Scand. J. of Economics 108(2), 213–229, 2006
DOI: 10.1111/j.1467-9442.2006.00447.x

Thomas Schelling's Contributions to Game Theory

*Avinash Dixit**

Princeton University, Princeton, NJ 08544, USA
dixitak@princeton.edu

I. Introduction

Thomas Schelling brings game theory to life, and does so in two senses.

First, instead of the abstract and austere beauty of mathematical theorems that we find in most of game theory going back to von Neumann, Morgenstern and Nash, he gives us lively and memorable examples to explain and illustrate his ideas. Schelling's (1960a, p. 21) pictures of bargaining—"When two dynamite trucks meet on a road wide enough for one, who backs up?"—and of tacit coordination—can two people who are told to meet each other in New York city on a specified day, but are not told the place or time and are unable to contact each other, successfully meet? (1960a, pp. 55–56)—are among the best known. But my personal favorite is Schelling's (1965, pp. 66–67) example of the strategy of "salami tactics" for defeating threats by incremental escalation:

> "Salami tactics, we can be sure, were invented by a child ... Tell a child not to go in the water and he'll sit on the bank and submerge his bare feet; he is not yet 'in' the water. Acquiesce, and he'll stand up; no more of him is in the water than before. Think it over, and he'll start wading, not going any deeper; take a moment to decide whether this is different and he'll go a little deeper, arguing that since he goes back and forth it all averages out. Pretty soon we are calling to him not to swim out of sight, wondering whatever happened to all our discipline."

Second, his ideas have direct relevance to many areas of political, social, economic and even personal lives. Whether you want to improve the security of your country in the nuclear age, or want to make sure you don't hit the snooze button on your alarm clock too often in the morning, you should know and apply Schelling's analysis and conclusions. If you want to learn how to prove theorems in game theory, you have your choice of many

* I am grateful to Vincent Crawford, Tore Ellingsen, Barry Nalebuff, Sylvia Nasar, Michael Spence and Richard Zeckhauser for comments on a previous draft.

excellent textbooks and technical journals; if you want to learn how to live game-theoretically, you have to read Schelling.

In statements and interviews on the day he was awarded the prize, Schelling decribed himself as a user of game theory rather than a game theorist.[1] I think he is being excessively modest. He elucidated many of the most basic concepts and strategies of game theory—credibility, commitment, threats and promises, brinkmanship, focal points and the tipping of equilibrium, to mention just a few. These ideas became the foundations for numerous applications in economics, political science, military strategy, and several other fields. Schelling himself launched many of these applications, thereby giving inspiration and guidance to others. He fully merits his place among the intellectual founders of game theory, and ranks as the theory's most inspired and effective visionary.

His theoretical vision is closely linked to his observations of life. The nuclear arms race provided much of the motivation and context for the conceptual framework of *The Strategy of Conflict*. Schelling was working actively in the field of defense and international relations at the time, at the Rand Corporation, as a consultant to government agencies, and so on. His concern for the problems of racial segregation and policies to counter it provided a starting point for much of the analysis in *Micromotives and Macrobehavior*. And his own (ultimately successful) efforts to quit smoking were an important input to his thinking about games played by divided selves and for *Choice and Consequence*.

Herein lies another lesson for the budding economic theorist. If you want to make incremental contributions to the literature in your field—let us say, if you want to generalize an existence theorem by relaxing the condition of semi-strict quasi-concavity to one of mere hemi-demi-proper pseudo-concavity—then stick to the technical journals. If you want to change your field in more fundamental ways, then obtain your primary motivation from life, and use it to look for fundamental shortcomings of previous thinking in the field.

This article is my attempt to give the readers a tiny glimpse of the riches of ideas and examples in Schelling's writings. Those who have not yet read the originals should surely do so; I hope they will find my brief account a tempting appetizer.[2]

---

[1] Schelling told reporters that "[Aumann] is a producer of game theory and I am a user of game theory" (BBC News, http://news.bbc.co.uk/2/hi/business/4326732.stm).

[2] Other appreciations of Schelling include Zeckhauser (1989), Crawford (1991) and Dixit and Zeckhauser (1996). Also noteworthy is a recent interview with him in the Richmond Fed's *Regional Focus*, Spring 2005. Detailed biographical and bibliographical information can be found on the Nobel Foundation's website, with links from http://nobelprize.org/economics/laureates/2005/index.html.

II. The First Decade of Game Theory

The publication of von Neumann and Morgenstern's (1944) pioneering book on game theory was immediately hailed as a milestone in the social sciences. Its promise took a long time to fulfillment, but is now essentially complete. Game-theoretic concepts, terminology and modes of analysis have come to dominate most areas of economics, and have made large inroads into political science and other social sciences. Thomas Schelling deserves a considerable part of the credit for this. His penetrating insights, expressed in simple prose or equally simple models, have given us powerful concepts like credibility, focal points and tipping of equilibria. But the very success of Schelling creates the risk of under-appreciation. So thoroughly have his ideas permeated our thinking that they may have become obvious. Therefore I shall begin by sketching a picture of game theory as it existed before Schelling.

