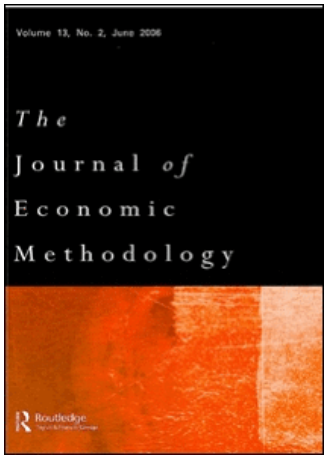

Journal of Economic Methodology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713704064

Player heterogeneity and empiricism in Schelling

Alessandro Innocenti [a]
[a] University of Siena,

Online Publication Date: 01 December 2007
To cite this Article: Innocenti, Alessandro (2007) 'Player heterogeneity and empiricism in Schelling', Journal of Economic Methodology, 14:4, 409 - 428
To link to this article: DOI: 10.1080/13501780701718649
URL: http://dx.doi.org/10.1080/13501780701718649

Journal of Economic Methodology 14:4, 409–428 December 2007

Player heterogeneity and empiricism in Schelling

Alessandro Innocenti

Abstract The main thesis of this paper is that Schelling's empiricism is deeply grounded in the assumption of player heterogeneity. He peoples games with real individuals and consequently postulates that there are differences in roles or identities among them. The subjects populating Schelling's thought or real experiments do not suppose that other players follow identical or symmetrical rules of logical inference to make their choices. This hypothesis is decisive in shaping Schelling's inductive game theory, which is applied by means of a three-step procedure. First, players are defined by making their differentiating features explicit. Secondly, heterogeneous players are embedded in a real environment to play the game. Thirdly, the game solution is derived inductively. This interpretation helps to explain why little progress has been made overall in developing Schelling's insights.

Keywords: player heterogeneity, inductive game theory, symmetry, coordination

JEL Codes: B40, C70

1 INTRODUCTION

Schelling's argumentative style may at times puzzle readers and reviewers. Indeed, his books are so full of real people and case studies that one often gets the impression that the vast amount of telling anecdotes makes the reader miss the key point. In a recent interview, Schelling replied as follows when asked to explain the many examples drawn from everyday life included in his papers:

> I always try to find something that I can put in the first paragraph to make the article sound interesting. […] So I think that has been part of my style. I wrote a textbook in international economics that had about a dozen policy chapters. I tried to have the first page of every chapter present an interesting puzzle or phenomenon that would get the interest of the readers.
>
> (Schelling 2005: 37)

Journal of Economic Methodology ISSN 1350-178X print/ISSN 1469-9427 online

DOI: 10.1080/13501780701718649

Although Schelling's aversion towards emphasizing his heterodox vein is well known,[1] the empiricism that characterizes his work is much more than a rhetorical device to capture the reader's attention. Rather, it is the basis of his unconventional method of applying the 'framework' of game theory.[2] This paper intends to discuss a specific feature of this approach. It will be argued that Schelling's originality is to a great extent the consequence of his definition of the game player. A fundamental assumption of standard game theory is external symmetry, according to which players are all alike outside the game and any differentiating characteristic has to be embedded into the game mathematical form.[3] In contrast, Schelling assumes that players are different from each other. In his games, player heterogeneity is not restricted to preference orderings or available strategies, but extends to include the players' labels and attributes, the mental representations of the game and the processes of preference formation. This method implies not only that players may perceive the same game differently but also that they are not assumed to think that other players reason identically or symmetrically to them. In terms of neoclassical economics, Schelling rejects the hypothetical construct of the representative agent. In his analyses, belief coordination is not a matter of guessing what a single abstract individual thinks, and the task of forming and coordinating player expectations is to be accomplished within the game, endogenously originated by real interaction among heterogeneous players.

This methodological trait permeates all of Schelling's contributions to economics. It provides the theoretical foundations of his bargaining theory and of the concept of focal point set out in *The Strategy of Conflict* (Schelling 1960). It is the key assumption of the racial segregation model proposed in *Micromotives and Macrobehavior* (Schelling 1978), which shows how individual rational choices do not necessarily lead to a socially desirable equilibrium. And it is also the motivation behind the focus in *Choice and Consequence* (Schelling 1984) on the issue of multiple selves, which is also a device for making the preference formation process both explicit as well as specific to the individual.

It will also be argued that this emphasis on player heterogeneity helps to explain why theoretical economics has made little progress in developing Schelling's contribution. Whereas Schelling supports an empirical and experimental approach to economics, taking into account individual heterogeneity, most economists conform to the fictitious notion of the representative agent because it makes for a fully abstract characterization of economic behaviour. Not surprisingly, this approach has been further reinforced by the application of game theory to economics as a purely deductive and mathematical tool.

The section below discusses Schelling's inductive conception of game theory, pointing out the relevance and some implications of the assumption of player heterogeneity. The third section analyses how this methodological

trait has been applied to economics in Schelling's books. Conclusions are drawn in section four.

