

HARSANYI'S GAMES WITH INCOMPLETE INFORMATION

by Roger B. Myerson

August 2004

This review of Harsanyi's great three-part paper on "Games with Incomplete Information Played by 'Bayesian' Players" has been written for a special 50th-anniversary issue of Management Science,

Address: Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 60637, USA. Phone: 773-834-9071. Fax: 773-702-8490.

E-mail: myerson@uchicago.edu

<http://home.uchicago.edu/~rmyerson/>

HARSANYI'S GAMES WITH INCOMPLETE INFORMATION

by Roger B. Myerson

The historical context

John Harsanyi's (1967-8) three-part essay "Games with Incomplete Information Played by Bayesian Players" should be read today with a view to its place in the history of economic thought. It is one of the great papers that gave birth to modern information economics. Other outstanding papers in this group include William Vickrey's (1961) paper on auctions, George Akerlof's (1970) "Market for Lemons," Michael Spence's (1973) "Job Market Signaling," and Michael Rothschild and Joseph Stiglitz's (1976) paper on insurance markets. But Harsanyi's contribution holds a unique place in this group of Nobel prize-winners. Each of the other papers analyzed one specific example of a market where people have different information. Harsanyi alone addressed the problem of how to define a general analytical framework for studying all competitive situations where people have different information. The unity and scope of modern information economics was found in Harsanyi's framework.

In the previous literature, as Rothschild and Stiglitz (1976) observed, economic theorists had a long tradition of banishing discussions of information to footnotes. The standard models of economic analysis assumed that all agents had the same information, and questions that went beyond the scope of these models were generally overlooked in any rigorous analysis. Even in the one great paper with a truly modern treatment of information before Harsanyi (1967-8), Vickrey (1961) omitted any mention of informational problems from the introductory overview of his paper.

Game theory was founded as an attempt to broaden the scope of economic analysis. But questions of informational structures also tended to be suppressed in the game-theory literature before 1965. This failure should be recognized as a direct consequence of the first success of game theory: the development of the normal form as a simple and general model of games.

Von Neumann's (1928) first paper on game theory began by defining a general mathematical structure for modeling multi-stage dynamic games, called games in extensive form. These general extensive-form models allow that players may get different information during the course of a game, in which case it is called a game with imperfect information. But these extensive-form models are mathematically complex and difficult to analyze in general. So immediately after his definition of the extensive form, von Neumann (1928) argued that any multi-stage extensive-form game can be reduced to an equivalent one-stage game, called the normal form, where all players act simultaneously and independently.

The key to this reduction is the idea of strategy. A strategy for any player in a game is defined to be a complete contingent plan of action that specifies, for every stage of the game and every possible state of the player's information at this stage, what the player would do at this stage if he got this information. A rational player should be able to choose his entire strategy right at the start of the game. The start of the game is before anybody has any opportunity to try to influence anybody else in the game, and so the players' initial strategy choices should be independent of each other. Once all players have chosen their strategies, the expected outcome of the game becomes a matter of mechanical calculation.

Thus, for any extensive-form game, von Neumann defined an equivalent one-stage game in normal form, where each player independently chooses his strategy for the given extensive-

form game. The mathematical structure of the normal form consists simply of a set of players, a set of strategies for each player, and a payoff function that specifies each player's expected payoff for every possibly combination of strategies.

The idea that this mathematically simple normal form is a general model for all games was the key simplifying insight that allowed the first great advances of modern game theory (see Myerson, 1999). Unfortunately, this reduction to normal form had the effect of suppressing all questions of information in games. From this viewpoint, no player gets any private information until after he has chosen his strategy for the whole game, and then there is nothing to do but mechanically implement his strategic plan.

The term "incomplete information" was used by von Neumann and Morgenstern (1944, p. 30) to refer to a game where part of the normal-form structure is unspecified. They argued against studying such incomplete models, because any serious theoretical investigation must be based on analysis of precisely specified models. They remarked that individuals' uncertainty about events in the game can be modeled by an extensive form with imperfect information.

