EXPERIMENTAL GAME THEORY

Werner Güth, Department of Economics, Humboldt-University of Berlin, Germany

Keywords: common knowledge, repeated play, ultimatum game, adaptation dynamics, strategic bargaining, characteristic function experiments, random price mechanism

Contents List:

Glossary:

Common knowledge of rationality: All players are rational, all players know that all players are rational, all players know that all players know that all players are rational etc.

Random price mechanism: A price is randomly selected from some interval including the true value. Trader chooses a price limit meaning that he buys (sells) at all prices below (above) his limit. Thus it is the only undominated strategy to bid one’s true value.

Tournament: Computerized strategies are collected and – mostly pairwise – matched in all possible pairs so that they finally can be ranked according to success.

Bounded rationality: Takes into account the limited cognitive and information processing abilities of human decision makers, should avoid any optimization hypothesis, concept should be based on empirical findings rather than on abstract reasoning.

Preplay communication: Strategically more or less structured communication phase before actually playing, can range from eye contact to non-binding promises what one will play.

Ultimatum bargaining: First the proposers offers a share of the “pie” to the responder who then can accept or reject the offer; in case of an acceptance the “pie” is shared as proposed, in case of a rejection both players receive 0-payments.

Alternating offer bargaining: Two parties take turns in proposing an agreement which the other can accept or reject where the “pie”, the monetary amount to be distributed, can depend on when an agreement is reached.

Summary: Whereas orthodox game theory relies on the unrealistic assumption of (commonly known) perfect rationality, participants in game playing experiments are at best boundedly rational. This makes it necessary to supplement orthodox game theory by a behavioral theory of game playing. We first point out that this applies also to (one person-) decision theory. After reviewing the influential experiments based on repeated games and the ultimatum game the typical reactions to the striking experimental results are categorized. Further sections are devoted to alternating offer bargaining and characteristic function experiments.

1. Introduction

Orthodox game theory relies on (common knowledge of) perfect decision rationality, i.e. unlimited cognitive and information processing capabilities of players. Even for finite games of perfect information like chess it is, however, obvious that these requirements are far beyond what human decision makers can accomplish.

Another problem in applying orthodox game theory is that it assumes individual cardinal utilities and subjective beliefs which can hardly ever be observed provided they exist at all. Of course, one may specify utilities by material payoffs like profits (which often can be observed) and beliefs by objective probabilities whenever possible, but then the game theoretic predictions are often not confirmed by experimental observations.

At present we can look back at half a century of experimental game theory. One has tried to implement game theoretic models as (laboratory) experiments and to test orthodox game theory (assuming its applicability, e.g. by specifying utilities by material payoffs and beliefs by appropriately designed chance moves) and to supplement orthodox game theory by behavioral concepts since narrowly defined orthodox game theory is often falsified.

In the following we start by discussing decision theoretic experiments. Regarding games with interpersonal strategic interaction we somewhat reverse the chronological order by first reporting about experiments, based on noncooperative games, although, similar to orthodox game theory, experimental game theory initially was dominated by cooperative game experiments.

2. One person – decision making

Testing experimentally rationality in one person-games means mainly to test (the axioms of) utility theory. There is probably no need to prove that human players will be unable to solve optimization tasks involving complex combinatorics. Clearly people differ in their capabilities, but it is a fact of human existence that even for the most capable among us one can find optimization problems with finite sets of choice alternatives which they cannot solve. Let us therefore review the experimental evidence for three choice problems whose degree of difficulty differs greatly, namely dynamic optimization, risky choices, and dominance solvability in one person-games.

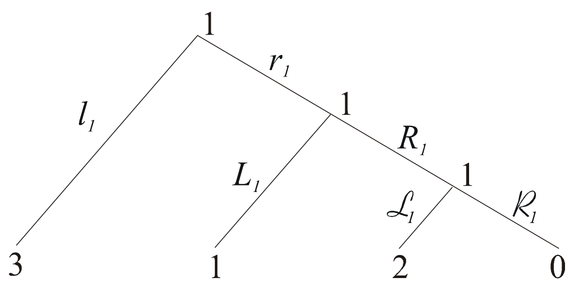

Figure 1: A simple dynamic one person-game

The simple dynamic one person-game in Figure 1 has the obvious solution (l_1, R_1, $\mathcal{L}_1$). It illustrates that the discussion whether (common knowledge of) rationality can be assumed at all decision nodes cannot be restricted to interpersonal strategic interaction. If player 1 has actually to choose among L_1 and R_1, his optimal choice R_1 anticipates own future rationality by preferring $\mathcal{L}_1$ over $\mathcal{R}_1$. But the fact that one must choose among L_1 and R_1 seems to question own rationality. A rational player 1 would choose l_1 (yielding the payoff of 3) and thus will never actually face the choice between L_1 and R_1.