2 SCHELLING'S INDUCTIVE GAME THEORY

Schelling's characteristic way to apply game theory has been the object of at least two seemingly conflicting interpretations. Zeckhauser's celebrative paper on Schelling's nomination as a Distinguished Fellow of the American Economic Association exemplifies the first view, which is well illustrated by the following quotation:

> Schelling, in essence, plays his games in a world that is richer than most game theory analyses. He acknowledges that players may choose 'dominated' strategies not only to create reputations, but to adhere to ethics, build self-respect, or reflect generosity. To say that Schelling is merely exploring metagames, while perhaps technically correct, seems to miss the richness.
>
> (Zeckhauser 1989: 158–9)

According to Zeckhauser, Schelling would solve games in an *ad hoc* environment that takes into account determinants of behaviour usually neglected by game theorists. This expanded set of decision criteria would aim at explaining a greater number of empirical findings, especially those conflicting with the postulate of individual maximization.[4] This interpretation is incidentally epitomized by Schelling in his criticism of Nash's solution to the bargaining problem:

> What a theory like Nash's needs is the premise that a solution exists; it is the observable phenomenon of tacit coordination that provides empirical evidence that (sometimes) rational expectations can be tacitly focussed on a unique (and perhaps efficient) outcome, and that leads one to suppose that the same may be possible in a game that provides nothing but mathematical properties to work on. The Nash theory is vindication of this supposition – complete vindication if it dominates all competing mathematical solutions in terms of mathematical esthetics. The resulting focal point is limited to the universe of mathematics, however, which should not be equated with the universe of game theory.
>
> (Schelling 1960: 289–90)

For Schelling, the mathematics of game theory is a subset of 'proper' game theory, which cannot be exhausted by the formal representation of the game. Consequently, an appraisal of Schelling's method of applying game theory would depend on the identification of this expanded set.

The second interpretation attributes Schelling with having pioneered the idea of common knowledge, a concept which has been fundamental for later developments of game theory. In this reconstruction an ideal line is traced

joining Schelling's analysis of the processes of tacit coordination to the formalization of the 'Harsanyi doctrine' given by Schelling's fellow laureate Aumann (1976). Schelling's method of solving games is seen as an empirical extension of the concept of correlated equilibrium that, by extending the validity of Nash's definition of rationality, subsumes his contribution under the umbrella of mainstream game theory and allows his empirically driven clues to be translated as abstract equilibria of rational behaviour (Janssen 2001; Sugden 2001; Dardi 2006). In this light, Schelling would have not rejected rational decision theory. Although the many anecdotes supporting his arguments point out that the hypothesis of 'homo rationalis' is not generally verified, they would not refute its validity as a framework of analysis.

Opportunely enough, Schelling makes his interpretation of game theory explicit on various occasions. A full chapter of *The Strategy of Conflict* is devoted to 'A reorientation of game theory' and a paper in *Choice and Consequence* directly addresses the question 'What is game theory?' A closer look at these and some other[5] works shows that the two interpretations outlined above miss a key point in Schelling's conception of game theory, which can be made explicit by addressing three questions. First, in what sense does game theory have normative value? Second, how does Schelling define the game player? Third, how does he describe the process of belief formation? The conclusion will be that Schelling applies game theory quite differently from the way that most economists do for reasons that are deeply rooted in his empiricism.

First, the normative value of game theory is defined as follows in *The Strategy of Conflict*:

> A normative theory must produce strategies that are at least as good as what people can do without them. More, it must not deny or expunge details of the game that can demonstrably benefit two or more players and that the players, consequently, should not expunge or ignore in their mutual interest. Two couples jockeying for space on a dance floor or two armies jockeying for a truce line may jointly suffer from decision processes that are limited to the abstract properties of the situation.
>
> (Schelling 1960: 98)

What this quotation points out is that, for Schelling, any prescriptive claim depends on the actual circumstances in which the game is played. Therefore, two formally identical games applied to different real environments may not have the same solution, and consequently, deductive reasoning cannot suffice to define the best strategy.

It is evident how this view relegates Schelling to a position outside conventional game theory. The common definition of rationality in game theory is a matter of obeying abstract rules of logical consistency. Aumann claims that

> game and economic theory are concerned with the interactive behavior of *Homo rationalis* – rational man. *Homo rationalis* is the species that always acts both purposefully and logically, has well defined goals, is motivated solely by the desire to approach these goals as closely as possible, and has the calculating ability required to do so.
>
> (Aumann 1985: 35)

Aumann's game theory is a deductive inquiry in which games are purely abstract representations and player rationality is defined by a list of theoretical conditions. Schelling diverges from this view in that players' ends, and hence the rational way of playing, are defined by inductive reasoning. He always starts thinking about games by considering puzzling examples drawn by the observation of real life. But how does Schelling choose the cases to be attached to the formal representation of the game? In the quotation above Schelling's technique consists of identifying the factual roles performed by players. His claim is that dancers and belligerent forces adopt different criteria to coordinate their choices. In general, he deems that individuals belonging to specific groups or playing different roles may take different decisions even in situations that are represented by the same game. Sometimes, Schelling makes reference to pairs of identical players – as above – while in other cases he considers pairs or groups of different players and sorts them according to some discriminating variable. This method characterizes all Schelling's work. The analytical index of *The Strategy of Conflict* lists a broad array of social and professional identities, such as bank robbers and children, parachutists and pirates, and gangs and mobs. The selection of roles and activities associated with games is even greater in *Micromotives and Macrobehavior.* The first chapter begins by discussing the seating arrangements in an auditorium and the first thirty pages alone present us with an unlimited number of player types involved in coordination problems.[6] The rest of the book includes a large variety of different roles including the often quoted example of professional ice hockey players choosing whether to wear or not wear a helmet.[7] *Choice and Consequence* and the more recent *Strategies of Commitment* also set out a vast assortment of real individuals whose identities are appropriately chosen to exemplify irrational kinds of behaviour explained by the notion of multiple selves and to analyse the corresponding techniques of self-command.