But questions about incomplete information would not go away. As soon as any normal-form game model was formulated to describe any social situation, it was natural to raise questions about how the analysis would change if some players do not actually know some parameters of this model. An example of unproductive early responses to such questions can be found in Luce and Raiffa's (1957) section 12.4. To avoid the assumption that everybody knows every player's payoff function, Luce and Raiffa considered a generalized normal form for n -person games which includes n^2 payoff functions that describe what each player j believes to be each player i 's payoff function.

Luce and Raiffa immediately acknowledged several difficulties with this approach. They were concerned that it might not be sufficiently general because uncertainty about strategies had not been considered. Even within the domain of uncertainty about payoffs, their model did not address the question of what another player k might believe about player j 's beliefs about player i 's payoffs. This model was not even rich enough to formulate the winner's curse effect, which arises when player i believes that another player j may know more about i 's payoffs than i himself knows. Furthermore, it was not clear how this generalized normal form should be analyzed.

Against this background, Harsanyi (1962) began to wrestle with the problems of extending Nash's (1950,1953) bargaining solution to situations where players do not know each others' payoff functions. In this work, he began to recognize problems of modeling players' beliefs about each others' beliefs in a game. A few years later he moved to confront these modeling problems at the most general and fundamental level. In 1965, he presented the first draft of his theory of games with incomplete information at game theory conferences in Princeton and Jerusalem.

The Jerusalem conference was particularly important because Harsanyi's audience there included Robert Aumann, Michael Maschler, and Reinhard Selten, who had been exploring a variety of related informational questions in their own work. Over the next several years, they all continued to meet and work together in a research project for the U.S. Arms Control and Disarmament Agency (see Aumann and Maschler, 1995, pp. xi-xvi). Problems of information in games became a central focus of this ACDA project, and it provided the peer group in which Harsanyi refined his ideas. So the ACDA project was the forum in which game theorists first became seriously concerned with the study of information in games, a development which has

continued to this day.

Part I: the Basic Model

Harsanyi (1967-8) started in Section 1 with an argument that incomplete information in a game may force us to consider an infinite hierarchy of beliefs. Suppose there is some parameter in our game model that some player i does not know. Then player i must have beliefs about this parameter which, according to Bayesian decision theory, can be mathematically described by a probability distribution over the possible values of this parameter. This probability distribution, which may be called i 's first-order belief, can be important for predicting i 's behavior, and so it must be included in our model of the game. But now we must ask whether everybody knows i 's first-order beliefs. If not, then any other player j must have beliefs about i 's first-order beliefs, which can be described by a probability distribution over the space of probability distributions over the original parameter space. This second-order belief of player j can be important for predicting j 's behavior, and so it must also be included in our model of the game. Third-order and higher-order beliefs may be defined similarly. The point is that, as long as our model includes any parameters that are unknown to some players, we will always find more parameters, the players' subjective probability distributions over the unknowns, which are relevant to predicting players' behavior and are measurable (in principle) and which must therefore be added to our model.

Thus we can avoid working with an underspecified model of the game only if we assume that all the parameters of our game model are known to all the players. This modeling principle implies that if some players are uncertain about a piece of information that other players know,

then this information must be represented, not by a parameter of the model, but by a random variable within the model. The model can include probability distributions that describe what various players believe about these random variables, but the model cannot specify the random variables' actual values. This basic insight underlies Harsanyi's Bayesian games.

Harsanyi began his construction of Bayesian games by methodically arguing that all uncertainty in games can be equivalently modeled as uncertainty about payoff functions. There are many points in Harsanyi's basic argument that have not been as carefully covered by later authors, who tend to assume that people have already learned these ideas from Harsanyi. For example, Harsanyi explains that uncertainty about the feasibility of a player's strategy can be converted into uncertainty about the player's payoffs, by specifying a very negative payoff for the player if he tries to use this strategy when it is supposed to be infeasible. Thus, Harsanyi addressed a basic modeling concern of Luce and Raiffa that was cited above.

