

Roger B. Myerson

The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2007

Autobiography

Becoming a game theorist

A scholar's greatest asset is his or her intuition about what questions to study and with what methodology. A scientific autobiography should shed some light on how this intuition grew and developed over time.

The interests that shape one's adult life generally have deep roots in childhood. I grew up in a comfortable suburb of Boston with a fine public school system, in a family that greatly valued reading and scientific learning. My father did research and engineering for a family business of manufacturing artificial teeth, and each of my parents returned to school to earn advanced degrees at different times in my youth.

Concern about the new risks of nuclear war was widespread in the 1950s, and, like many of my generation, I was aware of this terrible threat from a young age. I have early memories of telling my father that I was worried about political cartoons that depicted global dangers of the 1956 Suez crisis. My father reassured me that the leaders of the world were bringing all their wisdom and understanding to the task of managing the crisis peacefully. This perspective suggested, however, that perhaps it might be better if our leaders could have even more wisdom and understanding, to provide guidance for a safer and more peaceful world in the future.

When I was twelve, I read a classic science-fiction novel that depicted a future where advanced mathematical social science provided the guidance for a new utopian civilization. Ideas from this and other readings grew in long discussions with my friends. It was natural, perhaps, to hope that fundamental advances in the social sciences might help find better ways of managing the world's problem, as fundamental advances in the physical sciences had so dangerously raised the stakes of social conflict in our time.

I have always loved to read about history and found a fascinating beauty in historical maps. But I hoped for something more analytical, and so I was naturally intrigued when I first heard about economics. I began reading Paul Samuelson's basic economics textbook in a high school vacation. When I got to college, I chose to concentrate in economics and applied mathematics, but my high-school chemistry teacher predicted that I would switch to physical

science before the end. I was not sure whether he might have been right until I discovered game theory in 1972.

In the spring of 1972, as a third-year student in Harvard College, I took a beautiful course on decision analysis from Howard Raiffa. He taught us to see personal utility functions and subjective probability distributions as measurable aspects of real decision-making that are expressed (however imperfectly) in our daily lives. At the end of the course, he told us that the analysis of interactions among two or more rational utility-maximizing decision-makers is called game theory, and he described game theory as a field in which only limited progress had been made. This negative assessment provided a positive focus for my studies thereafter. I felt that, if I do not know how to analyze such obviously fundamental models of social decision-making, then how could I pretend to understand anything in social science? I started reading a book on game theory that summer.

There were no regular courses on game theory then at Harvard, and so I began to do independent reading on the subject, searching through the libraries for books and articles about game theory. My primary form of intellectual dialogue was scribbling notes into the margins of photocopied journal articles, which were written by the distant leaders of the field: Robert Aumann, John Harsanyi, John Nash, Thomas Schelling, Reinhard Selten, Lloyd Shapley, and others. Their published writings gave me good guidance into the field. In particular, when I discovered the work of John Harsanyi, some time in the fall of 1972, I really knew that I had found the research program that I wanted to join.

I was first attracted to Harsanyi's work by his (1963) paper that defined a general cooperative solution concept which included both the two-player Nash bargaining solution and the multi-player Shapley value as special cases. These two solution concepts had elegant axiomatic derivations, and their single-point predictions were much more appealing to me than the multiple sets of solutions that other cooperative theories identified. I worked for three days in the library to understand and reconstruct the derivation of Harsanyi's cooperative theory, simplifying it until I found that everything could be reduced to a simple balanced-contributions assumption. This was my first result in game theory.

But Harsanyi also wrote a series of papers in 1967-8 about how to model games with *incomplete information*, in which the players have different information at the beginning of the game. In 1972, Harsanyi and Selten published a new paper that suggested a generalized Nash

bargaining solution for two-person games with incomplete information. So I saw an important question about game theory that had not yet been addressed in the literature: How can we extend these cooperative solution concepts to games with more than two players who have incomplete information about each other? Nobody had any general theory for predicting what might happen when cooperative agreements are negotiated among rational individuals who have different information. This was the problem that I set out to solve in my dissertation research.