Reviewers of the von Neumann–Morgenstern (1944) book expected great things. Leonid Hurwicz (1945) said:

> "The potentialities of von Neumann and Morgenstern's new approach seem tremendous and may, one hopes, lead to revamping, and enriching in realism, a good deal of economic theory ... It would be doing the authors an injustice to say that theirs is a contribution to economics only. The scope of the book is much broader. The techniques ... are of sufficient generality to be valid in political science, sociology, and even military strategy."

Jacob Marschak (1946) praised their "astounding power of generalization", and prophesied that "[t]en more such books and the progress of economics is assured".

In the event the potentialities had to wait. Many mathematicians did not place game theory high in the ranking of pure mathematical fields. Most economists of the time had inadequate mathematical training to appreciate or use game theory, and just as this began to be remedied in the 1950s, along came general equilibrium theory, which fit more naturally into the economists' supply–demand perspective. But most of the fault for the delayed acceptance of game theory must be attributed to the state of that theory itself. Von Neumann and Morgenstern focused on two topics: two-person zero-sum games, and cooperative games. The former was too restricted to have much economic application; the latter rested on an assumption that some external agency would enforce the agreements arrived at by the players in the game. Nash proved a general theorem on existence of equilibrium in many-player non-zero-sum games played non-cooperatively, i.e., without external enforcement of the players' stipulated actions. But this was at too high a level of abstraction for application. That had to await the enrichment of the strategic and informational specification of games;

Schelling played a crucial and pioneering role in the development of these concepts and applications.

The reality after more than a decade of development is nicely captured in the classic text of Luce and Raiffa (1957). First, it has an excellent exposition of the two most important contributions of von Neumann and Morgenstern: expected utility theory (two chapters and an appendix) and the minimax theorem for two-person zero-sum games (one chapter and five appendices). Next, it has a lot of material (six chapters) on cooperative games. By contrast, it has only one chapter on many-player, non-zero-sum, non-cooperative games. After a brief proof of the existence of the Nash equilibrium, it develops just two special examples: the prisoners' dilemma and the battle of the sexes. The relative attention devoted to cooperative and non-cooperative games is almost exactly the opposite of the subsequent use these approaches have found.

Don't get me wrong. Luce and Raiffa's book is excellent; I still find frequent occasions to consult it. The fault lies with the state of development of the subject a dozen years after von Neumann and Morgenstern.

III. Changing the Game

In this setting, Schelling's 1956 essay on bargaining, reprinted as Chapter 2 of *The Strategy of Conflict* (1960a), strikes a dramatic new note and constructs a way forward. It begins with a basic and seemingly counterintuitive observation about bargaining: it is not a zero-sum game. True, the players' interests are in strict conflict when we compare any two Pareto-efficient outcomes. But failure of the negotiation is a possible outcome, and when comparing failure to any agreement in the core, the players' interests are in perfect agreement. Therefore bargaining is a variable-sum game.

Schelling next observes that one bargainer can secure an outcome better for himself if he can credibly persuade the other that the only alternative is the mutually undesired failure. "[W]hat one player *can* do to avert mutual damage affects what another player *will* do to avert it." This leads to several important logical implications.

First, it makes another seemingly counterintuitive idea obvious upon reflection. Restricting one's own freedom of choice may be strategically desirable in this context. If one player is unable to make any concessions, then the onus is on the other to concede in order to avoid disaster. This idea of strategic commitment has become one of the most fruitful in economics and politics.

Schelling generalizes the concept of commitment into a broader class of *strategic moves*; these are actions taken prior to playing a subsequent game with the aim of changing the available strategies, information structure or payoff functions of that game. This is where, in my judgment, he alters

traditional thinking of game theorists in the most fundamental way. Most game theorists insist on starting with a complete specification of the game. Once a game is fully specified, the theorist can determine the outcomes using a specified equilibrium concept. Schelling comes to the question from almost the opposite angle. His players ask themselves: "This is the outcome I would like from this game; is there anything I can do to bring it about?" This perspective is closer to the concept of mechanism design in information economics, but is richer and more complex in that all players in the game can simultaneously attempt to devise methods to alter its outcome in their own favor.

Schelling (1960a, p. 117, emphasis added) expresses his concept of strategic moves as actions that "can in some way *alter the game*, by incurring manifest costs, risks, or a reduced range of subsequent choice". This marvellously echoes Karl Marx in his *Theses on Feuerbach*: "Philosophers have only interpreted the world. The point, however, is to change it."