In more complex dynamic games initial suboptimal choices are the rule rather than the exception. Such problems are more thoroughly studied in experimental (and cognitive) psychology. In experimental game theory (and economics) one tests the quantitative or at least qualitative aspects of rational behavior based on models of intertemporal allocation behavior like rational addiction or the so-called life cycle saving models. The data even of experienced participants clearly reveal that participants, if they at all behave systematically, rely on heuristics rather than backward induction, respectively dynamic programming.

Most of the decision theoretic experiments by experimental game theorists focus on (axioms of) cardinal utilities where the experimental procedures vary a lot, e.g. from pure questionnaires without monetary incentives to those where non-optimality implies substantial losses. Partly one argues that deviations from optimal behavior should only be taken seriously when they imply substantial losses (saliency). Although the theory of rational decision making does not offer any guidance here, the saliency requirement can be justified by (uncontrolled and small) costs of optimizing, e.g. the disutility of deriving an optimal alternative. Mostly a participant is asked repeatedly to choose between two simple lotteries of the form $L = \left(\underline{P} \middle| w, \overline{P} \middle| 1-w\right)$ with $0 \leq \underline{P} < \overline{P}$ and $0 < w < 1$ where sometimes the monetary prizes $\underline{P}$ and $\overline{P}$ resulting with probability w, respectively $1-w$ are substituted by lotteries. Comparing all these choices reveals usually that certain axioms of expected utility maximization are violated.

Explaining some so-called paradoxes (referring to experimental results as paradoxical only shows how naive one's expectations have been) by individual risk attitudes can be experimentally excluded by applying the binary lottery technique. Here the payoff of a participant is the probability $1-w$ of winning the high prize $\overline{P}$ instead of the low prize $\underline{P}$ resulting with complementary probability w. By setting $u(\underline{P}) = 0$ and $u(\overline{P}) = 1$ the utility $u(L) = 1-w$ results, i.e. utility depends linearly on the experimental payoff variable $1-w$ (implying risk neutrality).

When trying to observe experimentally how a lottery L is evaluated one may rely on mechanisms for which truthful revelation of preferences is optimal, e.g. in the sense that truthful revelation is the only undominated value statement. Let $[a,b] \subset \mathbb{R}$ be an interval with $a < b$ and $\varphi(\cdot)$ a density over $\mathbb{R}$ with $\varphi(p) = 0$ for all $p \notin [a,b]$ and $\infty > \varphi(p) > 0$ for all $p \in [a,b]$. According to the random price mechanism the price p which one can receive or must pay is randomly determined according to $\varphi(\cdot)$. The only decision variable of a potential seller (whose willingness to accept one wants to observe) or buyer (in whose willingness to pay one is interested) is thus the price limit l meaning that one only sells at prices $p \geq l$, respectively buys at prices $p \leq l$.

Let v with $a < v < b$ denote one's true monetary value for lottery L. Clearly, $l = v$ is the only undominated price limit. Especially, the optimal decision $l = v$ does not depend on a, b, and $\varphi(\cdot)$ as long as $a < v < b$ and the qualitative requirements for $\varphi(\cdot)$ are satisfied. Nevertheless experimentally observed choices l react to such changes of a, b, and $\varphi(\cdot)$. Since dominant choices do not depend on risk attitudes, given true values v seem more a normative concept than a fact of life. Human decision makers do not have (complete) preferences. They rather must generate them by cognitively representing their decision environment, by imagining how choices affect their basic concerns etc. In such a dynamic process aspects, which are strategically irrelevant, may become influential. Consequently the normative concept of incentive compatibility may have little behavioral relevance and reliability. More generally, optimal mechanisms may perform rather poorly. An institution which is best when all parties act rationally can induce non-rational behavior and thus to bad results. This illustrates why orthodox (game) theory has to be supplemented by a behavioral theory of decision making whose formalization could be based on the (stylized) results of decision making experiments.

3. Experimental results in strategic games

The earliest experiments were based on markets which still require an auctioneer. One may thus be reluctant to view them as strategic game experiments. At the same time simple games like prisoners' dilemma games were studied experimentally.