I did not solve this problem in college or in graduate school, but it was a very good problem to work on. To try to build a theory of cooperation under uncertainty, I first needed to rethink many of the fundamental ideas of cooperative game theory and noncooperative game theory, and along the way I got a reasonable dissertation's-worth of results. My advisor Kenneth Arrow patiently read and critiqued a series of drafts that gradually lurched toward readability.

In 1976 I had the good fortune to be hired as an assistant professor by the Managerial Economics and Decision Sciences (MEDS) department in the (soon-to-be Kellogg) School of Management at Northwestern University. In the 1970s, game theory was a small field, and few schools would consider having more than one game theorist on their faculty, but Northwestern was actively building on strength in mathematical economic theory. The MEDS department was probably the only academic department in the world where game theory and information economics were not viewed as peripheral topics but as central strengths of the department. I had great colleagues, and every year we went out to hire more.

My first papers were largely about cooperative game theory, including results from my doctoral dissertation. Cooperative game theory generally begins with the assumption that people will agree on some feasible outcome that is *efficient*, in the sense that there is no other feasible outcome that all of the cooperating individuals would prefer. But in most situations we can find a wide range of such efficient allocations, with different alternatives being better for different individuals. An *equitable* bargaining solution should identify an efficient outcome in which each individual gets a payoff that is in some sense commensurate with his or her contribution to the collective agreement. In several early papers that were based on my work in graduate school, I showed how simple principles of equity between pairs of individuals could be consistently extended to situations where many individuals cooperate in coalitions.

Everything that I did in game theory was ultimately motivated by the long-run goal of developing a coherent general methodology for game-theoretic analysis. What kinds of game

models should we use to describe situations of conflict and cooperation among rational decision-makers, and what solution concepts should we use to predict rational behavior in these game models? In this quest, I was greatly influenced by three classic ideas of game theory: von Neumann's principle of strategic normalization for reducing dynamic games, Nash's program for subsuming cooperative games into noncooperative game theory, and Schelling's focal-point effect for understanding games with multiple equilibria. These ideas are so important that I cannot describe my work on game theory without explaining something about them.

Game theory could not have developed without a basic understanding that some conceptually simple class of models could be general enough to describe all the complicated game situations that we would want to study. John von Neumann (1928) argued that one-stage games where players choose their strategies independently can be recognized as such a general class of models, even if we want to study dynamic games where play may extend over time through many stages. The key to this argument is to define *strategy* for a player to be a complete plan that specifies what the player should do at each stage in every possible contingency, so that each player could choose his strategy independently before the game begins. By this principle of strategic normalization, for any dynamic extensive game, we can construct an equivalent one-stage game in *strategic form*, where the players choose strategies independently, and these strategy choices then determine everyone's expected payoffs.

After presenting this argument for the generality of the strategic form, however, von Neumann actually studied cooperative games in a different nonstrategic *coalitional-form* model of games. In response, John Nash (1951) argued that the process of bargaining should itself be recognized as a dynamic game where players make bargaining moves at different stages of time, and so it should be similarly reducible to a game in strategic form. To analyze games in strategic form, Nash defined a general concept of equilibrium. In a *Nash equilibrium*, the predicted behavior of each player must be his or her best response to the predicted behavior of all other players.

When we follow this Nash program and write down simple models of bargaining games, however, we regularly find that these games can have many equilibria, and so a predictive theory cannot be determined without some principles for selecting among all these equilibria. Schelling (1960) argued that, in a game that has multiple equilibria, anything that focuses the players' attention jointly on one equilibrium can cause them to expect it and thus rationally to act

according to it, as a self-fulfilling prophecy. The focal factors that steer the players to one particular equilibrium can be derived from the players' shared cultural traditions, or from the coordinating recommendations of a recognized social leader or arbitrator, or from any salient properties that distinguish one equilibrium as the focal equilibrium that everybody expects to play. In my view, the cooperative solutions which I was studying in my early papers were theories about how welfare properties of equity and efficiency can identify a focal equilibrium, which an impartial arbitrator could reasonably recommend, and which may be implemented by the players as a self-fulfilling prophecy.