Modern game theory treats strategic moves somewhat differently, namely as the first stage of a multi-stage game. The actions available to the players at the first stage are fully specified, and the solution concept is subgame-perfect Nash (further modified to include Bayesian updating of beliefs where that is relevant).[3] One player may have an advantage at this stage; this may be an accident of history, as in the case of an incumbent firm as opposed to one merely contemplating entry, or some technological or organizational asymmetry, as in the case of a firm that can expand its capacity faster. Or the players may have identical opportunities available for making strategic moves, but one may merely recognize and exploit them first; this is a non-equilibrium account of play in the given multi-stage game.

In reality there is a whole continuum, from perceiving and playing a given multi-stage game to a more fundamental alteration of the shared mental constructions of the players about the game they are playing and its equilibrium selection.[4] Observed behavior that comes closest to the former includes such experiments as Gneezy, Rustichini and Vostroknutov (2005), where the subjects play a nim-like game with fully explained rules, and only gradually figure out that they should be doing backward induction. Closer to the latter extreme are major constitutional changes in the rules of governance that require collective action to devise and implement.

[3] Reinhard Selten and John Harsanyi shared the Nobel Memorial Prize with John Nash in 1994 for their work on these concepts of subgame-perfect and Bayesian equilibria.

[4] Crawford (1985) develops a general theoretical framework for multi-stage non-cooperative games where individual players' actions in the earlier stages can endogenously change the rules of the game at the later stages. He argues that an infinite regress of changing the rules for changing the rules ... can only be avoided by specifying exogenous meta-rules, and that understanding the process of specifying these meta-rules is the right way to do cooperative game theory.

Schelling's different examples can be located at different points along this spectrum, but most of them require one or more of the players to exercise considerable creativity and ingenuity in devising and making the appropriate strategic moves. It would have been difficult for the other players to identify *ex ante* the technology or strategy as already existing in a larger multi-stage game. Therefore, the "changing the game" viewpoint is often more suitable in these applications.

Alternatively, we can think of the theory that explicates the solution of fully specified multi-stage games by backward induction as belonging to the science of game theory, and the imagination and creativity involved in perceiving or altering, and then playing, the early-stage actions as the art of game-playing. Both the science and the art are important aspects of reality.

Starting with the idea of commitment, or unconditionally taking an inflexible position, Schelling goes on to more complex and subtle conditional commitments, where a player commits to a response rule, which specifies the actual action the player will take in response to the other player's actions. Threats and promises are the basic conditional rules of this kind. They require a player to change the subsequent game so that he has the second move in it, and to commit in advance to how he will make that move in all conceivable eventualities.

All strategic moves—commitments, threats and promises—must be credible. The modern formal notion of credibility is subgame perfectness; the outcome that the player making the strategic move wishes to achieve must be the result of a subgame-perfect equilibrium in the enlarged game. Schelling had this essential idea, but his treatment of it was different; see Schelling (1960a, pp. 150–158). Even he could not escape the mathematical straitjacket of his time. Subgame perfectness is best understood in the extensive or tree form of a game, but in the 1950s game theory generally used only the strategic or normal form. All the strategies (complete contingent plans) of the respective players were laid out as rows and columns in a payoff matrix. This was natural for simultaneous-move games, but awkward for sequential-move games that were essential for analyzing strategic moves. Schelling developed the notion of credibility as the outcome that survives iterated elimination of weakly dominated strategies. We now know that in his context of generic extensive-form games with complete and perfect information, this procedure does indeed work; see Kreps (1990, p. 424). At least in this context, Schelling should get credit as a pioneer of the subgame-perfectness concept.

Practical implementation of the theoretical concept of credibility requires imaginative and application-specific thinking, and remains as much art as science. Schelling gives us wonderful examples of the art. Some are drawn from reality, while others are brilliant exaggerations that are even better at fixing the ideas in our minds because they are so memorable.

Schelling points out that to make a fixed position in bargaining credible, one can use an agent or delegate who has no authority to deviate from his given mandate (a device that is used by labor unions and others), but he also invokes the vision of a fictitious society in which a declaration of "cross-my-heart" is potent; see Schelling (1960a, pp. 29, 24). His example of the game of chicken, where each of two drivers heading toward each other wants to convince the other that he will not swerve, comes not from Hollywood movies, but from Homer's vivid account of a chariot race between Antilochos and Menelaos; see Schelling (1965, p. 117). And he tells us how apparent irrationality, recklessness or unreliability can be a good way to achieve credibility and can therefore be strategically rational: "hot-rods, taxis, and cars with Driving School license plates" are given wide berths; see Schelling (1965, p. 116).

Brinkmanship is probably the subtlest of all strategic moves, and Schelling's analysis of it is equally subtle; see Schelling (1960a, Chs. 7, 8; 1965, Ch. 3). The opposing players are on the verge of a precipice of risk such as a war or a strike, and brinkmanship is sometimes thought of as a threat made by one player to jump off carrying the other with him to disaster if the other does not comply with the first's demand. Schelling argues that brinkmanship cannot be understood as a cool and deliberate threat; that is too costly to the threatener himself to be credible. Rather, brinkmanship is "manipulation of shared risk"; see Schelling (1965, pp. 97, 99).

