sum games with a unique mixed-strategy equilibrium (O'Neill, 1987; Brown and Rosenthal, 1990; Ochs, 1995). When games have multiple equilibria, further restrictions are typically imposed on the data, such as those implied by subgame perfection, and the news is even worse (Güth, Schmittberger, and Schwarze, 1982; Thaler, 1988). The evidence is overwhelming that, as a predictive model of behavior, the standard theory of equilibrium in games needs an overhaul, and fortunately game theorists and economists have been working hard to improve the theory.[1] The need for doing this is made even more urgent since traditional game theory is regularly applied as a predictive tool in policy. For example, mechanism design theory applies game theory to develop and improve incentive structures for firms and regulatory agencies, to design better auction and bargaining mechanisms, and improve institutions for the provision of public goods. Game theory is also used to make policy prescriptions for government intervention in imperfectly competitive industries; and in fact this is one of the most significant innovations in the field of industrial organization and applied microeconomics in the last few decades. More recently, game-theoretic predictions have been applied to a variety of issues that arise in the study of international trade.[2] QRE is squarely placed in this broad effort to develop a better (and more useful) theory of games.

One possible response might be to abandon all three of the central tenets of Nash equilibrium and explain these deviations in terms of psychological propensities. The subject of this book, QRE, represents a more incremental approach. While judgment biases, and other sources of deviations from rational behavior, surely play some role in these departures, there is still a need for models that account for the equilibrium interaction of these effects. Initial attempts at this sought to explain systematic deviations from Nash equilibrium by presuming the existence of specific types of players who do not seek to maximize their expected payoffs in the experimental games. While this approach, and related applications of psychological theories of bias, have had success in explaining some of the departures from Nash equilibrium, they are open to the criticism of being post hoc, in the sense that the "invention" of types is specific to the game under consideration. What would be more desirable is a *general* model that does not need to be tailored to each game and each data set.

[1] See Camerer (2003).

[2] In fact, if one goes back further, to the 1950s when the first surge of research in game theory began, the bulk of the research was largely motivated (and funded) by applications to arms races and other problems of international conflict and national security.

This book is all about one such class of general models, QRE, and its applications to economics, political science, and pure game theory. By viewing choice behavior as including an inherently random component, QRE is a statistical generalization of Nash equilibrium. Its *statistical* foundation is closely related to quantal or discrete choice models that have been widely employed by statisticians in biometrics, psychology, and the social sciences, to study a wide range of scientific phenomena, including dosage response, survey response, and individual choice behavior. But it goes a step further by studying *interactive choice behavior*, and is based on the notion of rational expectations, a concept at the heart of modern economics and game theory. It is a *generalization* of the standard model of Nash equilibrium in the sense that Nash equilibrium is formally nested in QRE as a limiting case in the absence of decision error.

Thus, from a methodological perspective, QRE combines the stochastic choice approach, now commonly used in discrete choice econometrics, with the Nash equilibrium theory of games, the standard paradigm for studying strategic interaction in economics and political science. Strategies in a game are chosen based on their relative expected utility, but the choices are not necessarily best responses, which brings a flavor of limited rationality to the study of games, and does so in a general way. But, while closely related to ideas of limited rationality, it has its foundations in the theory of games of incomplete information. One of the central theoretical results of the next chapter is that there is a direct correspondence between quantal response equilibria of a game and the Bayesian equilibria of an expanded game with incomplete information.

The hybridization of a standard econometric approach with game theory opens up the possibility of using standard statistical models for quantal choice in a game-theoretic setting. With quantal response equilibrium, best-response functions are probabilistic (at least from the point of view of an outside observer) rather than deterministic. While better strategies are more likely to be played than worse strategies, there is no guarantee that best responses are played with certainty. Importantly, this imperfect response behavior is understood by the players. Like Nash equilibrium, this leads to a complicated equilibrium interplay of strategizing by the players, where they have to think about how other players are playing, realizing that those other players are thinking about them, too. The fixed point of such a process is a quantal response equilibrium. In the simplified perspective provided in figure 1.1, the fixed point requires that the probability distributions representing players' expectations (left-hand side) match the distributions of players' actual decisions (right-hand side), that is, the vector of equilibrium distributions gets mapped into itself. This equilibrium analysis can provide a limiting point of a learning process in which expectations evolve with observed distributions of others' decisions. In this book, we also discuss models of stochastic learning and introspection that have proved to be useful in dynamic or novel settings where players' expectations may not be aligned with the actual distributions of others' decisions.