I understood, however, that some general principles for eliminating some Nash equilibria might be also be appropriate. In particular, I knew that the theory of Nash equilibria for strategic-form games could yield predictions that seemed irrational when interpreted back to the framework of dynamic extensive games with two or more stages. The problem is that, if an event were assigned zero probability in an equilibrium, then the question of what a player should do in this event would seem irrelevant when the player plans his strategy in advance, and so a Nash equilibrium could specify strategic behavior that would actually become irrational for the player if this event occurred. In graduate school, I recognized that such irrational equilibria could be eliminated by admitting the possibility that players might make mistakes with some infinitesimally small probability. Reinhard Selten (1975) was also developing similar ideas about refinements of Nash equilibrium. When I read his important paper on perfect equilibria in the *International Journal of Game Theory*, I sent my paper on proper equilibria to be published there also. The motivation for these refinements of Nash equilibrium was to become clearer a few years later, when David Kreps and Robert Wilson defined *sequential equilibria* of dynamic extensive-form games (1982). In this sense, perfect and proper equilibria were attempts to recognize in strategic form the equilibria that would be sequentially rational in the underlying dynamic game. But Kreps and Wilson cogently argued for dropping our reliance on strategic normalization and instead analyzing sequential equilibria of games in the dynamic extensive form.

Learning to analyze incentive constraints in games with communication

Basic questions about information and incentives in economic systems were very much in the air at Northwestern in the late 1970s and early 1980s. There was great interest in Leonid

Hurwicz's ideas of incentive-compatibility, and these ideas influenced my search for a cooperative theory of games with incomplete information. But from my student days, I had learned to see games with incomplete information in the general framework of Harsanyi's Bayesian model. After I heard Alan Gibbard's early version of the revelation principle for dominant-strategy implementation, I became one of several researchers to see that it could be naturally extended to implementation with Bayesian equilibria in Harsanyi's framework. The revelation principle basically says that, for any equilibrium of any communication system, a trustworthy mediator can create an equivalent communication system where honesty is a rational equilibrium for all individuals. With the revelation principle, we can generally apply some mathematically simple constraints that summarize the problems of incentives for getting people to share their information honestly. So I wrote a paper showing how the revelation principle for Bayesian incentive compatibility could extend and simplify the feasible set for Harsanyi and Selten's (1972) bargaining solution for two-person games with incomplete information. This idea was the basis of my first article on the revelation principle, published in *Econometrica* in 1979.

In the Harsanyi-Selten bargaining solution, however, the objective is to maximize a little-understood multiplicative product of the players' expected utility gains. In discussions with Robert Wilson and Paul Milgrom about auctions, I began to see that there might be more interesting economic applications where other objective functions are maximized over the same feasible set. I realized that the revelation principle could become a general tool for optimizing any measure of welfare in any situation where there is a problem of getting information from different individuals. During a visit to the University of Bielefeld in Germany (1978-9), I wrote a paper that applied these ideas to the problem of designing an auction, where the objective is to maximize the seller's expected revenue, subject to the incentive constraints of getting potential buyers to reveal information about their willingness to pay. When I returned to Northwestern, I worked with colleagues on other important applications and extensions of these ideas. David Baron and I worked on optimal regulation of a monopolist with private information about costs. Mark Satterthwaite and I analyzed efficient mechanisms for mediation of bilateral trading problems, involving one seller and one potential buyer for a single indivisible good.

With some simple but natural technical assumptions, we were able to derive powerful *revenue-equivalence* theorems, which show how the expected profit for any type of individual

with private information can be computed from the expected net trades for other possible types of the same individual. In this calculation, an individual's profit is seen to depend on the way that his private information affects the ultimate allocation of valuable assets, but not on the details of how asset prices are determined in the market. Thus, people who have private information may be able to earn informational rents under any system of market organization in which the allocation of assets depends on their information.