In 1960, Schelling offers a methodological explanation for this profusion of identities:

> There is, consequently, no way that an analyst can reproduce the whole decision process either introspectively or by an axiomatic method. There is no way to build a model for the interaction of two or more decision units, with the behavior and expectations of those decision units being derived by purely formal deduction. An analyst can deduce the decisions

> of a single rational mind if he knows the criteria that govern the decisions; but he cannot infer by purely formal analysis what can pass between two centers of consciousness. It takes at least two people to test that. (Two analysts can do it, but only by using themselves as subjects in an experiment.)
>
> (Schelling 1960: 163)

If neither introspective nor deductive reasoning can disentangle a player's expectations from others, the differentiation of individuals becomes an essential pre-requisite to describe decision processes in strategic environments. Game players must be clearly defined to replicate and to investigate their actual reasoning processes. But it is not only Schelling's empiricism that motivates the assumption of player heterogeneity. It also aims at obtaining significant theoretical results. First, heterogeneous players can be assumed to perceive the same game differently. Second, the process of belief formation becomes a creative act individually performed by each player. Third, the game solution becomes the result of a time-consuming dynamic process, which does not necessarily converge to equilibrium.

The first consequence of labelling and differentiating between players is to reintroduce in the analysis those psychological elements removed by the assumption of the representative agent. According to the game theorist Ariel Rubinstein, the primary task of the game modeller is to choose 'only those factors which are perceived by the players to be relevant' (Rubinstein 1991: 919).[8] Once these elements are included in the formal description of the game, the game theorist has to conform to the property of external symmetry by assuming that the players are all alike. Schelling offers a clear-cut argument to claim that this approach results in too specific if not misleading interpretations: 'If two players are themselves mathematical game theorists, they may mutually perceive and be powerfully affected by potential solutions that have compelling mathematical properties' (Schelling 1960: 113). In contrast, non-mathematicians can exploit a wider range of clues to agree on a common solution:

> If the phenomenon of 'rational agreement' is fundamentally psychic – convergence of expectations – there is no presumption that mathematical game theory is essential to the process of reaching agreement, hence no basis for presuming that mathematics is a main source of inspiration in the convergence process.
>
> (Schelling 1960: 114)

Inspiration will be provided by mental reasoning in a way that is difficult to generalize:

> But where do the patterns come from? They are not very richly provided by the mathematical structure of the game, particularly since we have purposely made each player's value system too uncertain to the other to

> make considerations of symmetry, equality, and so forth, of any great help. Presumably, they find their patterns in such things as natural boundaries, familiar political groupings, the economic characteristics of states that might enter their value systems, Gestalt psychology, and any clichés or traditions that they can work out for themselves in the process of play.
>
> (Schelling 1960: 104)

This indeterminateness also extends to the way beliefs are formed. Heterogeneous players can follow different processes of belief formation, which become individually specific acts. As is well known, Nash equilibrium requires that all players expect the same strategy combination to be chosen. This requires that each player adopts the same rules of logical inference to form his or her beliefs on other players' choices. Schelling abandons such a perfect symmetry and, as a consequence, invalidates the theoretical foundation of Nash equilibrium.

In Schelling's games, each player first has to figure out how the other player mentally represents the game, and then what that other player's own beliefs are about his or her expectations. It is the awareness that introspection or 'putting oneself in another person's shoes' do not suffice for forming correct expectations that enables Schelling to describe the infinite regress issue better than his contemporaries. The first economist to point out this problem was Oskar Morgenstern, who applied it to the Conan Doyle's story of Holmes and Moriarty (Morgenstern 1935). In that context, Schelling provides the following definition of the infinite regress problem: 'the best choice for either depends on what he expects the other to do, knowing that the other is similarly guided, so that each is aware that each must try to guess what the second guesses the first will guess the second to guess and so on, in the familiar spiral of reciprocal expectations' (Schelling 1960: 87). It is exactly the assumption of heterogeneous players which makes the task of breaking this impasse a non-trivial one. If players have different expectations in the same game situation, not necessarily due to differences in their information or past experience, but because they are heterogeneous, the problem of infinite regress emerges in all its complexity. While Schelling leaves it open for empirical analysis to resolve, later game theorists will remove it by founding the assumption of common knowledge on the Harsanyi doctrine of common priors (Rizvi 1994).

Thirdly, Schelling's approach implies that game solutions do not necessarily converge to an equilibrium. The convergence of player strategies becomes the result of a dynamic process in which reciprocal expectations are adjusted through a time-consuming procedure. This path cannot be truncated by the instantaneous convergence to equilibrium because the individual processes of learning are differentiated. There is no way out from the modelling of this dynamic procedure: 'An equilibrium is simply a result. It is what is there after something has settled down, if something ever does

settle down. The idea of equilibrium is an acknowledgment that there are adjustment processes' (Schelling 1978: 26).

Schelling imposes no restrictions whatsoever on these processes. For example, in the negotiation process of the bargaining game any possible move before the deadline is admissible (Schelling 1960: 271–2). In this way, the agreement depends on the congeries of psychological factors characterizing actual reasoning and even rational players do not necessarily converge to an equilibrium. In the absence of common priors, non-equilibrium solutions can also be rationalized by the endless sequence of if-I-think-that-you-think-that-I-think (Crawford 1991: 279). This does not imply, however, that Schelling's method of solving games can be reconciled with the concept of rationalizability à la Bernheim, because deviations from perfectly consistent and optimizing behaviour are not necessarily correlated. As Schelling points out,[9] departures from rationality are in many directions also because individuals are different from each other.

In view of the two interpretations given above, Schelling applies game theory neither by empirically expanding its formal apparatus nor by anticipating the later concept of rationalizability. He considers games as frames to be peopled with real and heterogeneous players. The abandonment of the notion of representative agent provides the basis for Schelling's empiricism, according to which theoretical results cannot be defined separately from the actual environment to which they apply. It is as if Schelling always tried to solve games by performing an experiment, which may be just a mental simulation or even a full-fledged laboratory or field experiment.[10] Be it mental or real-life, Schelling follows a three-step procedure, which configures a sort of inductive game theory. The first step is the identification of the players or the player groups to make their distinctive characteristics explicit. Then, in the second step Schelling describes the real environment in which the game is played by specifying the actual case to be considered. In the last step inductive reasoning is used to predict or to prescribe the game solution. What will be argued in the next section is that Schelling has applied this same approach to all his contributions to economics.