But in asking what is the right way to think about games with incomplete information, Harsanyi also had to invest some time in exploring other less-fruitful formulations. For example, in Section 3, Harsanyi initially defined a variable a_{ki} to represent the information that player k has which is relevant to player i 's payoffs in the game. He then devoted several paragraphs to explaining how we should simplify this notation by letting the variable a_k denote the whole vector (a_{k1}, \dots, a_{kn}) , so that the symbol a_k denotes everything that player k knows that affects the payoffs of any of the n players. Next, Harsanyi introduced, for each player k , another variable b_k that denotes the other information that player k has that could affect his beliefs about what the other players know. Finally, Harsanyi introduced the symbol $c_k = (a_k, b_k)$ to denote all of player k 's private information in the game, and he called this c_k variable player k 's information vector or

attribute vector or type.

Today, of course, nobody needs to develop any such preliminary notation in a game-theory paper, because we have all learned from Harsanyi (1967-8) that a model of incomplete information should begin with a specification of the possible types for each player. The key to the simplification is Harsanyi's idea that each player's type should be defined to summarize everything that the player knows privately at the beginning of the game which could affect his beliefs about payoffs in the game and about all other players' types.

Actually, the simple term "type" itself was a late addition to the paper which Harsanyi adopted at the suggestion of Aumann and Maschler, who used it in their own work on repeated games with incomplete information (Aumann and Maschler, 1995). The term appears only in Part I, which Harsanyi revised last in June 1967. The longer equivalent phrase "attribute vector" is used in Parts II and III, which were not revised after May 1966. But whenever it came into the paper, Harsanyi made this concept of "type" an essential part of game theorists' vocabulary.

Harsanyi then argued that, without loss of generality, we can consider payoff functions that depend on the actions and types of all players but do not depend on any other random variables, because the rationality criteria for each player's actions depend only on his conditional expected payoff given his information. Any parameters that are unknown to all players can be removed from their payoff functions by computing, for each possible combination of types and actions, the expected payoffs with respect to the unknown parameters' conditional probability distribution given the players' types.

So Harsanyi defined an I-game or a Bayesian game with incomplete information to be a mathematical model consisting of (1) a set of players, (2) a set of possible actions for each player,

(3) a set of possible types for each player, (4) a payoff function that specifies each player's expected payoff for every possible combination of all players' actions and types, and (5) a probability function that specifies, for each possible type of each player, a probability distribution over the other players' possible types, describing what each type of each player would believe about the other players' types. This definition of Bayesian games has become one of the basic standard models in the game-theory literature.

In Section 4, Harsanyi introduced an alternative model which differs only in that the probability function specifies just one probability distribution, over the set of all possible combinations of all players' types. He referred to this alternative model as a C-game. From such a C-game, Bayes's rule yields an equivalent I-game that has the same sets of players, actions, and types, and the same payoff functions. In this equivalent I-game, for any type t_i of any player i , each possible combination of other players' types $(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n)$ should be assigned a subjective probability that is equal to the conditional probability of $(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n)$ given that i 's type is t_i . That is, if the C-game specifies that probability of any type profile (t_1, \dots, t_n) is $P(t_1, \dots, t_n)$, and $P_i(t_i)$ is the marginal probability of player i 's type being t_i in this C-game, then the equivalent I-game should specify that type t_i of player i assigns the subjective probability

$$R_i(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n | t_i) = P(t_1, \dots, t_n) / P_i(t_i)$$

to the event of $(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n)$ being the profile of other players' types.

Conversely, when we are given an I-game with probability functions that specify these $R_i(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n | t_i)$ conditional probabilities, we may say that the players' beliefs are consistent with a common prior if there exists a prior probability distribution P on the set of all possible combinations of all players' types such that these R_i type-conditional beliefs can be

derived from this prior P by Bayes's rule. So an I-game is equivalent to some C-game iff the players' beliefs in the I-game are consistent with a common prior.