> "If the brink is clearly marked and provides a firm footing, no loose pebbles underfoot and no gusts of wind to catch one off guard, if each climber is in full control of himself and never gets dizzy, neither can pose any risk to the other by approaching the brink ... [W]hile either can deliberately jump off, he cannot credibly pretend that he is about to. Any attempt to intimidate or to deter the other climber depends on the threat of slipping or stumbling ... [O]ne can credibly threaten to fall off accidentally by standing near the brink ... Deterrence has to be understood in relation to this uncertainty ... That is why deterrent threats are often so credible. They do not need to depend on a willingness to commit anything like suicide in the face of a challenge. A response that carries some risk of war ... through a compounding of actions and reactions, of calculations and miscalculations, of alarms and false alarms ... can be plausible, even reasonable, at a time when a final, ultimate decision to *have* a general war would be implausible or unreasonable."

Brinkmanship is a richer and more realistic dynamic version of chicken. Unlike the simple version where each driver chooses *whether* to swerve, the question is *when* to swerve. As the cars get closer to each other, the risk that they may no longer be able to swerve in time gradually rises. Each hopes that the other will swerve in time, but a positive probability of disaster remains.

Schelling developed a formal model of brinkmanship using the tools available at that time (1960a, Ch. 7). Nowadays we model it as the game called "war of attrition with asymmetric information" that has found so many applications in auction theory etc.; see Fudenberg and Tirole (1991, pp. 216–219, 239–240). But once again Schelling's pioneering analysis and examples convey the essential idea with ample logical clarity and simplicity.

Schelling's analysis of bargaining is quite general, but many of his illustrations and applications concern military strategy. Schelling points out that military strategy is very often "not ... the efficient *application* of force but the *exploitation of potential force*" (1960a, p. 5, italics in the original). Credibility of threats and promises was central to questions of nuclear deterrence, the arms race and attempts to control it. *The Strategy of Conflict* launched the whole field of strategic studies, and was a major influence on the emerging field of international relations.

One of Schelling's simple but beautiful insights in this area concerns the question of surprise attack—the possibility that one superpower may launch a pre-emptive attack fearing that the other might attack because of its fear that. ... Schelling (1960a, Chs. 9, 10), argues that if each side has a credible second-strike capability, that suffices to deter the other's attack, and a pre-emptive first strike is not needed. Therefore, the vital preventive action was to ensure that each superpower's missiles would survive the other's attack: "[S]chemes to avert surprise attack have as their most immediate objective the safety of weapons rather than the safety of people" (Schelling, 1960a, p. 233); populations are better protected by protecting missiles. That is why the MIRV missiles developed and deployed in the late 1970s and early 1980s were especially dangerous. They could hit the enemy's missile silos in great numbers and with great accuracy; they threatened his second-strike capability. This increased the need for a hair-trigger approach and a risk of pre-emptive strike by error or miscalculation. Fortunately the world survived, and we are able to appreciate and celebrate Schelling's early insight.

In an article (Schelling, 1960b) and book (Halperin and Schelling, 1961), Schelling examines many different strategies for stabilizing and controlling the nuclear arms race. These include design and deployment of forces, methods of explicit and tacit communication, and the building of mutual understanding and confidence; many were later adopted and have proved their worth.

IV. Meetings and Meetings of Minds

Schelling's emphasis on the mixture of conflict and common interest in game theory finds another striking application in the organization of tacit coordination when interests are mostly aligned. Schelling (1960a, p. 86)

emphasizes that pure coordination is a game of strategy just as much as one of conflict:

> "It is a behavior situation in which each player's best choice depends on the action he expects the other to take, which he knows depends, in turn, on the other's expectation of his own. This interdependence of expectations is what distinguishes a game of strategy from a game of chance or a game of skill."

The best-known example of this is the question of how to meet someone successfully when no place or time for the meeting has been arranged (or these matters had been arranged but have been forgotten); indeed, this may be the second most popularly known example of game theory, second only to the prisoners' dilemma. A more material economic application concerns an oligopoly whose member firms want to arrive at a tacit understanding to divide the market among them.

Schelling argues that a successful solution to these tacit coordination problems hinges on whether the players' expectations can converge on an outcome—a focal point. If I go and stand under the clock in Grand Central Station at noon, it must be because I think the person I am supposed to meet will do likewise. But he will do likewise only if he thinks that I think that he will be there at that time. And that will be the case only if I think that ... This sequence of thinking about thinking has to converge on a common expectation to ensure a successful meeting. In modern terminology of philosophy and game theory, we would say that the time and place where we are supposed to meet should be *common knowledge* between us.