3 PLAYER HETEROGENEITY AT WORK

Schelling began his career as an economist first by joining the economics faculty at Yale and then by being appointed Professor of Economics at Harvard in 1958. After some theoretical works on growth and international trade, in the 1950s he started to focus on microeconomic issues that could be formulated in terms of game theory. *The Strategy of Conflict* collected this work that was immediately perceived as innovative and insightful by the economic community.[11] Despite this positive reception, the co-founder of game theory, Oskar Morgenstern, reviewed the book in terms anything but

enthusiastic. His final judgment sounded almost sarcastic: 'the many misunderstandings about game theory to be found in this book would appear to be almost dangerous, should they have an influence upon policy making' (Morgenstern 1961: 104).

What Morgenstern's words document is that Schelling was immediately perceived by game theorists as an outsider.[12] Although in the 1950s Schelling was actively involved in the Rand Corporation, where the mathematical foundations of game theory were being laid, he addressed some criticisms to game theory that conflicted with some of the principles endorsed by that same community.[13] His main argument was that game theory 'has suffered from too great a willingness of social scientists to treat the subject as though it were, or should be, solely a branch of mathematics' (Schelling 1960: 10). This approach had produced models characterized by excessive abstractness, with the excessive emphasis on two-player zero-sum game being a major drawback. Positions such as these might be considered reason enough for the co-author of *Theory of Games and Economic Behavior* to express the critical remarks quoted above.

But Schelling also addressed his criticisms to the later Nobel Prize winners John Nash and John Harsanyi, who were developing – in the same years – the theoretical foundations of game theory.[14] The object was the assumption of symmetry that was introduced by Nash (1950, 1953) to single out a solution of his axiomatic bargaining model. Later, the axiom was discussed and extended by Harsanyi (1956), who proved the mathematical equivalence between Nash's and Zeuthen's (1930) bargaining models. In that paper Harsanyi gave the following definition of the axiom: 'The bargaining parties will follow identical (symmetric) rules of behaviour (whether because they follow the same principles of rational behaviour or because they are subject to the same psychological laws)' (Harsanyi 1956: 149). Schelling's attention was so attracted by this definition that he wrote a paper entitled 'For the abandonment of symmetry in game', which was included as Appendix B to *The Strategy of Conflict*.

Schelling gives a critical assessment of the axiom for two reasons: first, it converts a cooperative bargaining game into a tacit non-cooperative game by excluding any effect of the negotiation process; secondly, symmetry is only one of the many possible clues for determining the rational solution to the bargaining game. To support his view, Schelling uses the metaphor of a race in which all the runners know that they are equally fast. If the race is certain to end in a tie, the runners do not bother to run: 'The perfectly move-symmetrical cooperative game seems a little like that foot race. Bargaining in the one case is as unavailing as leg-work in the other; every player knows in advance that all moves and tactics are foredoomed to neutralization by the symmetrical potentialities available to his opponent' (Schelling 1960: 276). If different bargainers are formally treated as if they are identical, the model abstracts from real individuals and typifies them into artificial beings

which recall the neoclassical representative agent. For Nash and Harsanyi, this assumption does not necessarily commit one to the hypothesis that all players are equal due to the fact that they are univocally defined and differentiated in terms of their preferences. But Schelling deals with a strategic situation as bargaining quite differently. His theory of conflict is based on the hypothesis that 'coordination is not a matter of guessing what the "average man" will do' (Schelling 1960: 92). If it is assumed that bargainers may have different abilities and expectations, the convergence towards a solution becomes the outcome of an endogenous process of adjustment between heterogeneous players. Thus, there is no axiomatic way to define either equilibrium solution or even rational behaviour, because differences between bargainers' expectations can lead to differences between the agreements they expect.

This possibility also motivates Schelling's proposal of focal points. Indeed, a theory of tacit coordination is significant only if player expectations are not coordinated *ex-ante* by the mathematical definition of the game. When all the complications are swept aside by assuming that players are 'symmetric', the problem of coordination failure is practically sidestepped by the hypothesis of complete information. Player heterogeneity precludes such a way out and makes tacit coordination the outcome of the real interaction process: 'A prime characteristic of most of these "solutions" to the problems, that is, of the clues or coordinators or focal points, is some kind of prominence or conspicuousness. But it is a prominence that depends on time and place and who the people are' (Schelling 1960: 57–8).

This intrinsically empirical foundation justifies the basic indeterminacy of Schelling's theory of focal points. In this light, some recent attempts to formalize focal points in terms of team reasoning appear to be disputable. These models postulate a distinctive mode of rational choice in which the identification of the player with a group would facilitate the convergence of player expectations.[15] On the contrary, Schelling's inductive game theory takes full account of all the complexities inherent in the problem at hand. Thus, it does not aim to give prescriptions if there is an obvious way to play the game and each player knows what the others are going to choose. Rather, it addresses another problem: only if there is not an obvious play, do players have to find a focal point on which expectations of heterogeneous players may converge. In contrast with the standard definition of rationality, Schelling postulates a process of convergence in which if the player knows that the other player might choose one strategy because it is the maximizing strategy, it may not be a good tactic to coordinate his or her choice on such an expectation. In order to coordinate choices, it is better to reject the individual maximizing outcomes and converge over a different solution that can be identified separately by both players.