Harsanyi discussed several different stories about how the players' types might be generated in any Bayesian game. We might suppose that the types are determined by an initial lottery, according to the prior probability distribution P in the C-game, after which each player is informed of his own type but remains uncertain about the other players' types. But Harsanyi also suggested that we might instead imagine the existence of a separate agent for each possible type of each player in the game, and then the initial lottery would determine which of these type-agents actually get to act as players in the game. The latter interpretation has the obvious disadvantage of hypothesizing the existence of many fictitious type-agents who do not actually appear in the game. But Harsanyi suggested this story because it has the advantage of rendering meaningless any questions about the expected payoff of a player before he learns his type. Harsanyi's point here is that the type represents what the player knows at the beginning of the game, and so calculations of the player's expected payoff before this type is learned cannot have any decision-theoretic significance in the game.

Harsanyi recognized that we could reduce his C-game model to a conventional normal-form model. To do so, we would define the normalized strategies of each player to be functions from his set of types to his set of actions. Then for any combination of normalized strategies, we could compute each player's ex ante expected payoff, using the whole probability distribution P over the possible combinations of types.

But for Harsanyi, this normal form would be going too far, because it suppresses all the interesting informational questions and it adds nothing to the analysis. For example, if a player's

type includes a specification of his or her gender (about which some other players are uncertain), then the normal-form analysis would require us to imagine the player choosing a contingent plan of what to do if male and what to do if female, maximizing the average of male and female payoffs. But there is no need to force such absurd pre-birth calculations of expected payoff into our analysis. In providing us with a general definition of Bayesian games with incomplete information, Harsanyi was teaching us to study games where players may have different information right from the start of the game. In effect, Harsanyi defined Bayesian games by following von Neumann's construction of the normal form just as far as he could without suppressing any initial informational differences among the players.

Harsanyi also emphasized that, for game-theoretic analysis, we must take the perspective of a properly informed observer who knows only the information common to all the players in the game. The actual type of each player, being private information, must be treated as an unknown quantity in our game-theoretic analysis. Even if we are consultants to a particular player j who knows his actual type, we need to analyze the decision problems of other players, who do not know this type and so must formulate conjectures about what each possible type of player j would do. Thus, game-theoretic analysis requires that we deny ourselves any knowledge of any player's actual type, so that we can appreciate the uncertainty of all the other players who do not know it.

Our model of a game is a summary of what we know about the game. So this basic methodological assumption, that everything we know in our model is also known to all players, implies that the players know that they all know the model, and that the players know that they all know that they all know the model, and so on. Thus we are led to the assumption that our

Bayesian game model is common knowledge among the players (in the terms of Aumann, 1975).

Part II: Bayesian equilibrium

In Part II, Harsanyi (1967-8) defined the concept of Bayesian equilibrium as the basic noncooperative solution concept for Bayesian games. A Bayesian equilibrium specifies a (possibly randomized) action for each possible type of each player, such that each type's specified action maximizes his conditional expected payoff given his type, given his beliefs about the other players' types, and given the type-contingent behavior of all other players according to this equilibrium. For a consistent C-game, which could be (unsatisfactorily) reduced to normal form, Harsanyi noted that a Bayesian equilibrium is equivalent to an equilibrium in Nash's (1951) original sense. He also asserted the general existence of Bayesian equilibria for finite Bayesian games.

Harsanyi then analyzed the Bayesian equilibria of several bargaining games with incomplete information. His first example showed that players' incomplete information about each others' types may affect their rational equilibrium behavior. His second example showed that the equivalence of equilibrium strategies and maximin strategies for two-person zero-sum games may fail for Bayesian games with incomplete information (if the pessimism underlying player 1's maximin strategy includes an assumption that player 2's strategy will be the worst for 1's actual type).