Once again Schelling stands as a pioneer for a concept that became formalized in full generality later. Lewis (1969) is generally credited with the formal definition of common knowledge; Schelling's fellow laureate Aumann (1976) introduced it to economics. Geanakoplos (1994) has surveyed the subsequent literature.

Schelling points out (1960a, pp. 95–99) that the convergence of expectations essential for creating focal points usually arises from considerations beyond the pure formalism of a game. Many mathematical theorists prefer the solution concepts to be free from any particular labeling of the players or strategies. For them it should make no difference whether the players are the U.S. and the Soviet Union and the strategies are war or peace, or the players are A and B and the strategies are 1 and 2. For them it should make no difference whether a location is labeled "under the clock at Grand Central Station" or number 9 in a list of 20 numbers. But in reality it makes all the difference. Some historical, cultural, linguistic or even mathematical *salience* is essential for the creation of convergent expectations in the minds of players. This also explains why people coming from different social or cultural backgrounds are often less successful in solving tacit coordination

problems. Mathematicians may frown at this, but I think Schelling is right. The social sciences constitute the realm of the main application of game theory, and it is only fitting that the solutions involve social considerations in an essential way.

Even in situations like bargaining, where conflict of interests is more important, focal points can help the parties avoid the mutually bad outcome of no agreement. That may be why 50:50 division is observed so often in situations ranging from the ultimatum game to sharecropping, and similar conventions apparently override explicit rational calculation to determine the outcomes of many social interactions.

> "The coordination game probably lies behind the stability of institutions and traditions ... The force of many rules of etiquette and social restraint, ... that have been divested of their relevance or authority, seems to depend on their having become 'solutions' to a coordination game." [Schelling, 1960a, p. 91]

Here is the germ of an idea—social conventions, norms and institutions as creators of focal points for equilibrium selection—that many others have developed, often unaware that Schelling had said it before.

And finally in this context, Schelling considers the effect of allowing explicit communication: "The pure-coordination game not only ceases to be interesting but virtually ceases to be a 'game' if the players can concert [using] overt bargaining with uninhibited speech" (Schelling, 1960a, pp. 100–101). But "[t]he contrast with a zero-sum game and the particularly self-effacing quality of a minimax solution is striking. [In such a game,] [o]ne not only does not need to communicate with his opponent, he does not even need to know who the opponent is"; any communication would be rationally ignored (Schelling, 1960a, p. 105). Here is a clear understanding, at least for the two extreme cases, of the role of "cheap talk" in games. The seminal paper on the general theory of strategic communication between players whose interests have something in common but are not identical is Crawford and Sobel (1982); it was reading Schelling's discussion of the two extremes and puzzling about what might lie between them that led Crawford and Sobel to their celebrated work.[5]

V. When Bad Equilibrium Happens to Good People

The focus of Schelling's *Micromotives and Macrobehavior* (1978) is on social and economic interactions among large numbers of participants. One might have thought that game theory would not matter in such a context,

[5] Personal communication from Vincent Crawford.

because the strategic significance of each player would be negligible in a large population. But that is true only in a very special class of games, namely market competition for purely private goods. In other settings, there are subtle local or global externalities, and individually rational choices can lead to suboptimal game equilibria, sometimes in surprising ways. Schelling elucidates and illustrates these with his customary combination of precision and telling anecdotes.

His motivation for much of this work came from his deep concern for problems of race and segregation. Must this be the result of deeply held racism by most individuals, or might milder preferences for being among neighbors like oneself (micromotives) translate into starkly segregated residential patterns (macrobehavior)? To understand the interaction of individual preferences into social outcomes in this setting, he conducted extensive experiments by moving pennies around on a checkerboard. Each piece can be surrounded by up to eight other pieces; these are its "neighbors". A decision rule determines whether a piece stays or moves to a new location, depending on its own color and the number of neighbors of the same color. Tracing the dynamics of this process, Schelling found that even very mild preference for having a few neighbors of the same color can lead to fully segregated outcomes. The mechanism is a positive feedback effect: when one person of a given color moves away, others near his previous location have fewer similar neighbors, which shifts the balance in their decision toward moving also. Then a small change in preferences or a small exogenous disturbance to the composition can set in motion a process of "tipping" into segregation. And once a segregated outcome is reached, it can be very hard to halt and harder to reverse. Therefore persuading the few initial movers to remain is essential for the successful stabilization of an integrated neighborhood.

These insights have now passed into the folklore of urban economics and sociology. Their scope has expanded to the numerous other applications Schelling developed in *Micromotives and Macrobehavior*. The idea of "tipping" has even passed into the world of bestsellers and popular culture; see Gladwell (2000; see Endnote on p. 282). And Schelling's early experiments on the chessboard have evolved into much more elaborate computer programs, and many versions of his simulation game are available on the internet.[6]

The final chapter of this book (Schelling, 1978, Ch. 7) is a brilliant analysis and application of multi-person "dilemma" games. The theory is simple and the exposition is purely geometric, but the insights are quite profound and the applications are telling.