The first paper on QRE, published two decades years ago in *Games and Economic Behavior*, proposed a structural version of QRE based on the additive random-utility model of choice that provides the choice-theoretic foundation for discrete choice econometrics (McKelvey and Palfrey, 1995). The basic idea is that there are many different possible "types" of players, but these types are known only to the agents themselves, and cannot be observed directly by the econometrician or, in the case of interactive games, by the other players. The statistical variation in behavior that is always observed in experiments can thus be interpreted as the result of some underlying distribution of these types. Hence observed or anticipated behavior will appear to be stochastic, even if all agents know their own type and always adopt pure strategies. This way of viewing QRE owes significant debt to Harsanyi's (1967, 1973) landmark contributions to our understanding of games of incomplete information, which today is called Bayesian Nash equilibrium; in fact, as originally defined by McKelvey and Palfrey (1995, 1998), QRE corresponds to a Bayesian equilibrium relative to some distribution of additive random-utility disturbances to the game.

These connections help explain why QRE has become an important tool for the statistical analysis of data from experimental games. It serves as a formal structural econometric framework to estimate behavioral parameters using both laboratory and field data, and can also lead to valuable insights into theoretical questions such as equilibrium selection and computation of equilibrium. There is now an extensive literature that either uses QRE in the analysis of data, or is focused exclusively on studying the theoretical properties of QRE in specific classes of games. The time is ripe for a book that pulls this work together into a single volume. The aim of the book is to lay out for an informed reader the broad array of theoretical and experimental results based on QRE. It contains some genuinely new material, offering new directions and open issues that have not been explored in depth yet, as well as delving into a range of peripheral issues, tangencies, and more detailed data analysis than the typical journal-article format allows for. There is also a "how-to" aspect to the book, as it provides details and sample programs to show readers how to compute QRE and how to take it to the data.

1.2 WHY A STATISTICAL VERSION OF NASH EQUILIBRIUM?

The simple answer is that it is a better descriptor of real human behavior, which is inevitably error prone. But there are deeper methodological reasons. Laboratory

TABLE 1.1. A dominance solvable game (Lieberman, 1960).

	a_1	a_2	a_3
a_1	0	15	−2
a_2	−15	0	−1
a_3	2	1	0

studies of games, ranging from pedantically simple to highly complex, all share two features. First, as pointed out above, systematic contradictions to the most basic predictions of standard game theory are commonplace. But a more severe problem for formally testing or rejecting theories is that, even where classical game theory predicts reasonably well "on average," there is still a huge amount of variation in behavior that is inconsistent with the theory. This second problem is particularly troublesome for games with a pure-strategy equilibrium, because standard theoretical models generally don't allow for any variation in the data.[3] Point predictions about behavior are simply too sharp and easy to reject, even if they can approximate observed behavior reasonably well on average. When games have multiple equilibria, one typically needs to place further restrictions on the data, such as the restrictions implied by subgame perfection and other refinements or selection criteria, and the performance of standard theory is even worse (e.g., in ultimatum and centipede games). However, these further restrictions generally share the same weaknesses (or are more severe) with respect to point predictions.

These two problems—systematic contradictions to predicted behavior and the overly restrictive nature of point predictions—are nicely illustrated by the following symmetric two-player zero-sum game that is strictly dominance solvable (and conceptually much easier than many children's games such as tic-tac-toe). Each player has three possible action choices. The payoff matrix is shown in table 1.1, with the entries being Row's payoffs (and the negative of Column's payoffs). Action a_2 is strictly dominated by a_3. Once these dominated actions are removed, then a_1 is strictly dominated by a_3. Thus the unique prediction from any standard equilibrium notion, or rationalizability, is for both players to choose a_3. Yet, in the laboratory, one observes all three strategies being played by both the Row and Column players. Indeed, both Row and Column very often choose a_1.[4]

In order to model behavior in games like this in a rigorous way for statistical analysis, one needs a model that admits the statistical possibility for any strategy

[3] A similar problem arises with respect to market environments in the laboratory, where the theoretical models being studied (competitive equilibrium) are typically inconsistent with price dispersion. One imagines it may be possible to adapt the quantal response equilibrium approach to develop a statistical theory of market equilibrium. There were some early attempts to apply this approach to demand theory and utility theory (e.g., Block and Marschak, 1960).

[4] A laboratory experiment based on this game was conducted by Lieberman (1960).

to be used. Indeed, without such a statistical model, for any game with a unique pure-strategy equilibrium, a single observation of a player using a different strategy (as is invariably the case in the laboratory) leads to immediate rejection of Nash equilibrium at any level of significance. This is sometimes referred to as the zero-likelihood problem, since the theoretical model assigns zero-likelihood to some data sets.[5] Call a model that assigns positive probability to all data sets a *statistical model*. Without a statistical model, standard maximum-likelihood methods are virtually useless for analyzing most data sets generated by laboratory experiments, due to unspecified variation outside the model.