In a third example, Harsanyi considered a divide-the-dollar game for two players where each has two possible types: a strong type that can keep all the money that he earns in the game and a weak type that must pay half of his earnings to some secret organization. Each player is

equally likely to be strong or weak, but each type of each player assigns a very high probability (almost 1) to the event that the other player is the opposite type. So if a player is weak, then he is almost sure that the other player is strong, and vice versa. With this example, Harsanyi showed the subtle difficulties of extending cooperative solution concepts like the Nash (1950, 1953) bargaining solution to games with incomplete information. When we look at the players' expected payoffs ex ante, before anybody learns his type, the Nash bargaining solution is the unique symmetric point on the Pareto-efficient frontier, which is achieved by the equilibrium where a strong type would claim the entire dollar, while a weak type would claim nothing. This Bayesian equilibrium yields an ex ante expected payoff of almost 0.50 to each player, before he learns his type. But when a player knows that he is weak, he would be unlikely to settle for such a solution, because claiming nothing is a weakly dominated strategy for him.

Nash (1953) recognized that his bargaining solution was one of many equilibria of the divide-the-dollar game, but he formulated an ingenious argument to show that his bargaining solution could be the most stable equilibrium. This argument breaks down in Harsanyi's example when the players learn their types before they make their payoff demands, thus revealing the fragility of Nash's noncooperative justification of his bargaining solution. But Nash (1950) originally justified his bargaining solution by axioms that included a scale-invariance axiom, and the two types in this example differ only in that the weak type's payoffs are multiplied by $1/2$. So an obvious extension of Nash's axioms for this example would suggest that the each player should claim the same amount, half of the dollar, regardless of his type. From the ex ante perspective, this solution appears to be Pareto-dominated, because each player's expected payoff is 0.375 before he learns his type. But Harsanyi has already taught us that ex-ante expected

payoffs are meaningless. Instead we should consider each player's conditional expected payoff given his type, for each possible type. These conditional expected payoffs of all types together constitute the interim allocation vector of Holmstrom and Myerson (1984). The equal-division equilibrium yields an interim allocation vector (.5, .25, .5, .25) which is indeed on the interim-efficient frontier for this example.

Part III: Consistent priors

Harsanyi's main concern in Part III is with the question of when a Bayesian game with incomplete information (an I-game) can be modeled by a C-game with a consistent common prior. This question is equivalent to the question of when the players' beliefs about each others' types can be derived by Bayes's rule from a common prior.

Harsanyi recognized that I-games can easily be constructed such that the players' beliefs are not consistent with any common prior. For example, consider a two-person game where each player's type can be weak or strong, and it is common knowledge that either type of player 1 would assign probability 0.9 to the event that 2's type is the same as 1's type, while either type of player 2 would assign probability 0.5 to the event that their types are the same. These beliefs cannot be derived from any common prior probability distribution on the four possible type pairs.

But Harsanyi felt that there was something unsatisfactory about models that assume such inconsistent beliefs. He tried to argue that, in any real situation, a reasonable game model should have a common prior, so that we can use C-games without loss of generality in economic analysis. This argument has become known as the Harsanyi doctrine.

But we can analyze I-games just as well as C-games, and so a failure of this Harsanyi

doctrine would cause no difficulties for game-theoretic analysis. In this regard, it seems unfortunate that Harsanyi (1967-8) formally stated the definition of Bayesian equilibrium only in the context of the C-game model with its consistent common prior. Perhaps Harsanyi may have thought that a definition of equilibrium should be limited to games that can be reduced to the normal form which Nash originally studied, so that we can talk about equivalence to Nash's definition. But any reader can see that Harsanyi's definition of Bayesian equilibrium can be applied just as well to I-games without any common prior. As Harsanyi has emphasized, the rationality of each player's strategy in an equilibrium depends only on his type-conditional expected payoffs, which are calculated by the same formula in I-games and C-games.