[6] Two such programs are at http://ccl.northwestern.edu/netlogo/models/Segregation and http://www.econ.iastate.edu/tesfatsi/demos/schelling/schellhp.htm.

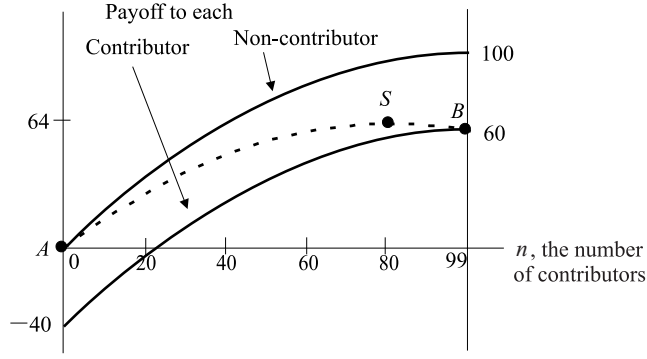

Fig. 1.

The key simplification is to assume that each player has a purely binary choice, and that the interaction is impersonal in the sense that each player's payoff depends only on the numbers of others making the one choice or the other, and not on the identities of these choosers. Then the whole payoff structure can be displayed in a simple diagram. I will explain the method using two examples; Schelling develops many other applications.

My first example is a multi-person prisoners' dilemma. Consider a population of 100 people facing a collective action problem of providing a non-excludable good. If n people contribute to its provision, the gross benefit to each of the 100 people is given by

$$2n - 0.01n^2. \tag{1}$$

This function starts at 0 and levels off at 100 when $n = 100$; thus there are diminishing returns to the good and the last person's contribution has very low gross marginal benefit. Each contributor incurs a cost of 40, whether in money or time or some other currency.

View the choice from the perspective of any one selected player. Figure 1 shows his payoffs. On the horizontal axis is the number n of other contributors, where n ranges from 0 to 99. The lower solid curve is the selected player's payoff when he does not contribute so there are n total contributors; this is just the function (1) above. The upper solid curve is the selected player's payoff when he does contribute, so there are $(n + 1)$ total contributors and he incurs the cost 40; that is:

$$2(n + 1) - 0.01(n + 1)^2 - 40. \tag{2}$$

In the figure, the non-contribution payoff curve is everywhere above the contribution payoff curve. Therefore for every n in the range from 0 to 99, the selected player gets a higher payoff if he does not contribute

than if he does; non-contribution is his dominant strategy.[7] But when all players choose their dominant strategy, $n = 0$ and the selected player is at point A at the left endpoint of the upper of the two curves, with payoff 0 to each. If instead everyone had contributed, the outcome would be at B, the right extreme of the lower curve, and the payoff to each would be 60.

This is just a simple way of depicting the multi-person dilemma. But the figure leads us to a further consideration that is more subtle. Fully resolving the dilemma and inducing everyone to contribute to reach point B is not socially optimal in this instance! If there are n contributors and $(100 - n)$ non-contributors, the aggregate payoff is the gross benefit of all 100 minus the cost incurred by the n contributors; that is:

$$100(2n - 0.01n^2) - 40n = 160n - n^2. \tag{3}$$

The average payoff is $1.6n - 0.01n^2$; in the figure this is shown by the dotted line.

Because the gross marginal benefit gets very small as n approaches 100, the net benefit is maximized for a smaller value of n. In the figure, this is point S. It is easy to calculate that the optimal choice is $n = 80$. The resulting average net benefit is 64. But this comes about because the 80 contributors get 56 each and the 20 non-contributors get 96 each. If the group wants to achieve this highest aggregate or average benefit, it must solve the additional and potentially more difficult problems of deciding who contributes, and whether and how the non-contributors compensate the contributors for their cost.

My second example concerns network effects, and the possibility of getting locked in to an inferior equilibrium. Examples like the choice between QWERTY and DVORAK keyboards, and Windows and Linux operating systems, have been offered (David, 1985) and disputed (Liebowitz and Margolis, 1990). As usual Schelling has a more vivid example, namely the adoption of helmets by professional ice hockey players.

Figure 2 shows the payoffs for an individual player from each of his two choices, as a function of the percentage of the population of players who have chosen to wear helmets. When very few others are wearing helmets, a wearer suffers the shame of being thought to be a coward, whereas a non-wearer gets sufficient ego-gratification for his macho to offset the risk of injury. When almost everyone else is wearing helmets, refusing to wear one seems foolish bravado, and the safety value of wearing a helmet looms large. The value of safety or, conversely, the cost of injury, is likely to

[7] The difference between the expressions in (1) and (2) is easily calculated to be $0.01(2n + 1) + 38$, which is always positive.