The problem of providing incentives for people to share private information honestly, which has been called the problem of *adverse selection*, was not the only incentive problem that economists were learning to analyze in the 1970s. There was also a growing literature on the problem of providing incentives for individuals to choose hidden actions appropriately, which has been called the problem of *moral hazard*. Robert Aumann (1974), in his definition of correlated equilibria for games with communication, formulated a version of the revelation principle that applied to moral-hazard problems. In 1982 I published a paper to extend the revelation principle in a unified way to problems with both moral hazard and adverse selection, yielding a general theory of incentive-compatible coordination systems for Bayesian games with incomplete information. Later, in 1986, I showed how to extend the revelation principle to dynamic multi-stage games where sequential rationality is required in all events.

In these extensions, I began to see a basic methodological dichotomy between the revelation principle and the principle of strategic normalization: If we use one then we cannot use the other. When we allow that players can communicate with each other through mediated channels that are not specified explicitly in the game model, then the revelation principle can be applied, and the complicated nonlinear mathematics of Nash equilibrium can be replaced by the simple linear constraints of incentive compatibility. The set of incentive-compatible mechanisms includes all equilibria that can be achieved by adding any mediated communication system to the given game, and so this set can be viewed as the feasible set for a focal arbitrator or leader who can design the communication system. But the existence of implicit opportunities for communication among the players during the game invalidates the argument for strategic normalization, and so one-stage strategic-form games can no longer be considered as the general model for all situations. Thus, when we analyze games with communication, we study Bayesian games and more general extensive-form games, setting aside the principle of strategic normalization, as Kreps and Wilson suggested.

Bengt Holmstrom and I published a paper in 1983 on how economic concepts of efficiency should be extended to situations where people have private information. According to Pareto's basic concept of efficiency, the functioning of the economy is efficient when there is no feasible way to make everyone better off. But we must re-think many parts of this definition when we admit that each individual may have different private information, which is represented in our Bayesian games by a random privately-known *type* for each player. An economist's evaluation of efficiency or inefficiency cannot depend on information that is not publicly available, and a realistic concept of feasibility must take account of incentive constraints as well as resource constraints. Thus, the concept of incentive-efficiency should be applied to the entire mechanism or rule for determining economic allocations as a function of individuals' privately-known types, not just to the final allocation itself. Holmstrom and I suggested that an incentive-compatible mechanism should be considered *incentive-efficient* (in the *interim* sense) when there is no other incentive-compatible mechanism that would increase the expected utility payoff for every possible type of every individual. If a change would make even one possible type of one individual worse off, then an economist cannot say for sure that the change would be preferred by everyone, when each individual privately knows his or her own type.

In the early 1980s, I returned to the problem of defining general cooperative bargaining solutions for Bayesian games where players have different information. To identify one bargaining solution among the many mechanisms on the interim incentive-efficient frontier, we need to define some principles for equitable compromise, not just among the different individuals, but also among the different possible types of any one individual. In particular, to avoid leaking private information, a player's bargaining strategy may need to be an inscrutable compromise between the payoff-maximization goals of his actual type and the goals of the other possible types that the other players think he might be. Such an inscrutable intertype compromise must be defined by some principles for measuring how much credit each possible type can claim for the fruits of cooperation. For example, when a player could be a good type or bad type and would prefer to be perceived as the good type, so that the incentive constraint that bad types should not gain by imitating good types is binding and costly in incentive-efficient mechanisms, then an inscrutable intertype compromise might put more weight on the goals of his good type than his bad type. In the early 1980s, I began to see that this idea can be mathematically formalized and measured by using the Lagrange multipliers of the informational

incentive constraints in the mechanism-design problem. Economists have long recognized the importance of Lagrange multipliers of resource constraints, which correspond to prices of these resources, but economists had not previously done much with Lagrange multipliers of incentive constraints. With this mathematical insight, I was at last able to solve the problem of extending the Nash bargaining solution and the Shapley value to Bayesian games with incomplete information, which I published in *Econometrica* and the *International Journal of Game Theory* in 1984.