A solution concept for a given class of games, e.g. the class of finite games in normal form, partitions the class into equivalence classes. Two games from the same equivalence class are said to be strategically equivalent. Most solution concepts, for instance, allow for positively affine utility transformations. Experimentally one could try to induce such transformations by appropriately changing the monetary payoffs. The typical result is that such game theoretically irrelevant differences can nevertheless change behavior rather dramatically.

For the case of a prisoners' dilemma games where defection from cooperation always leads to the same payoff advantage, regardless what the other chooses, one can even decompose the same game in infinitely many ways. This leads to a one parameter family of decomposed games which all are identical to the same prisoners' dilemma game. Nevertheless average cooperation ranges from 25 % for certain decompositions up to 75 % for others.

More generally, a strategic game (regardless whether is modelled as a stage game, in extensive or normal form) does not seem to account for all determinants of behavior. On the other hand, as will be illustrated below, it may pay attention to subtle strategic details like the exact timing of decisions and specific information feedback which are often neglected by experimental participants. There will be the need to supplement strategic equivalence by concepts of behavioral equivalence.

A very dominating paradigm in experimental game theory has been the once or repeatedly played prisoners' dilemma game. Notice that in an experiment a finite upper bound for the number of repetitions cannot be avoided and should thus be commonly known. Thus Folk Theorems do not apply and mutual defection is the solution in finitely often repeated prisoners' dilemma games. Nevertheless the robust results of many experiments are that

- players cooperate in most rounds
- although they defect towards the end

what, however, does not unravel cooperation altogether as suggested by backward induction. Since participants try to prevent being preempted, i.e. that their partner terminates cooperation earlier, they seem to be fully aware of the backward induction idea. Because of its detrimental consequences they, however, do not follow its recommendations.

Since Folk Theorems do not apply, this has posed quite a puzzle for game theorists and inspired the innovative reputation (other names are "crazy perturbation" or "gang of the four"-) approach. The basic idea is to allow for (a little bit of) incomplete information concerning the other player's type (in repeated prisoners' dilemma games (s)he may be an unconditional cooperator, i.e. a bit "crazy"). Thus one can try to build up the reputation (the other's posterior probability) of being (confronting) an unconditional cooperator. Although the reputation approach has been extremely successful (in the sense of inspiring a large literature), some qualitative aspects of reputation equilibria are not at all or only poorly supported by experimental data like

- the gradual decline of the probability for cooperation leading to sure defection in the last period (some participants, for instance, cooperate in the last round)

- their specific mixing.

Nevertheless reputation equilibria illustrate how orthodox game theory (reputation equilibria do not question rationality, only that is surely expected) can be enriched by paying attention to robust experimental findings. Other applications study theoretically and experimentally signaling, trust, bargaining and centipede games.

Repeated games are also the main playground of tournament studies. Here participants are invited to develop (mostly computerized) strategies which then encounter each other in a tournament. Usually participants learn about the differential success of their strategy (often the only payoff is just one's ranking) and can try to improve it in the light of such experiences. Such studies definitely induce participants to develop rather general behavioral plans (what to do in all contingencies). They, however, are also biased in the sense that one first of all tries to ensure cooperation. If even on large markets with many pairwise interactions one would have to expect such a drive towards cooperation, antitrust policy would be hopeless and useless. Actually some experiments show that even very narrow (in one study two duopolists interact for 60 rounds) markets may yield results closer to the competitive equilibrium than to the duopoly or cooperative solution.

Another thoroughly explored topic is strategic bargaining (often referred to as the noncooperative approach to bargaining). Experimental game theory mostly concentrated on very simple models of strategic bargaining where the problem is hardly the difficulty of deriving the game theoretic solution, but more to accept it, especially when it is unfair. In ultimatum bargaining experiments a positive amount p of money, the "pie", can be distributed by

- first allowing the proposer to decide on his offer o with $0 \le o \le p$ to the responder

- who then can accept the offer o (so that the proposer gets $p - o$ and the responder o) or reject it (both players get nothing).

Whereas the game theoretic solution predicts that the proposer offers 0 or the smallest positive money unit and that the responder accepts all (positive) offers, the typical and robust experimental findings are that

- responders reject even substantial positive offers o in the range $0 < o < p/2$, which they apparently view as unfair,

- proposers shy away from too low offers o (the most frequent offer is usually the equal split $o = p/2$).

Slight doubts regarding the responder's rationality would not solve the problem here. To justify the offer $o = p/2$ instead of $o = \varepsilon$ (with ε denoting the smallest positive unit of money) the proposer must be extremely risk averse or the responder rather likely irrational or "crazy". Furthermore, this would not explain at all why responders should reject positive offers. How did game theorists react to this challenge?

