The Nash equilibrium: A perspective

Charles A. Holt* and Alvin E. Roth

Department of Economics, University of Virginia, Charlottesville, VA 22904-4182; and Department of Economics and Harvard Business School, Harvard University, Cambridge, MA 02138

In 1950, John Nash contributed a remarkable one-page PNAS article that defined and characterized a notion of equilibrium for *n*-person games. This notion, now called the "Nash equilibrium," has been widely applied and adapted in economics and other behavioral sciences. Indeed, game theory, with the Nash equilibrium as its centerpiece, is becoming the most prominent unifying theory of social science. In this perspective, we summarize the historical context and subsequent impact of Nash's contribution.

In a brief 1950 communication to PNAS (1), John Forbes Nash formulated the notion of equilibrium that bears his name and that has revolutionized economics and parts of other sciences. Nash, a young mathematics graduate student at Princeton, was a part of the Camelot of game theory centered around von Neumann and Morgenstern. They had written *Theory of Games and Economic Behavior* (2) to expand economic analysis to allow economists to model the "rules of the game" that influence particular environments and to extend the scope of economic theory to include strategic small-group situations in which each person must try to anticipate others' actions. von Neumann and Morgenstern's definition of equilibrium for "noncooperative" games was largely confined to the special case of "two-person zero-sum" games, in which one person's gain is another's loss, so the payoffs always sum to zero (3). Nash proposed a notion of equilibrium that applied to a much wider class of games without restrictions on the payoff structure or number of players (1, 4, 5). von Neumann's reaction was polite but not enthusiastic.† Nevertheless, the Nash equilibrium, as it has become known, helped produce a revolution in the use of game theory in economics, and it was the contribution for which Nash was cited by the Nobel Prize committee at the time of his award, 44 years later.

Equilibrium Points in *n*-Person Games

The first part of the 1950 PNAS paper introduces the model of a game with *n* participants, or "players," who must each select a course of action, or "strategy":

> One may define a concept of an *n*-person game in which each player has a finite set of pure strategies and in which a definite set of payments to the *n* players corresponds to each *n*-tuple of pure strategies, one strategy being taken for each player.
> (ref. 1, p. 48)

The notion of a strategy is quite general, and it includes "mixed" strategies that are probability distributions over decisions, e.g., an inspector who audits on a random basis or a poker player who sometimes bluffs. Another interpretation of a mixed-strategy is that of a population of randomly matched individuals in the role of each player of the game, some proportion of whom make each of a number of available choices. The idea of the Nash equilibrium is that a set of strategies, one for each player, would be stable if nobody has a unilateral incentive to deviate from their own strategy:

> Any *n*-tuple of strategies, one for each player, may be regarded as a point in the product space obtained by multiplying the *n* strategy spaces of the players. One such *n*-tuple counters another if the strategy of each player in the countering *n*-tuple yields the highest obtainable expectation for its player against the $n-1$ strategies of the other players in the countered *n*-tuple. A self-countering *n*-tuple is called an equilibrium point.
> (ref. 1, p. 49)

That is, a Nash equilibrium is a set of strategies, one for each of the *n* players of a game, that has the property that each player's choice is his best response to the choices of the $n-1$ other players. It would survive an announcement test: if all players announced their strategies simultaneously, nobody would want to reconsider. The Nash equilibrium has found many uses in economics, partly because it can be usefully interpreted in a number of ways.

When the goal is to give advice to all of the players in a game (i.e., to advise each player what strategy to choose), any advice that was not an equilibrium would have the unsettling property that there would always be some player for whom the advice was bad, in the sense that, if all other players followed the parts of the advice directed to them, it would be better for some player to do differently than he was advised. If the

advice is an equilibrium, however, this will not be the case, because the advice to each player is the best response to the advice given to the other players. This point of view is sometimes also used to derive predictions of what players would do, if they can be approximated as "perfectly rational" players who can all make whatever calculations are necessary and so are in the position of deriving the relevant advice for themselves.