New directions

In 1980 I met Gina Weber, and we were married in 1982. Our children Daniel and Rebecca were born in 1983 and 1985. Everything in my life since then has been a joint venture shared with them.

The biggest focus of my work in the later 1980s was on the writing of a general textbook on game theory, which was published in 1991. In this book, I presented, as coherently as I could, the general methodology for game-theoretic analysis that so many of us had been working to develop. Other textbooks on game theory were also published in the early 1990s, and general textbooks on economic theory also began to treat game theory, information economics, and mechanism design as essential parts of microeconomic analysis. When the importance of game theory in economics was recognized by the Nobel Committee in October 1994, by its award to John Nash, John Harsanyi, and Reinhard Selten, I opened bottles of champagne to celebrate with my colleagues at Northwestern.

The last section of my game theory book considered markets with adverse selection, where theorists since Michael Rothschild and Joseph Stiglitz (1976) have found that simple equilibrium concepts can have complicated nonexistence problems. I suggested that equilibria may be sustainable in great generality by a version of Gresham's law, that bad types may circulate more than good types in the market, and I formalized this argument in a paper in 1995.

In the late 1980s, I began to work on game-theoretic models of politics. I had always felt that game theory's applications should go beyond the traditional scope of economics. In constitutional democracies, political constitutions and electoral systems define the rules of the game by which politicians compete for power. So game-theoretic analysis should be particularly valuable for understanding how changes in such constitutional structures may affect the conduct

of politicians and the welfare-relevant performance of the government. I saw some analogy here with well-known ideas from the economic theory of industrial organization. In particular, the ability of political organizations to extract profits of power (corruption) should be expected to depend on barriers against the entry of new competitors into the political arena.

My first paper on comparison of electoral systems with Robert Weber was followed by several more papers in which I explored various models for evaluating the competitive implications of electoral reforms. For example, a politician could try to appeal broadly to all voters or could try to concentrate narrowly on appealing to small subgroups of the voting population; and I developed game models to show how different electoral rules can systematically affect politicians' competitive incentives to appeal more broadly or narrowly. In other papers, I showed how some electoral rules can yield a wider multiplicity of equilibria, including discriminatory equilibria in which some good candidates may not be taken seriously by the voters. I argued that such discriminatory equilibria can become barriers to entry against new politicians, thus allowing established political leaders to claim more profit from their political power. Democratic competition is intended to reduce such corrupt profit-taking by political leaders, but my analysis suggested that the effectiveness of democracy against such corruption can depend on the specific structure of the electoral system. In several models, I found that approval voting could induce stronger incentives for politicians to compete vigorously at the center of the political spectrum.

The constitutional distribution of power among different elected officials can also affect political incentives. Daniel Diermeier and I wrote a paper to show how presidential veto powers and bicameral subdivision of the legislature can decrease legislators' incentives maintain coherent discipline in political parties or legislative coalitions.

The analysis of games with incomplete information has shown economists how probabilistic analysis of decision-makers' uncertainties can offer practical insights into competitive strategies, but my MBA students at Kellogg did not seem well prepared to apply such analysis. So I felt that it was important to seek new ways of teaching that could make practical probability models accessible to them. After trying many different approaches throughout the 1980s and 1990s, I found that the best pedagogical solution was to use simulation modeling in spreadsheets. Models and techniques that seemed too difficult for students when I wrote on the blackboard became intuitively clear to them when we worked together in electronic

spreadsheets. So I wrote an MBA-level textbook on probability models for economic decisions which was published in 2005.

Since 1996, I have written several papers on the history of game theory itself. I had the privilege of speaking about the importance of John Nash's contributions to economic theory at the American Economic Association's luncheon in honor of his Nobel Prize in January 1996. This presentation developed into a longer paper that was published in the *Journal of Economic Literature* in 1999, for the 50th anniversary of the day when Nash submitted his first paper on noncooperative equilibrium.