One escape is to "repair" the game theoretic representation of the experimental situation, e.g. by assuming that utilities depend not only profits, but also on its distribution or on a desire for reciprocity or on what one thinks is expected by the other. These repairs do not (allow to) question decision rationality, but only that experimental situation is adequately represented by the game model. The problem is, of course, that nearly all results can be “saved” in this way. So “repairs” should be reasonable and intuitive. It is, for instance, obvious that we often care for the distribution of rewards, but when and why we do so seems to be already a cognitive result rather than a given.

A more recent route of escape is what one can call the “pour out the baby with the bath water”-approach. Inspired by the development of evolutionary game theory, which partly is based on evolutionary biology and partly revives an earlier research tradition in psychology (of learning theories), one either denies cognition at all (by assuming preprogrammed phenotypical behavior as in orthodox evolutionary biology) or by restricting rationality more or less severely (in reinforcement or stimulus-response learning, for instance, one prefers what previously proved to be better). Initially behavior can be anything, but partly the claim is that it will converge to equilibrium behavior. The most prominent evolutionary dynamics are the replicator dynamics. Other adaptation dynamics are best reply dynamics, reinforcement or stimulus-response-learning, and imitation processes. To determine the stable behavior one partly relies on static concepts, e.g. the one of evolutionarily stable strategies.

This theoretical discussion and the progress in developing software packages for computerized experiments has inspired a very new type of experiments. The same group of participants plays again and again the same simple game (e.g. a 2 x 2-bimatrix game) with randomly changing partners. This research tradition is too recent to justify any general conclusions how people adapt to past experiences and how to combine such path dependence with undeniable strategic deliberations. One rather robust result is that in 2 person-coordination games behavior converges to strict equilibria but not necessarily to the payoff dominating strict equilibrium. It is, however, impressive how closely theoretical exercises of adaptive dynamics and experimental studies are interrelated. If the same simple game is played very often, boredom might, of course, induce variety seeking. In studies of robust learning where one confronts participants repeatedly with a variety of related games instead of just one such game this seems less likely.

A third reaction to the typical results of strategic game experiments is the “bounded rationality approach” which denies all the givens of normative decision and game theory like

- given strategy sets
- given evaluations of plays
- given Bayesian beliefs and
- unlimited cognitive and information processing capabilities.

When to decide and which options one seriously considers is often the first step in one’s strategic deliberation process. Furthermore the evaluation of results is not readily available. Rather one has to derive reasonable goals and their (relative) importance, e.g. in the sense of a multi objective-satisficing approach. In case of conflicting concerns like between the desire to be fair and to earn as much as possible as a proposer in ultimatum bargaining one often avoids cognitive dissonance by suppressing one of the goals (e.g. by viewing the situation simply as a fair division task and offering $o = p/2$). Rather than forming Bayesian beliefs how likely all possible offers o will be accepted one further will simply distinguish offers o (close to $o = p/2$) which are surely acceptable and choose the lowest offer o among those. Finally explanations of experimental results should rely on deliberations which human decision makers are likely to perform when confronting such a decision problem (if game theorists need months to solve a game how should experimental participants derive the solution in an experiment?).

4. Alternating offer bargaining

The experimental results of ultimatum bargaining provoked a lively debate among game theorists whether orthodox game theory is just a normative, but unapplicable exercise. Like in one person-games the question is whether gamesmanship explains (at least experienced) experimental behavior or whether orthodox game theory has to be supplemented by a behavioral theory of game playing which pays more attention to psychological ideas than to elegant axiomatic characterizations of rationality concepts like cardinal utilities and the equilibrium for noncooperative games.

One typical reaction to striking experimental results is to ask how the results would change when the experimental setup is more or less seriously varied. In case of ultimatum bargaining it has been argued that fairness may matter less when parties have more than just one negotiation round for reaching an agreement. The guiding game model for this line of experimental research is the model of alternating offer bargaining: In odd rounds t player 1 offers and player 2 responds, in even rounds t the roles are reversed. An agreement is achieved if an offer is accepted. Otherwise one proceeds to the next round (except in the last round when non-acceptance means conflict). The usual assumption is furthermore that delaying the agreement is costly; i.e. of a shrinking pie.