When the goal is prediction rather than prescription, a Nash equilibrium can also be interpreted as a potential stable point of a dynamic adjustment process in which individuals adjust their behavior to that of the other players in the game, searching for strategy choices that will give them better results. This point of view has been productive in biology also: when mixed strategies are interpreted as the proportion of a population choosing each of a set of strategies, game payoffs are interpreted as the change in inclusive fitness that results from the play of the game, and the dynamics are interpreted as population dynamics (6, 7). No presumptions of rationality are made in this case, of course, but only of simple self-interested dynamics. This evolutionary approach has also been attractive to economists (e.g., ref. 8).

A third interpretation is that a Nash equilibrium is a self-enforcing agreement, that is, an (implicit or explicit) agreement that, once reached by the players, does not need any external means of enforcement, because it is in the self interest of each player to follow

the agreement if the others do. Viewed in this way, the Nash equilibrium has helped to clarify a distinction sometimes still made between "cooperative" and "noncooperative" games, with cooperative games being those in which agreements can be enforced (e.g., through the courts), and noncooperative games being those in which no such enforcement mechanism exists, so that only equilibrium agreements are sustainable. One trend in modern game theory, often referred to as the "Nash program," is to erase this distinction by including any relevant enforcement mechanisms in the model of the game, so that all games can be modeled as noncooperative. Nash took initial steps in this direction in his early and influential model of bargaining as a cooperative game (9) and then as a noncooperative game (10).

Nash's 1950 PNAS paper not only formulated the definition of equilibrium but also announced the proof of existence that he obtained using Kakutani's (11) fixed point theorem. This technique of proof subsequently became standard in economics, e.g., the notion of a competitive equilibrium as a vector of anticipated prices resulting in production and consumption decisions that generate the same vector of prices. In a personal communication to one of the authors, Nash remarked, "I know that S. Kakutani's generalized fixed point theorem was actually inspired to improve on some arguments made by von Neumann in an economic context in the 1930s."

Nash shared the 1994 Nobel Prize with John Harsanyi and Reinhard Selten. Harsanyi was cited for extending the Nash equilibrium to the larger class of games called games of incomplete information, in which players need not be assumed to know other players' preferences and feasible choices (12). Selten was cited for his work on equilibrium refinements, which takes the point of view that the requirements of the Nash equilibrium are necessary conditions for advice to perfectly rational players but are not sufficient conditions, and there may be superfluous equilibria that can be removed from consideration by appropriate refinements that focus attention on a nonempty subset of Nash equilibria (13, 14). The Nash equilibrium has been extended, refined, and generalized in other directions as well. One noteworthy generalization of mixed strategy equilibrium is "correlated equilibrium" (15), which considers not only independently randomized strategies for each player but also jointly randomized strategies that may allow coordination among groups of players.

Equilibrium and Social Dilemmas

The Nash equilibrium is useful not just when it is itself an accurate predictor of how people will behave in a game but also when it is not, because then it identifies situations in which there is a tension between individual incentives and other motivations. A class of problems that have received a good deal of study from this point of view is the family of "social dilemmas," in which there is a socially desirable action that is not a Nash equilibrium. Indeed, one of the first responses to Nash's definition of equilibrium gave rise to one of the best known models in the social sciences, the Prisoners' Dilemma. This model began life as a simple experiment conducted in January 1950 at the Rand Corporation by mathematicians Melvin Dresher and Merrill Flood, to demonstrate that the Nash equilibrium would not necessarily be a good predictor of behavior. Each of the two players in that game had to choose one of two decisions, which, for expositional purposes, we will call "cooperate" or "defect." The game specifies the payoffs for each player for each of the four possible outcomes: (cooperate, cooperate), (cooperate, defect), (defect, cooperate), and (defect, defect). The payoffs used were such that each player's best counter to either of the other's choices was to defect, but both players would earn more if they both cooperated than if they both chose their equilibrium decision and defected.