In fact, the consistent common prior assumption has become virtually universal in economic analysis, mainly because applied economists have found that inconsistent models are more difficult to motivate. To understand why, we must first recognize that applied economic modeling becomes most interesting when it provides a new explanation for some observed behavior that seemed surprising or paradoxical. So games with incomplete information become interesting in applied economic analysis when differences in people's beliefs can account for behavior that would be otherwise difficult to explain. But such an explanation naturally begs the question of what might have caused these critical differences in people's beliefs. If our Bayesian game model has a consistent common prior, then these differences of belief can be entirely explained by differences in the players' past experiences which have given them different information. On the other hand, if there is no common prior, then we can only say that these difference in people's beliefs are just a fundamental assumption of our model. But then we must face the question: if we can assume any arbitrary characteristics for the individuals in our model,

then why could we not explain the surprising behavior even more simply by just assuming that each individual has a payoff function that is maximized by this behavior? Thus, to avoid such trivialization of the economic problem, applied theorists have generally limited themselves to models that satisfy Harsanyi's consistency assumption.

At some point in Part III, a reader may sense that the meaning of "incomplete information" may have evolved over the course of this work. Von Neumann and Morgenstern's (1944) original meaning of "incomplete information" was that some parameters that are necessary for game-theoretic analysis have been left unspecified. But of course, we cannot do serious theoretical analysis of an underspecified model. So at the beginning of Part I, the term "incomplete information" was applied instead to fully specified models where some players might not know some parameters of the model. But Harsanyi argued that, under the assumptions of Bayesian decision theory, if any player is uncertain about some parameter, then we must add to our model a specification of the players' beliefs about this parameter, and then we should integrate the parameter itself out of our analysis when we define the players' type-conditional expected payoff functions. Thus, Harsanyi banished from our models any parameters that are not common knowledge among the players. With this modeling methodology, what can it now mean to say that a game has "incomplete information"?

The answer seems to be that, in Harsanyi's advanced sense, a game has incomplete information if, at the beginning of the game, some players have incomplete information about what other players know or believe. If the players' beliefs in an I-game are consistent with a common prior, then we can construct an equivalent C-game model which has complete information in the sense that we can imagine the game beginning at a point in time before the

players learn their types. But Harsanyi has taught us that this ex-ante common-prior position is not essential for noncooperative analysis, and it may even be seriously misleading for cooperative welfare analysis. So when we read Harsanyi's Part III, it may be more useful to focus more on the distinction between consistency or inconsistency of beliefs, rather than between completeness or incompleteness of information. That is, instead of reading "C-games" and "I-games" as games with "complete" and "incomplete" information (as Harsanyi originally suggested in Part I), it may be more useful to reinterpret Harsanyi's C-games as Bayesian games where beliefs are Consistent with a common prior, while Harsanyi's I-game model allows that beliefs may be Inconsistent with any common prior.

Conclusion

Harsanyi gave us a general analytical methodology for studying situations where people have different information. In this methodology, we always formulate a game model that is assumed to be common knowledge among the players, but we do not assume that all players have the same information. In Harsanyi's Bayesian games, the players' different information is described by a collection of random variables, called the players' types, each of which is the private information of one player. The actual value of each players' type is omitted from the model, which includes instead a precise probabilistic description of what each type of each player would believe about the other players' types. In his definition of Bayesian games, Harsanyi followed von Neumann's simplifying construction of the normal form as far as possible without suppressing initial informational differences among the players. It may seem simple now, but at first it was hard to see.

Harsanyi himself sometimes seemed uncomfortable with this common-knowledge approach. At various points throughout the paper, even in the last section of Part III, Harsanyi sometimes referred to a Bayesian game model as having been assessed only by some particular player j . But if we take this qualification seriously, then we find ourselves dragged back to the beginning of the Part I again. After all, if we only know that player j thinks that this model describes the game, then we must ask what do other players believe about j 's model, and what does j believe about other players' models? To escape from this infinite regress, the reader has no choice but to overlook Harsanyi's parenthetical remarks about a game being only "assessed by player j ". If our model is j 's model, then we must assume that j 's model is correct and common knowledge among the players. But within this methodological framework, we can admit as much uncertainty as may seem appropriate in any situation by enlarging the set of types that represent one player's private information and other players' uncertainty.