As the end of the 20th century approached, I began to ask whether the advances in competitive analysis that meant so much to me might actually offer hope for a better 21st century. So I wrote a theoretical retrospective on the history of Germany's Weimar republic, the failure of which was so central among the disasters of the 20th century. The establishment of the Weimar republic was framed by the treaty of Versailles and the Weimar constitution, which were written in 1919 with expert advice from leading social scientists, including John Maynard Keynes and Max Weber. I wanted to ask whether any recent advances in political and economic theory might offer a better framework for understanding such practical problems of institutional design, so that mistakes like those of 1919 should be less likely in the future. In this retrospective, I found that the advances that seemed to offer the most valuable insights for improving international relations were based largely on the ideas of Thomas Schelling and Reinhard Selten, in particular, the focal-point effect and the analysis of strategic credibility.

So when America set out to invade Iraq in 2003, I applied Schelling's ideas about credible deterrence to show that how America's rejection of multinational military restraint could exacerbate threats against America. When a powerful nation uses military force without clear limits, instead of deterring potential adversaries, it can actually motivate potential adversaries to invest more in militant counter-forces. Thus, I have argued, the greatest superpower in the world may have the greatest need to articulate clear and credible limits on its use of military force, according to rules and principles which the rest of the world can judge.

In recent years, my research agenda has been increasingly shaped by a broad study of political history. To understand the great problems of political change and economic development, we need new theoretical models that can help us to better understand the functional logic of the traditional systems from which many nations are now evolving. To have any hope of

finding such broader models, however, a theorist needs the broadest possible understanding of different economic and political institutions of different societies throughout the world, from ancient times until today. For such a program of study, I found Samuel Finer's *History of Government* (1997) to be a particularly valuable introduction to the global history of political institutions, but it needs to be followed by much more reading. From this historical perspective, I have come to believe that new theoretical models of oligarchy or feudalism or tribalism could become as important for political analysis as the Arrow-Debreu general equilibrium model for analysis of competitive markets.

In searching for universal principles that underlie all kinds of institutional structures, I have come to see agency incentive problems and reputations of political leaders as critical factors that are fundamental to the establishment of any political institution. The rules of any political system must be enforced by government officials, and getting these officials to act according to these rules is a moral-hazard problem. The incentives for lower government officials depend on rewards and punishments that are controlled by higher political leaders. So we find, at the heart of the state, a problem of moral hazard for which the solution depends on the individual reputations of political leaders. The institutions of any political system must be organized by political leaders whose first imperative is to maintain their reputation for rewarding loyal supporters. The problem of cultivating a democracy can then be recognized as a problem of creating opportunities for more politicians to begin cultivating good democratic reputations, that is, reputations for serving the mass of voters even as they reliably reward their active agents and supporters with patronage benefits.

As an application of these ideas, I have written critically about American policies to establish democracy in occupied Iraq. I have argued that the first priority of the occupation authority in 2003 should have been to create elected and well-funded local councils, in which local leaders throughout the country could begin building independent reputations for responsible democratic governance.

In 2001 I accepted a job in the economics department at the University of Chicago. This was only my second academic position. Northwestern's long tradition of great strength in economic theory made it an ideal home for me to work with great colleagues in game theory and political economics. But after 25 years in one school it seemed time for a change, especially when the alternative on the other side of town was a university with another outstanding tradition

of scholarship in economics. I still value my former colleagues at Northwestern as much as my new colleagues at Chicago. Also, I am proud of another connection that I still have with Northwestern University, as my wife Gina is now a dean in Northwestern's McCormick School of Engineering.

In 2007, the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel was awarded to Leonid Hurwicz, Eric Maskin, and me. I am very proud to have my name so prominently linked with Leo's and Eric's, and with the advances in analysis of coordination mechanisms and incentive constraints to which we contributed along with many other great economists. The opportunity to share this honor with my wife and my family and to celebrate with so many friends and colleagues has been a wonderful experience for me. But I understand that the true laureates are the ideas of incentive analysis and mechanism design, and I want to continue working to understand them more deeply and to present them more clearly. There is much that we still need to learn about how our social institutions operate, and how they can be better designed.