Nash's thesis advisor, Albert Tucker, was preparing a talk on recent developments in game theory to be given to the Stanford Psychology Department when he saw the Dresher and Flood payoff numbers on a blackboard at the Rand Corporation. Tucker then devised the famous story of the dilemma faced by two prisoners who are each given incentives by the prosecutor to confess, even though both would be better off if neither confesses than if they both do (16, 17). In the initial experiment (18) and in innumerable experiments that followed, players often succeed, at least to some degree, in cooperating with one another and avoiding equilibrium play (19).[‡]

Social scientists across many disciplines have found prisoner's dilemmas helpful in thinking about phenomena ranging from ecological degradation (20) to arms races. What the Nash equilibrium makes clear, even in a game like the prisoner's dilemma in which it may not be an accurate point predictor, is that the cooperative outcome, because it

_____

[‡]H. Raiffa independently conducted experimens with a Prisoner's Dilemma game in 1950, but he did not publish them (see ref. 19).

is not an equilibrium, is going to be unstable in ways that can make cooperation difficult to maintain. This observation has been confirmed in many subsequent experiments on this and more general "social dilemmas" (see, e.g., refs. 21–23). You can put yourself into a social dilemma game by going to the link: http://veconlab.econ.virginia.edu/tddemo.htm and playing against decisions retrieved from a database. This Traveler's Dilemma game is somewhat more complex than a prisoner's dilemma, in that the best decision is not independent of your beliefs about what strategy might be selected by the other player (24).

Design of Markets and Social Institutions

One of the ways in which research on dilemmas and other problems of collective action has proceeded is to look for the social institutions that have been invented to change games from prisoner's dilemmas to games in which cooperation is sustainable as an equilibrium; see e.g., Elinor Ostrom's 1998 presidential address to the American Political Science Association (25). For example, just as firms selling similar products may undercut each other's price until price is driven down to cost, it is possible for a series of actions and reactions to force players in a game into a situation that is relatively bad for all concerned, which provides strong incentives for restrictions on unilateral actions. This kind of "unraveling" is encountered in some labor markets in which employers may try to gain an advantage by making early offers. In the market for federal appellate court clerks, for example, positions began to be arranged earlier and earlier, as some judges tried to hire clerks just before their competitors. This continued until offers (for jobs that would begin only on graduation from law school) were being made to law students 2 years in advance, only on the basis of first-year law school grades (see ref. 26). This situation was widely viewed as unsatisfactory, because it forced both judges and law students to make decisions far in advance, on the basis of too little information. The most recent of many attempts to reform this market took the form of a year-long moratorium on the hiring of clerks by appellate judges, which ended the day after Labor Day 2003, with only third-year law students to be hired. It is still too early to know whether this relatively mild intervention will finally solve the unraveling of the law clerk market. But a moratorium by itself does not change the rules of the game sufficiently to alter the dilemma-like properties of the equilibrium, and so we predict that fur-

Holt and Roth

ther changes will be needed if the familiar problems are to be avoided in the long term.

An alternative solution to this prisoner's dilemma problem of timing in such markets is to arrange an organized clearinghouse in which both employers and job candidates participate. Such clearinghouses arose, for example, as a result of situations in which medical students were being hired well over 1 year before graduation; this happened in the U.S. in the 1940s and in the U.K. in the 1960s. Today, graduates of American medical schools (and others applying for residencies at American hospitals) submit rank order lists of preferred positions to a clearinghouse called the National Resident Matching Program. Roth (27, 28) has studied these U.S. and U.K. matching markets. It turns out that one important factor in whether such a labor market clearinghouse succeeds or fails is whether the clearinghouse is designed so that it is a Nash equilibrium for applicants and employers to participate in a way that produces a matching of workers to jobs that is stable, in the sense that no employer and applicant who are not matched to one another would both prefer to be (29, 30).

Participation in such a clearinghouse is even more straightforward if the clearinghouse is constructed so that it is a Nash equilibrium for applicants to simply put down their true preferences, regardless of how likely they think they are to receive each of the jobs for which they have applied, or how other applicants are ranking those jobs. For matching markets like these entry level labor markets, this kind of equilibrium is possible for an appropriately designed clearinghouse. Thus the current version of the National Resident Matching Program algorithm, designed by Roth and Peranson (31), has the property that applicants can confidently be advised it is in their best interest to submit rank order lists of residencies that correspond to their true preferences. That game theorists have started to play a role in designing such clearinghouses and other markets is an indication of how game theory has grown from a conceptual to a practical tool.