There is something fundamentally counter-intuitive about the art of modeling with Bayesian games. To describe a situation where players have great uncertainty about others' information, we must formulate a Bayesian model where their sets of possible types are large, and then we must assume that this large model is common knowledge among the players. In this sense, more uncertainty apparently requires us to assume that more is common knowledge. The mathematical difficulties of this modeling dilemma have been studied by Mertens and Zamir (1985) to establish the general applicability of Harsanyi's approach.

But in fact Harsanyi's approach has proven its great practical value as a general framework for applied economic modeling. Since his great three-part study of incomplete information was published in Management Science, its influence has been the basis for a

profound revolution in social science.

Any academic discipline must rely on a general methodology to provide a framework for inquiry and debate. Academic methodologies enable scholars to see connections that may be obscure to the untrained layman. But scholars must also be aware that our expertise is diminished beyond the scope of our methodology, and we learn to stay within its boundaries. Thus, Vickrey's (1961) study of auction games was under-appreciated when it first appeared, because its advanced informational assumptions placed it beyond the standard analytical frameworks of economic analysis. After Harsanyi's (1967-8) great study of incomplete information, Vickrey's auction games, Akerlof's (1970) market for lemons, and Spence's (1973) labor markets, and Rothschild and Stiglitz's (1976) insurance markets could all be viewed as interesting examples of people playing Harsanyi's Bayesian games. Having this common framework for such informational problems enabled us to apply insights from the study of any one of them to all the others, and thus the new economics of information was born.

REFERENCES

- George Akerlof, "The market for lemons: quality uncertainty and the market mechanism," Quarterly Journal of Economics 84 (1970), 488-500.
- Robert J. Aumann, "Agreeing to disagree," Annals of Statistics 4 (1976), 1236-39.
- Robert J. Aumann and Michael Maschler, Repeated Games with Incomplete Information, MIT Press (1995).
- John C. Harsanyi, "Bargaining in ignorance of the opponent's utility function," Journal of Conflict Resolution 6 (1962b), 29-38.
- John C. Harsanyi "Games with incomplete information played by Bayesian players," Management Science 14 (1967-1968), 159-182, 320-334, 486-502.
- Bengt Holmstrom and Roger B. Myerson, "Efficient and Durable Decision Rules with Incomplete Information," Econometrica 51 (1983), 1799-1819.
- R. Duncan Luce and Howard Raiffa, Games and Decisions Wiley (1957).
- Jean-Francois Mertens and Shmuel Zamir, "Formulation of Bayesian analysis for games with incomplete information," International Journal of Game Theory 14 (1985). 1-29.
- Roger B. Myerson, "Nash Equilibrium and the History of Economic Theory," Journal of Economic Literature 37 (1999), 1067-1082.
- John Nash, "The bargaining problem," Econometrica 18 (1950), 155-162.
- John Nash, "Non-cooperative games," Annals of Mathematics 54 (1951), 286-295.
- John Nash, "Two-person cooperative games," Econometrica 21 (1953), 128-140.

John von Neumann, "Zur Theories der Gesellschaftsspiele." Mathematische Annalen 100 (1928), 295-320. English translation by S. Bergmann in R. D. Luce and A. W. Tucker, eds., Contributions to the Theory of Games IV (1959), pp. 13-42, Princeton U. Press.

John von Neumann, and Oskar Morgenstern, Theory of Games and Economic Behavior. Princeton University Press (1944).

Michael Rothschild and Joseph Stiglitz, "Equilibrium in competitive insurance markets: an essay on the economics of imperfect information," Quarterly Journal of Economics 90 (1976), 629-649.

Michael Spence, "Job market signaling," Quarterly Journal of Economics 87 (1973), 355-374.

William Vickrey, "Counterspeculation, auctions, and competitive sealed tenders," Journal of Finance 16 (1961), 8-37.