Aside from statistical problems, the zero-likelihood problem also causes headaches for game theory itself. The classical problem of specifying beliefs "off the equilibrium path" in extensive-form games has never been resolved satisfactorily. Bayes' rule does not apply if the denominator is zero. And once Bayes' rule is lost, one is in a theoretical wilderness without a compass, left with ad hoc assumptions about subsequent behavior. In signaling games, for instance, what one ends up with is a morass of "belief-based refinements" that led theorists down a dead end in search of the holy grail of equilibrium concepts.[6] In repeated games, we are left with subtle and questionable models of "renegotiation" when a player happens to wander off the equilibrium path.

In contrast, a statistical approach to game theory has no wilderness, no off-path behavior, because no move is ever completely unexpected and hence no player is ever completely surprised. The problem, of course, is how to formulate such a statistical theory, the subject of this book.

1.3 SIX GUIDING PRINCIPLES FOR STATISTICAL GAME THEORY

Quantal response equilibrium theory is based on six guiding principles. The first principle, *completeness* (or *interiority*), requires the model to avoid the zero-likelihood problem discussed above. That is, all possible actions occur with positive probability. The second principle, *continuity*, requires that choice probabilities vary continuously in expected payoffs. The third, *responsiveness*, requires that the probability of choosing an action rises with its expected payoff. The fourth principle, *monotonicity*, requires that one action is more likely to be chosen than another if its expected payoff is higher. The fifth, *equilibrium*, requires rational expectations. That is, each player's choice probabilities correspond to quantal

[5] In the example of table 1.1, the Nash equilibrium assigns zero likelihood to nearly all data sets.

[6] Brandts and Holt (1993) report a series of laboratory experiments that provide strong evidence that equilibrium refinements are not useful in terms of explaining which Nash equilibrium is most commonly observed. See chapter 7.

response behavior relative to the expected payoffs in a game, evaluated at the actual choice probabilities of others. The sixth principle, *generality*, requires the model to be specified in sufficient generality to apply to arbitrary games. A general model is not ad hoc; it does not need to be specially tailored for each game and/or each data set. Variations are possible, of course, and the book discusses several that have been developed that follow only a subset of these principles.

1.4 A ROAD MAP

The remaining chapters of this book explore the theoretical properties of QRE, its implications for the analysis of experimental and field data, and applications to a variety of games of particular interest to political scientists, economists, and game theorists. We discuss how QRE can organize and explain many observed behavioral anomalies, how QRE can be computed efficiently, and how QRE can be used as a structural econometric model to fit data and to measure parameters from decision-theoretic and behavioral models.

The book is organized into two parts: “Part I: Theory and Methods” (chapters 2–6), and “Part II: Applications” (chapters 7–9), although part I also includes a number of simple examples to illustrate the basic principles and for a clearer presentation of the theory.

Part I begins by laying out the general theory of quantal response equilibrium for normal-form and extensive-form games, in chapters 2 and 3 respectively. Proofs of some of the main propositions are included. Chapter 4 explores issues related to individual heterogeneity with respect to player error rates, or what we refer to as *skill*. That chapter also describes some extensions of QRE that relax the assumption that player expectations about the choice behavior of other players are correct. For example, in games that are played only once, players are not able to learn from others’ prior decisions, and expectations must be based on introspection. Chapter 4 develops the implications of *noisy introspection* embedded in a model of iterated thinking. Chapter 5 explores questions related to learning and dynamics. Part of that chapter lays out the theory for how to apply QRE to repeated and dynamic stochastic games with an infinite horizon and stationary structure. Chapter 6 provides some user-friendly examples to illustrate how QRE can be computed numerically in games, and how QRE can be used as a structural model for estimation. Sample computer programs in MATLAB for specific examples are provided, which the reader can use as a template to adapt to specific applications.

Part II analyzes in detail many different applications of QRE to specific classes of games, in order to give a sense of the wide range of problems that QRE is useful for, and to illustrate many of the ideas that were developed in chapters 2–6. The applications in chapter 7 include a variety of simple games that illustrate the effects of QRE in several canonical settings that are of widespread interest to game theorists and social scientists: social dilemmas (the traveler's dilemma), adverse selection (the compromise game), signaling (the poker game), and social learning. Chapter 8 examines the effect of QRE in a range of games of particular interest to political scientists, including free riding, voter turnout, information aggregation in committees and juries, crisis bargaining in international relations, and dynamic bargaining in legislatures. Many of the sharply nonintuitive Nash predictions for voting and political bargaining games are reversed or modified by adding sensible amounts of noise in an equilibrium analysis. Chapter 9 is oriented toward economic applications, including games with continuous strategy spaces, such as auction, pricing, and bargaining games. Most of the applications in these three chapters also include a summary of experimental results that have used QRE in the analysis of the data.

The final chapter 10 speculates about directions in which future developments in QRE theory and experimental work could contribute new insights about strategic behavior.