Auctions are another kind of market in which it is becoming increasingly common for game theorists to be asked for design advice (see, e.g., refs. 32 and 33). And the economic theory of auctions [for which a Nobel Prize was given to William Vickrey in 1996, in large part for his seminal 1961 paper (34)] is a perfect example of how game theory and the Nash equilibrium have changed economics. Before game theory, economists often analyzed markets simply in

terms of the supply and demand of the goods to be sold, with no way to discuss the rules of the game that make one kind of auction different from another or make auctions different from other kinds of markets (such as stock markets or shopping malls). Today, that discussion is most often carried forward by analyzing the Nash equilibria of the auction rules.

Experimental Economics

It is worth mentioning that Nash both commented on and participated in early experiments in economics (see ref. 35). It was a natural progression to move from a fascination with mathematical models of strategic behavior to the observation of decisions made by people who are playing for real money payoffs under controlled conditions in the laboratory. Indeed, as game theory started its move to the forefront of economic theory, it generated scores of testable predictions and helped lay the groundwork for the introduction of experimental methods into economics (36, 37).

The increasing use of experimental methods in economics and the growing interaction between economics and psychology was itself recognized by the 2002 Economics Nobel Prize that was awarded to a psychologist, Daniel Kahneman (see, e.g., ref. 38) and an economist, Vernon Smith (see, e.g., ref. 39). Before Smith's experiments, it was widely believed that the competitive predictions of supply/demand intersections required very large numbers of well-informed traders. Smith showed that competitive efficient outcomes could be observed with surprisingly small numbers of traders, each with no direct knowledge of the others' costs or values. An important developing area of game theory is to explain these and other experimental results in the context of well-specified dynamic models of the interaction of strategic traders.

Another emerging connection between game theory and experimentation is the increased use of experimental methods in teaching. A well designed classroom experiment shows students that the seemingly abstract equilibrium models can have surprising predictive power. The Internet makes it much easier to run complex games with large groups of students. For example, >30 different types of games, auctions, and markets can be set up and run from a site, (http://veconlab.econ.virginia.edu/admin.htm) that also provides sample data displays from classroom experiments. Most of the data displays and dynamically generated data graphs have options for hiding the relevant Nash predictions when the results are being

discussed and then showing the Nash predictions subsequently.

Modeling Learning and Stochastic Equilibrium

Experimentation has helped move game theorists to focus on approaches that are able to predict how people actually behave when the perfect foresight and perfect rationality assumptions of classical game theory are not satisfied. The experimental literature is full of examples both of games in which observed behavior quickly converges to equilibrium behavior and games in which equilibrium is a persistently poor predictor. This has helped to reinforce the trend, already apparent in the theoretical literature, to extend the static, often deterministic, formulation of equilibrium and consider dynamic and stochastic models.

Experiments make clear that players often do not conform to equilibrium behavior when they first experience a game, even if it is a game in which behavior quickly converges to equilibrium as the players gain experience. One reaction to this has been to develop models of learning that converge to equilibrium in the limit, starting from nonequilibrium behavior (see, e.g., refs. 40 and 41). Another has been the beginning of attempts to develop models of learning that can predict observed behavior in simple experimental games (see, e.g., ref. 42). In particular, learning models are useful for explaining patterns of adjustment, e.g., whether prices converge from above or below, as well as the ultimate steady-state distributions (e.g., refs. 43 and 44).

Some games are played only once, e.g., the exact strategic environments in many military, legal, and political conflicts are unique to the particular time and place. In that case, there is no history that can be used to form precise predictions about others' decisions. Therefore, learning must occur by introspection, or thinking about what the other person might do, what they think you might do, etc. Such introspection is likely to be quite imprecise, especially when thinking about others' beliefs or their beliefs about your beliefs. There has been some recent progress in formulating models of noisy introspection, which can then be used to predict and explain "non-Nash" behavior in experiments using games played only once (24, 45).