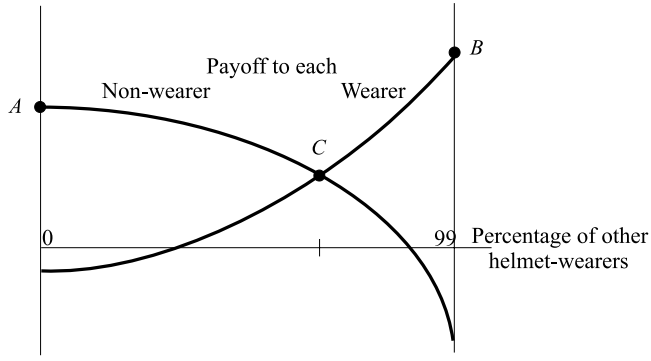

Fig. 2.

be large when almost everyone is wearing helmets, because then they may play more aggressively. Therefore the two curves cross. If no one else is wearing helmets, my preferred strategy is not to wear one myself; this is a Nash equilibrium, shown at point *A*. If everyone else is wearing one, it is optimal for me to do likewise; this is another Nash equilibrium, shown at point *B*. The payoffs in the figure are such that the *B* equilibrium is better for everyone than the *A* equilibrium. But unless enough players can be persuaded to wear helmets to take the population past the crossing point *C*, a further cumulative process of switches leading to *B* may not get established, and the players may stay locked in at *A*.

Examples like this emphasize that individually rational choice is no guarantee of a collectively desirable equilibrium. Any contrary idea that people may have garnered from a mistaken reading of Adam Smith needs to be discarded; social optimality of perfectly competitive markets for private goods is a very special property that does not generalize to other kinds of interactions. Bad outcomes can arise either because individuals make irrational or mistaken choices, or because of the nature of the interaction of individual choices, no matter how carefully and rationally calculated. Both possibilities exist, and the recent fashion among economists is to emphasize the former by invoking psychological or behavioral aspects of decision-making. But I think that the latter possibility is both more frequent, and more serious. Errors of perception and calculation in individual decisions can often be cured by giving the individuals better information and presenting them with better default options, as has been found to be the case in saving and healthcare choices; e.g. Bernheim and Garrett (2003); Choi, Laibson, Madrian and Metrick (2003). But inefficiencies that arise at the level of the interaction or equilibrium of individual choices require collective action to change the equilibrium, which is considerably more difficult.

VI. The Enemy Is Us

It need not take two distinct individuals to play a game of strategy; one person may be playing against his own future self. In *Choice and Consequence* (1984), Schelling examines and elucidates many such situations facing people and societies.

We have all taken actions while our rational self was in charge, to guard against the temptations to which a future irrational or addicted or merely weak-willed self would succumb. This is a kind of commitment, but its purpose is to alter one's own future action and not that of someone else. We arrange for a portion of our paycheck to be withheld and put into a pension plan, because if we had it in hand we might be tempted to buy that plasma TV or multi-megapixel digital camera. We place the alarm clock on a desk far from the bed, ruling out the possibility of hitting the snooze button in the morning. Conversely, in situations where our current self thinks that the preferences of the future self may be good and worth incorporating into the current calculation of payoffs, we take actions that facilitate the formation of such preferences; we cultivate tastes and form habits that will improve our selves. Schelling illustrates and classifies many such actions (1984, Ch. 3; 1996).

His motivation for this research came from the problem of quitting smoking. He had been a smoker himself, and he got deeply immersed in the field, knowing all the studies, etc. He ran a substantial smoking project at Harvard for many years, and somewhere in the middle of those years his book *Choice and Consequence* emerged.

Similar problems writ large confront societies, not just individuals, and at that level they sometimes go unsolved because the necessary collective action cannot be coordinated. Large amounts of resources are spent to rescue people who are stranded, or to treat particular sufferers from serious diseases, when much smaller sums spent on risk-reduction or disease-prevention would have helped many more. Schelling observes that "[w]e often know who died for lack of safety; we rarely know who lived because of it" (1984, Ch. 5). This essay offers a clear-headed but considerate analysis of the willingness-to-pay approach to the valuation of life and similar touchy issues.

Recent work on hyperbolic discounting and saving is another example of Schelling's concepts about self-control; see Harris and Laibson (2002). A person with such preferences, when making decisions in 2006, discounts outcomes in 2007 and beyond heavily relative to 2006, so he is not willing to save much in 2006, but discounts 2008 and beyond much less relative to 2007, so he plans to save a lot in 2007. Of course his preferences shift when 2007 arrives; now he discounts 2008 heavily relative to 2007, and wants to postpone the high-saving plan for another year. If in 2006 he looks

ahead and foresees this, he will gladly make an irreversible commitment to a saving plan that starts in 2007. Researchers have found that good choices of default options (opting in versus opting out for saving plans at work) can serve as useful commitment devices to counter the temptation to procrastinate; see Madrian and Shea (2001). As usual, Schelling anticipates some of this discussion with several striking examples; see Schelling (1984 Ch. 3; 1996).

