AUTOBIOGRAPHY OF A GAME THEORIST

Roger B. Myerson

February 2014

Deciding what to study

My great-grandfather was a poor immigrant in America, but he brought from Europe a deep appreciation of scholarship. He liked to say that a scholar needed to be like a pin, with a sharp end for finding new insights and a dull end for the long work of studying and writing between flashes of insight.[1] I would suggest that the scholar's sharp end itself depends on two elements: a mastery of the analytical methods of one's academic discipline, and a passionate desire to understand something better with the help of these tools. The mastery of analytical tools comes from long study (where the dull end helps). But the ultimate source of the vital curiosity that drives our desire to study and learn remains somewhat mysterious to me. Let me illustrate with a personal story.

In a press conference at the 2007 Nobel celebration in Stockholm, a reporter asked the new Nobel laureates about how we could encourage more children to become interested in science. As my good fortune that day included not only being a new laureate but also having my parents there with me for the celebration, I took the question as a welcome opportunity to express my gratitude to them for years of encouragement and support. So I described to the reporter how my mother had actively encouraged me to develop an early interest in science by reading me a wonderful Disney book about atomic physics when I was in first grade. True, I did not become a physicist (though my brother did), but I wanted to honor my mother's decision to introduce science to me at such a young age. After the press conference, my mother thanked me for the nice words but admitted that she actually preferred reading fiction to children, and she included this science book in our bedtime reading only because I had insisted on it, bringing it from my big brother's bookshelf. This revelation was very surprising to me at the time, but I should have known from my own experiences as a parent that children's interests and passions must develop from within according to their own mysterious internal logic. Parents and teachers can encourage this development but cannot direct it.

When I talk to prospective doctoral students, I regularly ask them about what they might

[1] See Abraham Myerson, *Speaking of Man* (New York: Alfred A. Knopf, 1950), p 7.

want to study in their thesis research, but many students hesitate to respond because they feel that they do not know enough yet to identify a good research questions in economics. I then try to explain that any scholar's most precious asset is his or her sense of what is a good question that could be worth months or years of careful study. Students should expect that their teachers in graduate school will introduce them to new areas of research that may cause their interests to grow and change, but their potential to make unique contributions in science will ultimately depend on their personal interests and curiosity that guide their choice of research questions.

Starting with a passion to better understand some big questions about the world, scholars use the analytical tools of their discipline to rigorously study some small parts of these big questions, and in this way we make small careful steps toward the better understanding that we need. A researcher's passionate curiosity about a big question and mastery of the analytical tools of a discipline are both essential for the hard task of finding a small part of an important question that is ready for new analysis and understanding.

A scholar's motivation can be applied or theoretical. Indeed, an alternation between methodological interests and applied substantive interests is common in many academic careers. A concern for applied problems in the real world can lead one to develop new tools of theoretical analysis, and a sense of the untapped power of new analytical tools can lead one to explore other applied problems for which these tools seem potentially useful.

Since I was a student, I have always felt that questions of social science are important, because the peace and prosperity of our world may depend on a deeper understanding of how our societies function. When I learned in college that there were gaps in economists' basic ability to analyze interactions between rational decision-makers who have different information, I very much wanted to be part of filling those gaps. So my early research as a graduate student and as a young professor was motivated largely by methodological concerns, as I worked along with many others to develop fundamental methods of analysis in game theory and information economics. But in later decades, when these methodological gaps had been somewhat reduced, my research became more motivated by applied problems for which the methods of game theory seemed potentially useful, such as the comparison of political incentives under different electoral systems, or the problems of democratic state-building. The subsequent sections of this essay summarize the broad themes of this progression: from finding my way into the methodological questions of game theory, to developing a framework for analysis of incentive constraints in

games with communication, and then exploring new research directions with more applied motivations from political science and even macroeconomics.

Finding my way into economics and game theory

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. But 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 by political cartoons that depicted global dangers of the 1956 Suez crisis, when I was five years old. 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 problems, 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 we do not know how to analyze such obviously fundamental models of social decision-making, then how could one 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 knew that I had found a part of the research frontier that I wanted to help advance.

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. Harsanyi called these models *Bayesian games*, and the different possible states of each player's private information were called the player's possible *types*. In 1972, Harsanyi and Selten published a new paper that suggested a generalized Nash bargaining solution for two-person games with incomplete information. This Harsanyi-Selten solution maximized a particular social welfare function (computed from the expected gains from cooperation for every possible type of each player) over a rather complicated feasible set, but it was defined only for games with two

players. 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 three or more 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.

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 the play may extend over time through many stages. The key to this argument is to define a *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 or her 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. Thomas 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 developed more formally a few years later by David Kreps and Robert Wilson when they defined *sequential equilibria* of dynamic extensive-form games (1982). From this perspective, perfect and proper equilibria may be seen as 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.

Developing tools for analysis of incentive constraints in games with communication

Basic questions about information and incentives in economic systems were frequent topics of discussion among my colleagues 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 formulate some mathematically simple constraints that summarize the problems of

giving people appropriate incentives to share their private information. So I wrote a paper showing how these incentive constraints 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, which was 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 over the feasible set. 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 could affect 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.

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* 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. This recommended efficiency criterion considers each individual's welfare at the *interim* stage, after he has learned his own type but before he can learn anyone else's type. Holmstrom and I also discussed other alternative criteria that consider individuals' *ex ante* expected welfare before learning anyone's type, or *ex post* welfare after learning everyone's types.