If a game is repeated, e.g., with random matchings from a population of players, some noise may persist even after average tendencies have stabilized. The "quantal response equilibrium" is based on the idea that players' responses to differences in expected payoffs are

sharper when such differences are large and are more random when such differences are small (see ref. 46 for an existence proof and ref. 47 for application to bidding in an auction). This notion of equilibrium is a generalization of the Nash equilibrium in the sense that the quantal response predictions converge to a Nash equilibrium as the noise is diminished. But the effect of nonnegligible noise is not merely to spread decisions around Nash predictions; strategic interactions cause feedbacks in some games that magnify and distort the effects of noise. This approach has been used to explain data from some laboratory experiments in which observed behavior deviates from a unique Nash equilibrium and ends up on the opposite side of the set of feasible decisions (24, 43).

Still another approach seeks to reconcile experimental evidence and equilibrium predictions by considering how those predictions would differ if systematic regularities in participants' preferences were modeled. One lesson that consistently emerges from small-group bargaining experiments is that people are often as concerned with fairness issues as they are with their own payoffs (see, e.g., ref. 48). The incorporation of fairness and other notions of nonselfish preferences into standard models often brings economic game theory into contact with evolutionary explanations of human behavior (see, e.g., refs. 49 and 50).

Nash's Contributions in Perspective

In the last 20 years, the notion of a Nash equilibrium has become a required part of the tool kit for economists and other social and behavioral scientists, so well known that it does not need explicit citation, any more than one needs to cite Adam Smith when discussing competitive equilibrium. There have been modifications, generalizations, and refinements, but the basic equilibrium analysis is the place to begin (and sometimes end) the analysis of strategic interactions, not only in economics but also in law, politics, etc. The Nash equilibrium is probably invoked as often in small-group (and not-so-small-group) situations as competitive equilibrium is used in large markets. Students in economics classes today probably hear John Nash's name as much as or more than that of any economist.

In the half century after the publication of Nash's PNAS paper, game theory moved into center stage in economic theory. Game theory has also become part of a lively scientific conversation with experimental and other empirical scientists and, increasingly, the source of practical advice on the design of markets and other economic environments. Looking ahead, if game theory's next 50 years are to be as productive, the challenges facing game theorists include learning to incorporate more varied and realistic models of individual behavior into the study of strategic behavior and learning to better use analytical, experimental, and computational tools in concert to deal with complex strategic environments.