The effect of player differentiation also concerns the mental representation of the game. It is not so much the mere inclusion in the analysis of the

psychological determinants of behaviour that allows one to identify the saliency of the focal point, but the interpretations specifically given by each player to any element useful for this purpose. Schelling's three-stage procedure serves to single out the most suitable cognitive procedure among the many possibilities. Holmes and Moriarty's pursuit across train stations is again revealing of Schelling's method. The pursuit can be ended by using the stations as coordinating hints but the solution cannot be found deductively:

> But in the common-interest version they must somehow *use* the labeling of the stations in order to do better than pure chance; and how to use it may depend more on imagination than on logic, more on poetry or humor than on mathematics. It is noteworthy that traditional game theory does not assign a "value" to this game: how well people can concert in this fashion is something that, though hopefully amenable to systematic analysis, cannot be discovered by reasoning a priori. This corner of game theory is *inherently* dependent on empirical evidence.
>
> (Schelling 1960: 97–8)

Coordination strictly depends on how Sherlock Holmes and Professor James Moriarty mentally represent the environment in which they play the game. To solve the game, the formal framework has to be complemented with any device the two players may find useful for their specific purposes. Consequently, there does not exist a correct way to play disjoint from the actual play and the real identities of the two players.

If players are heterogeneous with regard to decision processes they have also to be considered from different each other in order to derive a social welfare function. In this case, as argued by Kirman (1992), the hypothesis of representative agent can indeed be seriously misleading. Schelling's book *Micromotives and Macrobehavior* focuses on this topic by relying on the hypothesis of heterogeneous players to obtain its main result, which is the non-intentionality of collective choices.

Schelling's starting point is the criticism of the way economists deal with this issue: 'Actually, economists do not usually make careful observations, compare what they observe with alternatives they can imagine, and judge the results to be good. What they do is to infer, from what they take to be the behavior characteristics of people, some of the characteristics of the system as a whole, and *deduce* some evaluative conclusions' (Schelling 1978: 22). In contrast, in Schelling's inductive game theory player heterogeneity impedes the drawing of this inference from the micro to the macro level. As the following quotation shows, Schelling's analysis displays a great variety of social behaviour:

> To continue with our listing, the subject includes systems of deference, etiquette, social status, and hierarchy. It includes 'street behavior' – being

> on the streets or staying off; staring ahead or nodding hello; asking for directions, matches, the time of day, or spare change; and carrying weapons. It includes the formation of mobs and riots, panic behavior, rules of the road, traffic conventions, and the signals and insignia by which people recognize each other. It includes style and taste, hairstyles and cosmetics, clothing styles and jewelry, patterns of eating and drinking, coffee breaks and cocktail hours, tobacco, marijuana, littering and jaywalking, obeying and disobeying the law, and coming or not coming to help if somebody is in distress. I want to avoid any suggestion that there is some single mechanism that underlies all of these behaviors. Quite the contrary.
>
> (Schelling 1978: 41)

To deal with this multiform universe, Schelling uses the concepts of *critical mass* and *critical number* – which define respectively the existence of network externalities making a player's utility of a good dependent on the number of players already owning said good and the numerical threshold to be overcome to form the critical mass – and the three-step inductive procedure outlined above.

The first stage is to introduce heterogeneity by assuming that the critical number for one player differs from that of another. Then, in the second stage, Schelling defines the real environment in which the game is played. In this case, he considers the phenomenon of 'tipping', which is the cumulative processes of migration occurring in the 1960s within US cities of even a small fraction of non-whites causing the neighbourhood to 'tip' from totally white to totally non-white. Finally, in the third stage Schelling defines inductively the solution of the game by performing simple simulations on a checkerboard of sixty-four squares representing places where people live. By moving two different kinds of pieces according to various decision rules, he shows that very small changes in the preference for proximity to pieces of the same kind or minor exogenous disturbances can trigger a chain reaction that results in extreme segregation. More significantly, the simulation shows that this outcome, as with those actually verified in many US neighbourhoods, can hardly ever be reversed.

Significantly enough, Schelling's exercise reveals the basic unpredictability of the final outcome: 'What is instructive about the experiment is the ''unravelling'' process. Everybody who selects a new environment affects the environments of those he leaves and those he moves among. There is a chain reaction. It may be quickly damped, with little motion, or it may go on and on with striking results' (Schelling 1978: 150–1). The cascade effects are triggered by the incremental decisions taken by locally interacting players that produce non-linear transformations in aggregate behaviour primarily by virtue of the multiple differentiation of the subjects. In these interactions, players are different from each other not only in terms of social roles and

personal attributes – blacks and whites, and in other applications of the same model, boys and girls, surfers and swimmers, smokers and non-smokers, slow truck drivers and fast drivers – but also with respect to the spatial dimension, which makes their choices dependent on the composition of the neighbourhood.

In *Micromotives and Macrobehavior*, the same method is also applied to explain the emergence of multiple equilibria. The final chapter of the book discusses the case of the adoption of helmets by professional hockey players. What initially seems to motivate Schelling is the variety of reasons for wearing helmets, which are discussed in the *Newsweek* quotation reported in the first part of the chapter. Then, Schelling develops the model according to the now-standard paradigm in network economics of binary choices with externalities. But in another example he postulates that the players are represented by two kinds of skaters, the professionals who like empty ice rinks and the amateurs preferring neither too crowded nor empty rinks (Schelling 1978: 107–9). The aggregation of the preferences of both groups can produce the outcome of no attendance in that the presence of few professional skaters attracts enough amateurs so as to make the rink unattractive to the professionals. But if the number and preferences of professional skaters is such that the above effect is not triggered, two equilibria are possible: one in which the number of amateurs is too high to attract the others, and a second in which too few professionals skate.