VII. Concluding Thoughts

Schelling's work shows that it is possible to be clear, precise and logically rigorous without being overtly mathematical. Schelling's writing shows that it is possible to explain deep and important ideas in a simple and transparent way. His examples stay in our mind long after the algebra and geometry are forgotten.

Schelling stands as an exemplar of a research style that relies on deep integration between observations from life and theoretical thinking, where each motivates and enriches the other. Newcomers to research in economics and game theory could find no better role model.

Schelling has done more than create many of the most important concepts of modern game theory. He has given us more than mere game theory to help us become better social scientists. He has given us game theory to practice, and game theory to remember.

References

Aumann, R. J. (1976), Agreeing to Disagree, *Annals of Statistics 4*, 1236–1239.

Bernheim, B. D. and Garrett, D. M. (2003), The Effects of Financial Education in the Workplace: Evidence from a Survey of Households, *Journal of Public Economics 87*, 1487–1519.

Choi, J. J., Laibson, D., Madrian, B. C. and Metrick, A. (2003), Optimal Defaults, *American Economic Review 93*, Papers and Proceedings, 180–185.

Crawford, V. (1985), Efficient and Durable Decision Rules: A Reformulation, *Econometrica 53*, 817–836.

Crawford, V. (1991), Thomas Schelling and the Analysis of Strategic Behavior, in R. J. Zeckhauser (ed.), *Strategy and Choice*, MIT Press, Cambridge, MA.

Crawford, V. and Sobel, J. (1982), Strategic Information Transmission, *Econometrica 50*, 1431–1452.

David, P. (1985), Clio and the Economics of QWERTY, *American Economic Review 75*, Papers and Proceedings, 332–337.

Dixit, A. and Zeckhauser, R. (1996), Thomas Schelling, in W. J. Samuels (ed.), *American Economists of the Late Twentieth Century*, Edward Elgar, Cheltenham.

Fudenberg, D. and Tirole, J. (1991), *Game Theory*, MIT Press, Cambridge, MA.

Geanakoplos, J. (1994), Common Knowledge, in R. J. Aumann and S. Hart (eds.), *Handbook of Game Theory*, Vol. II, Elsevier, Amsterdam.

Gladwell, M. (2000), *The Tipping Point: How Little Things Can Make a Big Difference*, Boston, MA: Little, Brown.
Gneezy, U., Rustichini, A. and Vostroknutov, A. (2005), Backward Induction as a Cognitive Process, working paper, University of Chicago Graduate School of Business.
Halperin, M. H. and Schelling, T. C. (1961), *Strategy and Arms Control*, Twentieth Century Fund, New York.
Harris, C. J. and Laibson, D. (2002), Hyperbolic Discounting and Consumption, in M. Dewatripont, L.-P. Hansen and S. Turnovsky (eds.), *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, Vol. 1, Cambridge University Press, Cambridge.
Hurwicz, L. (1945), Review of *The Theory of Games and Economic Behavior*, *American Economic Review 35*, 909–925.
Kreps, D. (1990), *A Course in Microeconomic Theory*, Princeton University Press, Princeton, NJ.
Lewis, D. (1969), *Convention: A Philosophical Study*, Harvard University Press, Cambridge, MA.
Liebowitz, S. J. and Margolis, S. E. (1990), The Fable of the Keys, *Journal of Law and Economics 33*, 1–25.
Luce, R. D. and Raiffa, H. (1957), *Games and Decisions*, John Wiley, New York.
Madrian, B. C. and Shea, D. F. (2001), The Power of Suggestion: Inertia in 401(k) Participation and Savings, *Quarterly Journal of Economics 116*, 1149–1187.
Marschak, J. (1946), Neumann and Morgenstern's New Approach to Static Economics, *Journal of Political Economy 54*, 97–115.
Schelling, T. C. (1960a), *The Strategy of Conflict*, Harvard University Press, Cambridge, MA.
Schelling, T. C. (1960b), Reciprocal Measures for Arms Stabilization, *Daedalus 134*, 101–117.
Schelling, T. C. (1965), *Arms and Influence*, Yale University Press, New Haven, CT.
Schelling, T. C. (1978), *Micromotives and Macrobehavior*, W. W. Norton, New York.
Schelling, T. C. (1984), *Choice and Consequence*, Harvard University Press, Cambridge, MA.
Schelling, T. C. (1996), Coping Rationally with Lapses from Rationality, *Eastern Economic Journal 22*, 251–269.
Von Neumann, J. and Morgenstern, O. (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.
Zeckhauser, R. (1989), Distinguished Fellow: Reflections on Thomas Schelling, *Journal of Economic Perspectives 3*, 153–164.