1. Nash, J. F. (1950) *Proc. Natl. Acad. Sci. USA* **36,** 48–49.
2. von Neumann, J. & Morgenstern, O. (1944) *Theory of Games and Economic Behavior* (Princeton Univ. Press, Princeton).
3. von Neumann, J. (1928) *Math. Annal.* **100,** 295–320.
4. Nash, J. F. (1951) Ph.D. thesis (Princeton University, Princeton).
5. Nash, J. F. (1951) *Ann. Math.* **54,** 286–295.
6. Maynard-Smith, J. (1974) *J. Theor. Biol.*, **47,** 209–221.
7. Hoffbauer, J. & Sigmund, K. (1988) *The Theory of Evolution and Dynamical Systems* (Cambridge Univ. Press, Cambridge, U.K.).
8. Weibull, J. W. (1995) *Evolutionary Game Theory* (MIT Press, Cambridge, MA).
9. Nash, J. F. (1950) *Econometrica* **18,** 155–162.
10. Nash, J. F. (1953) *Econometrica* **21,** 128–140.
11. Kakutani, S. (1941) *Duke Math. J.* **8,** 457–459.
12. Harsanyi, J. (1967–68) *Manage. Sci.* **14,** 159–182, 320–334**,** 486–502.
13. Selten, R. (1965) *Z. Gesamte Staatswissenschaft* **121,** 301–324, 667–689.
14. Selten, R. (1975) *Int. J. Game Theor.* **4,** 25–55.
15. Aumann, R. J. (1974) *J. Math. Econ.* **1,** 67–96.
16. Straffin, P. D., Jr. (1980) *UMAP J.* **1,** 102–103.
17. Tucker, A. W. (1950) *UMAP J.* **1,** 101.
18. Flood, M. M. (1958) *Manage. Sci.*, **5,** 5–26.
19. Raiffa, H. (1992) in *Toward a History of Game Theory*, ed. Weintraub, E. R. (Duke Univ. Press, Durham, NC), pp. 165–175.
20. Hardin, G. (1968) *Science* **162,** 1243–1248.
21. Rapaport, A. & Chammah, A. M. (1965) *Prisoner's Dilemma: A Study in Conflict and Cooperation* (Univ. of Michigan Press, Ann Arbor).
22. Axelrod, R. (1984) *The Evolution of Cooperation* (Basic Books, New York).
23. Ledyard, J. (1995) in *Handbook of Experimental Economics*, eds. Kagel, J. & Roth, A. (Princeton Univ. Press, Princeton), pp. 111–194.
24. Goeree, J. K. & Holt, C. A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 10564–10567.
25. Ostrom, E. (1998) *Am. Polit. Sci. Rev.* **92,** 1–22.
26. Avery, C., Jolls, C., Posner, R. A. & Roth, A. E. (2001) *Univ. Chicago Law Rev.* **68,** 793–902.
27. Roth, A. E. (1984) *J. Polit. Econ.* **92,** 991–1016.
28. Roth, A. E. (1990) *Science* **250,** 1524–1528.
29. Gale, D. & Lloyd, S. (1962) *Am. Math. Month.* **69,** 9–15.
30. Roth, A. E. & Sotomayor, M. (1990) *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, Econometric Society Monograph Series (Cambridge Univ. Press, Cambridge, U.K.).
31. Roth, A. E. & Peranson, E. (1999) *Am. Econ. Rev.* **89,** 748–780.
32. Milgrom, P. R. (2004) *Putting Auction Theory to Work* (Cambridge Univ. Press, Cambridge, U.K.).
33. Wilson, R. B. (2002) *Econometrica* **70,** 1299–1340.
34. Vickrey, W. (1961) *J. Finance* **16,** 8–37.
35. Kalisch, G. K., Millnor, J. W., Nash, J. F. & Nering, E. D. (1954) in *Decision Processes*, eds. Thrall, R. M., Coombs, C. H. & Davis, R. L. (Wiley, New York), pp. 513–518.
36. Kagel, J. H. & Roth, A. E., eds. (1995) *Handbook of Experimental Economics* (Princeton Univ. Press, Princeton).
37. Davis, D. D. & Holt, C. A. (1993) *Experimental Economics* (Princeton Univ. Press, Princeton).
38. Kahneman, D. & Tversky, A. (1979) *Econometrica* **47,** 263–291.
39. Smith, V. L. (1962) *J. Polit. Econ.* **70,** 111–137.
40. Fudenburg, D. & Kreps, D. (1993) *Games Econ. Behav.* **5,** 320–367.
41. Fudenberg, D. & Levine, M. (1998) *Learning in Games* (MIT Press, Boston).
42. Erev, I. & Roth, A. E. (1998) *Am. Econ. Rev.* **88,** 848–881.
43. Capra, C. M., Goeree, J. K., Gomez, R. & Holt, C. A. (2002) *Int. Econ. Rev.* **43,** 613–636.
44. Goeree, J. K. & Holt, C. A. (2002) in *Encyclopedia of Cognitive Science*, ed. Nadel, L. (McMillan, London), Vol. 2, pp. 1060–1069.
45. Goeree, J. K. & Holt, C. A. (2003) *Games Econ. Behav.*, in press.
46. McKelvey, R. M. & Palfrey, T. R. (1995) *Games Econ. Behav.* **10,** 6–38.
47. Goeree, J. K., Holt, C. A. & Palfrey, T. R. (2002) *J. Econ. Theor.* **104,** 247–272.
48. Bolton, G. E. & Ockenfels, A. (2000) *Am. Econ. Rev.* **90,** 166–193.
49. Fehr, E. & Gächter, S. (2002) *Nature* **415,** 137–140.
50. Nowak, M. A. & Sigmund, K. (1998) *Nature* **393,** 573–577.