Among the areas in which I had the privilege of contributing to information economics, I see a uniquely central importance of this paper with Holmstrom, even though it was not actually mentioned in the citation for the Nobel prize of 2007. Pareto's concept of efficiency is the most important normative criterion for economic analysis, and our paper was the first to carefully

consider how this fundamental concept should be extended to economic situations where people have different information. It was published in the prestigious journal *Econometrica* after the usual scrutiny by anonymous referees, and I still remember their reports on this paper. One referee said that the first half of the paper (which surveyed six alternative definitions for extending Pareto efficiency as a normative concept) was a fundamental publishable contribution but then said that the second half (which proposed a complicated extension of Pareto efficiency as a positive predictive concept) should be cut, as it was speculative and unconvincing. The other referee said that the first half should be cut, as it was obvious and trivial, but then said that the second half should be published as it was a subtle and interesting new idea. Fortunately the editor chose to publish the union of the sections that referees liked, rather than their empty intersection.

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 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 should 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.

In the years since then, my two papers on cooperative games with incomplete information have not attracted much interest, and probably very few people have ever read them. But my interest in these questions, even if it was not shared by many others, led me to approach other questions of information economics from a different perspective which was essential for my other contributions in mechanism design and auction theory which have been more widely valued.

New directions and research problems from political science and macroeconomics

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. These were bottles that I had been saving, in eager anticipation of such an event, for so long that the champagne had gone flat, but it still tasted good that day.

The last section of my game theory book considered markets with adverse selection, for which theorists since Michael Rothschild and Joseph Stiglitz (1976) have found that simple concepts of competitive equilibrium may be impossible to satisfy in simple examples. I suggested that equilibria for such markets may be sustained by a version of *Gresham's law*, that bad types would circulate in the market more than good types. I formalized this argument in a paper in 1995 which showed that such sustainable equilibria exist for general markets with adverse selection.

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 might 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 profit more 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 to 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. As I argued in a paper in 2008 in the *American Political Science Review*, the institutions of any political system must be organized by political leaders whose first imperative is to maintain their reputation for rewarding loyal supporters. I suggested that constitutional constraints on powerful leaders may be naturally derived from their fundamental need to maintain the long-term trust of their active 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 for serving the voters at large, even while rewarding active supporters with patronage benefits. In 2006 in the *Quarterly Journal of Political Science*, I published a formal model to show how federal division of powers across national and local governments can increase such opportunities for promoting new leaders with good democratic reputations.

As an application of these ideas, I have written critically about American policies to establish democracy in Iraq and Afghanistan. I have argued that, in 2003, the first priority of the

occupation authority in Iraq 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. The centralized constitutional structure of the government in Afghanistan after 2003 may have similarly frustrated democratic state-building efforts there.

After the financial crisis of 2008, I began to think more about the great questions of macroeconomic stabilization, the importance of which had unfortunately become much more evident to people everywhere. My guiding intuition was that, just as the problems of moral hazard in government should be fundamental to our understanding of politics, so the problems of moral hazard in finance should be fundamental to our understanding of macroeconomics. Moral hazard arises in finance because agents who are entrusted with responsibility for prudently investing other people's savings may be tempted to profit inappropriately by abusing their power over other's money. Problems of moral hazard in financial institutions were widely evident in the financial crisis, but such problems were often ignored in traditional macroeconomic theory, because the basic economic principles for analyzing moral-hazard agency problems were only developed as part of information economics in the 1970s.

In a recent (2012) paper on moral-hazard credit cycles, I have shown that a simple standard model of repeated moral-hazard problems can offer a fundamental explanation of macroeconomic instability. The key point is that efficient incentive plans involve large late-career rewards, to motivate an agent to maintain a record of good service throughout his or her career. The promise of a late-career reward for a good performance record may be seen as an investment in a relationship of trust, which is in many ways analogous to investments in capital that economists have traditionally studied. In each case, a large expense at one point in time can buy productive services that yield returns over a long interval of time. But whereas the greatest expense in a capital investment normally comes before the returns are realized, the largest moral-hazard incentive payments come late in the agent's career, after investors have earned returns from the agent's service. More importantly, whereas capital investments are normally assumed to depreciate over time, the motivational value of late-career incentive payments increases during an agent's career, as normal time discounting causes the agent to value the prospective late-career rewards more as the time of their realization draws nearer. Changing depreciation to appreciation is mathematically like changing the direction of time, which can change a dynamically stable system to an unstable system. Thus, where simple economic models with

capital investment tend to be dynamically stable, I found that equally simple economic models with moral-hazard agency problems tend to be dynamically unstable, falling into boom-bust cycles that can go on forever.

I have had other big changes late in my own career. 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.

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, the importance of which seemed compelling to me long before they became the subject of a Nobel prize.

Thus, I have had the privilege of exploring some new ideas that were eventually recognized as worthy of study by students of subsequent generations. This is ultimately greatest honor that any scholar can hope for, of which Nobel prizes can highlight only a small part. I have also worked on many other ideas and questions which excited me equally but which have not attracted much subsequent attention or interest from others. I have regularly shared ideas and questions that intrigue me with colleagues and students, but I do not try to steer students toward these questions, because good research takes a long time and nobody can know what will prove fruitful in research until it is done. Any scholar must make his or her own choices about what questions to study. For me, the greatest prize has been to work in a long career as a professor with opportunities to explore so many ideas, in search of the better understanding that we need.