With the benefit of hindsight, these stylized examples were bound to inspire the more recent revision of a significant part of macroeconomic literature. Today, Schelling concedes that he was not so prophetic:

> I published my results, and it got quite a bit of attention at the time. But it wasn't until 25 years later that I realized that this game had pioneered some of the work in what is called 'agent-based modeling' and which is used in a variety of disciplines in the social sciences. At the time I was working out this example I didn't realize that I was engaged in an area of research that would one day have a formal name.
>
> (Schelling 2005: 41)

In fact, the key hypothesis of this approach is that economies are not just collections of homogeneous agents but complex dynamic systems characterized by dispersed interaction among heterogeneous agents acting locally on each other in some space (Arthur *et al.* 1997). Schelling was not the first but he was presumably the most persuasive scholar to show that the mistakes caused by inferring the whole from a part, the so-called fallacy of composition, could be corrected by postulating player heterogeneity.

The direct reference to real cases is also the main reason which induced Schelling to propose the technique of self-management in the 1980s. He gives this account of how he started thinking about it:

> I argued that there was an important area of behavior that they were overlooking, namely, the ways that people can cope with their own misbehavior. I remember that they argued, especially with respect to heroin addiction, that there was very little that individuals could do for themselves. But as I listened to them describe relapse, I remembered a description of a reformed heroin addict who took a bus through town. The bus happened to stop at the very place where he had previously met with others to shoot up heroin. He began to get withdrawal symptoms, and before he'd gone very far, he'd gotten off the bus, walked back, looked for his old friends, and was soon back on heroin. And I thought, well now, there's a piece of advice you could give an addict, namely, don't take the bus that goes past the place where you're likely to succumb to the physical urge to seek out the heroin. This seemed to me a simple thing that one could do for oneself: avoid the friends, places, or other cues or stimuli associated with one's old habit. So I began to observe the things that people do or try to do in disciplining their own behavior.
>
> (Iglehart 1990: 110–11)

As the quotation suggests, Schelling follows his three-step method even to analyse temporally inconsistent behaviour. First he considers real players like drug addicts or smokers, then he observes real choices such as the situation quoted above, and finally he draws, inductively, some general precepts. The consequence of this empirically driven approach is again to take into full account, as with interpersonal relations, all the complexities of the preference formation process. Individual choices become the outcome of a game played by competing multiple selves, which are not hierarchically ordered. In fact, there is no 'super-self' nor is there a 'metapreference system' to define a solution to this game:

> In this case I want to discuss, that superself, that dynamically programming referee, does not exist. Instead, there is a succession or alternation of impermanent selves, each in command part of the time, each with its own needs and desires during the time it is in command but having – at least some of them – strong preferences about what is done during the period that another one is in command.
>
> (Schelling 1984a: 86–7)

Similarly to what he does for the fallacy of composition in macroeconomics, Schelling raises a microeconomic issue of considerable theoretical importance and starts to empirically inquire how it can be solved:

> The conclusion I come to is that this phenomenon of rational strategic interaction among alternating preferences is a significant part of most people's decisions and welfare and cannot be left out of our account of the consumer. We ignore too many important purposive behaviors if we

> insist on treating the consumer as having only values and preferences that are uniform over time, even short periods of time.
>
> (Schelling 1984b: 5)

Schelling offers many 'characterizations' to individuate the various identities playing the intrapersonal game. Their common feature is that they 'invite efforts at anticipatory self-command' (Schelling 1984b: 4) but each of them require an *ad hoc* analysis. However, if the behaviour of a single individual is the result of the interaction of a collective of – albeit rational – agents, there will not be any certainty to observe choices consistent with an abstract definition of rationality. Individual payoff maximization becomes only one of the possible decision criteria because both any intertemporal basis for calculating the best preferred alternative and any formal device for defining how later choices are determined by past decisions are rejected.

Schelling invalidates time preference as an analytical tool with examples so common and yet so witty as to be hardly forgettable:

> But the person who is simply not thinking of the future, who shuts his eyes to avoid it (especially when the future is not a decade away but tomorrow, when he knows he'll suffer remorse and disgust and the disapproval of family and friend who witness the collapse of resolve), like the person who furiously scratches his hives, would have to be someone whose time discount is 100 percent per hour or per minute, compounding to an annual rate too large for my calculator.
>
> (Schelling 1984a: 62–3)

Moreover, the assumption of multiple selves implies that individual preferences are neither unchanging nor exogenous with respect to the act of choice. They become necessarily an *ex-post* concept which may be defined only after having made the potential choices actual. This is because choices are not only the result of the interaction of multiple selves but also the product of the activity of self-management, through which the decision maker tries to effectively do what he has planned to do. From a theoretical point of view, the final outcome of this process becomes hardly predictable because it depends on the complex set of tactics and techniques that individuals learn by trying 'to influence and constrain the behavior of others' (Schelling 1984a: 69). In this reversal of the causal arrow going from introspective to external sources of knowledge, it is easy to recognize the crucialness of the assumption of player heterogeneity. In effect, if there is no single agent to represent the consumer but a collection of heterogeneous selves alternating among each other, the rankings of preferences are multiple or, even if a single ranking is computable, it does not necessarily imply maximization (Rizvi 2001). The players do not adhere to deterministic decision rules and individual choices do not allow inferring individual values. The approach of revealed preference is

basically falsified. Once again, notwithstanding his efforts to be conciliatory with the mainstream, Schelling departs from standard decision theory in a way that asks for rebuilding of its main foundations. By also taking the hypothesis of heterogeneity within the self, Schelling makes it the keystone of his methodology.

4 CONCLUSIONS

The main thesis of this paper is that Schelling's empiricism is deeply grounded in the assumption of heterogeneous players. He peoples games with real individuals and consequently postulates that there are differences in roles or identities among them. The subjects populating Schelling's thought, field or laboratory experiments do not assume that other players follow symmetrical or identical rules of logical inference to decide their strategies. This hypothesis is decisive in shaping Schelling's inductive game theory, which is applied by means of a three-step procedure. First, players are defined by making their differentiating features explicit. Secondly, heterogeneous players are embedded in a real environment to play the game. Thirdly, the game solution is derived inductively.

The abandonment of the neoclassical construct of representative agent has important theoretical implications. There is no way to define deductively how player beliefs converge because this process is the outcome of individually specific processes. In the same game different players may have different expectations of the other players' strategies. As a consequence, the game solution is not the result of an instantaneous convergence to equilibrium but the final step of the dynamic adjustment process of the beliefs of heterogeneous players.

This interpretation helps to explain why little progress has been made overall in developing Schelling's insights. His inductive game theory does not aim either to integrate mathematical game theory with decision criteria other than self-seeking maximization or to found, empirically, the later concept of correlated equilibrium. By removing the assumption of external symmetry, according to which players are all alike out of the formal description of the game, Schelling challenges standard economic theory and, more specifically, the way in which it has incorporated the mathematics of game theory. The hard core of game theoretical economics is represented by the modelling of a player's behaviour in terms of deductive and abstract rationality. This method heavily relies on the nearly mythical figure of the player, which is able to perform all the tasks needed to replicate the game theorist's intuition. Schelling's heterogeneous players are only required to conform to their actual behaviour and the theorist's task becomes that of embedding them in their specific environment.

There is a recurrent question addressed to Schelling whenever he is interviewed, namely, whether he considers himself a dissenter from

mainstream economics. Invariably, Schelling gives an open-ended answer. He starts by admitting that the rational approach to economics is a valuable one, and then he acknowledges being interested in something different. What this paper has meant to illustrate is that Schelling has steadfastly turned aside from the abstract and deductive approach of mainstream economics. His focus has consistently been towards the production of empirical generalizations that, by taking into account all the complexities of real life, could be directly relevant for describing and prescribing behaviour. In the development of this research programme the assumption of player heterogeneity has been essential.

Alessandro Innocenti
University of Siena
innocenti@unisi.it

NOTES

1 '[…] even though his [Schelling's] writings are anything but mainstream, he hardly ever writes in an adversarial, disputatious tone. He leaves it to others to expatiate on the "failures" or the "irrelevance" of economics, on "What's wrong with economics". He simply walks away from the whole scene. […] It may therefore be conjectured that Tom refrains from the imprecations because he has absolutely no use for the related activity: the construction of an alternative paradigm. He does not want to fall into the trap of exchanging one straitjacket for another.' Albert Hirschmann quoted in Zeckhauser (1989: 162–3). On this point see also Aydinonat's (2005) interview with Schelling.

2 'One of the first things that strike a social scientist when he begins to experiment with illustrative matrices is how rich in variety the relationship can be even between two individuals, and how many different meanings there are for such simple notions as "threat," "agreement," and "conflict." He is struck by how many configurations of information and misinformation there are, how many different communication systems, and what a variety of alternative "legal" constraints on bargaining and tactics. Even the simplest of situations, involving two individuals with two alternatives apiece to choose from, cannot be exhaustively analyzed and catalogued. Their possibilities are almost limitless. For this reason, game theory is more than a "theory," more than a set of theorems and solutions; it is a framework for analysis. And for a social scientist the framework can be useful in the development of his own theory' (Schelling 1984a: 221–2).

3 For a discussion of the conception of the individual in game theory, see Davis (2003).

4 Dixit (2006) can also be considered sympathetic with this view: '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' (Dixit 2006: 216–17).

5 Schelling deals with this issue in other papers of his books and in some interviews (Iglehart 1990; Aydinonat 2005; Schelling 2005).

6 The list includes car and helicopter drivers, people escaping from a burning building, social scientists, rangers, dairy farmers, taxi drivers, and so on.

7 Dixit (2006: 225–7) examines the theoretical implications of this 'vivid' example, which is also discussed in section 2 above.

8 'If we adopt the view that a game is not a rigid description of the physical rules of the world, then a game-theoretic model should include only those factors which are perceived by the players to be relevant. Modeling requires intuition, common sense, and empirical data in order to determine the relevant factors entering into the players' strategic considerations and should thus be included in the model' (Rubinstein 1991: 919).

9 'Rationality is a collection of attributes, and departures from complete rationality may be in many different directions. Irrationality can imply a disorderly and inconsistent value system, faulty calculation, an inability to receive messages or to communicate efficiently; it can imply random or haphazard influences in the reaching of decisions or the transmission of them, or in the receipt or conveyance of information; and it sometimes merely reflects the collective nature of a decision among individuals who do not have identical value systems and whose organizational arrangements and communication systems do not cause them to act like a single entity' (Schelling 1960: 16).

10 Schelling's empiricism leads him fairly naturally to embrace laboratory testing as a useful tool for the economist. A whole chapter of *The Strategy of Conflict* is devoted to illustrating the usefulness of experimental research for game theory.

11 See Bishop's (1961) review published in *The American Economic Review*, Baumol's (1961) in *The Review of Economics and Statistics*, Barbier's (1963) in *Econometrica*, and Meade's (1963) in *The Economic Journal*.

12 Martin Shubik (1961), Morgenstern's student, also harshly criticized Schelling's book in a review published in *The Journal of Political Economy*.

13 Schelling acknowledges his debt to The RAND Corporation in the preface of *The Strategy of Conflict*. After having defined as 'superb' the 'collection of people' working there, Schelling writes: 'RAND is not responsible for the shapes my ideas have taken – the "views herein expressed" – but I hope it will, as a corporation, take satisfaction from its responsibility for some of the ideas' taking any shape at all' (Schelling 1960: VI).

14 This debate and its historical consequences are discussed in Innocenti (2005).

15 Janssen (2001) and Sudgen and Zamarrón (2006) provide a survey of these contributions.

REFERENCES

Arthur, B. W., Durlauf, S. N. and Lane, D. (1997) 'Introduction', in B. W. Arthur, S. N. Durlauf and D. Lane (eds). *The Economy as an Evolving Complex System* II, Santa Fe Institute Studies in the Sciences of Complexity Lecture Notes, Santa Fe Institute: Perseus Books, pp. 1–14.

Aumann, R. J. (1976) 'Agreeing to disagree', *Annals of Statistics* 4: 1236–9.
Aumann, R. J. (1985) 'What is game theory trying to accomplish?', in K. J. Arrow and S. Honkapohja (eds) *Frontiers of Economics*, Oxford: Basil Blackwell, pp. 28–76.
Aydinonat, N. E. (2005) 'An interview with Thomas C. Schelling: interpretation of game theory and the checkerboard model', *Economic Bulletin* 2: 1–7.
Barbier, M. (1963) 'Book review: *The Strategy of Conflict*', *Econometrica* 31: 614.
Baumol, W. J. (1961) 'Book review: *The Strategy of Conflict*', *Review of Economics and Statistics* 43: 391–2.
Bishop, R. L. (1961) 'Book review: *The Strategy of Conflict* by Thomas C. Schelling', *American Economic Review* 51(4): 674–6.
Crawford, V. (1991) 'Thomas Schelling and the analysis of strategic behavior', in R. J. Zeckhauser (ed.) *Strategy and Choice*, Cambridge, MA and London: MIT Press, pp. 265–94.
Dardi, M. (2006) 'Schelling's "clues" and Bayesian rationality', University of Firenze, mimeo.
Davis, J. B. (2003) *The Theory of Individual in Economics. Identity and Value*, London and New York: Routledge.
Dixit, A. (2006) 'Thomas Schelling's contributions to game theory', *Scandinavian Journal of Economics* 108: 213–29.
Harsanyi, J. (1956) 'Approaches to the bargaining problem before and after the theory of games: a critical discussion of Zeuthen's, Hicks', and Nash's theories', *Econometrica* 24: 144–57.
Iglehart, J. K. (1990) 'Perspectives of an errant economist: a conversation with Tom Schelling', *Health Affairs* 9: 109–21.
Innocenti, A. (2005) 'Linking strategic interaction and bargaining theory: the Harsanyi–Schelling debate on the axiom of symmetry', *Quaderni del Dipartimento di Economia Politica Università di Siena*. 468 (forthcoming, *History of Political Economy*).
Janssen, M. (2001) 'Rationalizing focal points', *Theory and Decision* 50: 119–48.
Kirman, A. P. (1992) 'Whom or what does the representative individual represent', *Journal of Economic Perspectives* 6: 117–32.
Meade, J. E. (1963) 'Book reviews: *Conflict and Defense: A General Theory*; *The Strategy of Conflict*; *Fights, Games and Debates*', *Economic Journal* 73: 284–95.
Morgenstern, O. (1935) 'Vollkommene Voraussicht und wirtschaftliches Gleichgewicht', *Zeitschrift für Nationalökonomie* 6: 337–57, English translation 'Perfect foresight and economic equilibrium', in A. Schotter (ed.) *Selected Economic Writings of Oskar Morgenstern*, 1976, New York: New York University Press, pp. 169–83.
Morgenstern, O. (1961) 'Book reviews: *Fights, Games and Debates*; *The Strategy of Conflict*', *Southern Economic Journal* 28: 103–5.
Nash, J. (1953) 'Two-person cooperative games', *Econometrica* 21: 128–40.
Nash Jr, J. F. (1950) 'The bargaining problem', *Econometrica* 18: 155–62.
Rizvi, S. A. T. (1994) 'Game theory to the rescue', *Contributions to Political Economy* 13: 1–28.
Rizvi, S. A. T. (2001) 'Preference formation and the axioms of choice', *Review of Political Economy* 13: 141–59.
Rubinstein, A. (1991) 'Comments on the interpretation of game theory', *Econometrica* 59: 909–24.
Schelling, T. C. (1960) *The Strategy of Conflict*, Cambridge, MA: Harvard University Press.

Schelling, T. C. (1978) *Micromotives and Macrobehavior*, New York and London: W.W. Norton & Co.
Schelling, T. C. (1984a) *Choice and Consequence*, Cambridge, MA and London: Harvard University Press.
Schelling, T. C. (1984b) 'Self-command in practice, in policy, and in a theory of rational choice', *AEA Papers and Proceedings* 74: 1–11, reprinted in Schelling 2006.
Schelling, T. C. (2005) 'Interview: Thomas Schelling', *Region Focus* Spring: 36–41.
Shubik, M. (1961) 'Book reviews: *The Strategy of Conflict*; *Fights, Games and Debates*', *Journal of Political Economy* 69: 501–3.
Sugden, R. (2001) 'The evolutionary turn in game theory', *Journal of Economic Methodology* 8: 113–30.
Sudgen, R. and Zamarrón, I. E. (2006) 'Finding the key: the riddle of focal points', School of Economics and School of Philosophy, University of East Anglia, mimeo.
Zeckhauser, R. (1989) 'Distinguished fellow: reflections on Thomas Schelling', *Journal of Economic Perspectives* 3: 153–64.
Zeuthen, F. (1930) *Problems of Monopoly and Economic Warfare*, London: Routledge & Kegan Paul.