There are by now several experimental studies of alternating offer bargaining. Notice that without any cost of delaying an agreement nothing much would be changed since all what matters is who is the final proposer and therefore obtains (nearly) all of p, the so-called “pie”. The studies essentially vary the time preferences, e.g. in the form of equal or unequal discount factors, and the horizon (the maximal number of rounds which, of course, is finite although one study tries to create an illusion of an infinite horizon). Another study assumes that every periodic proposer can declare his offer to be an ultimatum and that the “pie” p is increasing instead of decreasing. Like in the centipede experiment participants here can both gain by trusting (not terminating early), what happens in nearly half of the plays although this could be exploited. The explanation of the “centipede results” by (expected) altruism (in the tradition of reputation equilibria) cannot account for the increasing pie-results.

Whereas all these models rely on asymmetric bargaining rules among the symmetric bargaining models the so-called (Nash)-demand game (all parties simultaneously choose their demand which is what they obtain whenever the vector of demands is feasible, otherwise they get their conflict payoffs) received the most attention. An interesting study applies the binary lottery technique when studying demand bargaining (where one, however, allows for several rounds of simultaneous demands). Parties can earn an individual positive monetary prize and bargain just about the probability of winning their prize (with complementary probability the other party wins its prize). What is systematically varied is the information about the other’s prize. When prizes are private information, parties usually agree on equal winning probabilities. If both prizes are commonly known, they choose winning probabilities equating their monetary expectations what, of course, violates independence of affine utility transformations.

Due to the usually large number of strict equilibria (all efficient vectors of demands exceeding conflict payoffs) participants, playing the demand game, face an additional coordination problem what might justify to introduce preplay communication. The main findings are that the (Nash)-bargaining solution maximizing the product of agreement dividends must be focal for being selected and that, although the (Nash)-axioms are normatively convincing, they are behaviorally questionable. Experimentally the monotonicity axiom is better supported. Other topics in experimental non-cooperative games, often related to strategic market games, can be found in handbooks and introductions to experimental economics.

5. Characteristic function experiments

Like orthodox game theory also experimental game theory has been initially dominated by characteristic function models. Of course, a priori it is not clear at all how to implement a given characteristic function as an experiment (after all it is no strategic game). The usual procedure is to allow for free face to face-communication and to let coalitions announce payoff agreements which become binding if no coalition member quits it in so and so many minutes.

Given the usual heterogeneity of individual behavior value concepts were little used (although they may become important in accounting experiments of which the early reward or cost allocation experiments may be early, but (too) simple precedents). Thus the well-known set solutions like the core, the internally and externally stable sets, or the various bargaining sets were mainly tested or new related concepts developed.

Robust effects are that

- players in the same coalition obey the power structure (by granting to a more powerful coalition partner at least as much as to a less powerful one),
- equal payoff distributions are often proposed and often used as counterproposals when trying to argue against a previous proposal,
- smaller coalitions than the grand coalition are formed even when this is inefficient.

Characteristic function experiments were not only performed by game theorists, but also by (social) psychologists. A typical situation is to rely on majority voting games $(w_1,...,w_n;m)$ where w_i with $0 \le w_i \le 1$ and $w_1 + ... + w_n = 1$ denotes the voting share of player $i = 1,...,n$ and m with $1/2 \le m \le 1$, mostly $m = .5$, the majority level which a winning coalition S with $\sum_{i \in S} w_i > m$ must obtain. The characteristic function $v(\cdot)$ allowing for side payments assumes $v(S) = 1$ if S is winning and $v(S) = 0$ otherwise. In case of $n = 3$, $m = .5$ and $w_1 = .49$, $w_2 = .39$, $w_3 = .12$ one has $v(\{i\}) = 0$ for $i = 1, 2, 3$, and $v(S) = 1$ for any coalition S with at least two members. Thus the power structure, as reflected by the winning coalitions, is completely symmetric in spite of the large differences in voting shares. In such a situation experimentally observed payoff distributions are often influenced by both, the power structure and the voting shares.

Due to predominance of strategic game models in the industrial organization literature characteristic function experiments became a less popular research topic. Since strategic models seem to account for every possible result (see the discussion of "repairs" above) without serious restrictions on what to assume and what not there may, however, come a revival of characteristic function experiments for special situations where cooperative solutions are informative, e.g. in the sense of a small, but non-empty core. The main advantage would be that such informative solutions do not depend on subtle strategic aspects which often are behaviorally irrelevant although crucial for the non-cooperative solution. An example is the sequential timing of moves in ultimatum bargaining whose characteristic function is, however, symmetric in the sense of $v(\{i\}) = 0$ for both players i and $v(N) = p$ for the grand coalition N consisting of both players.

6. Quo vadis, experimental game theory?

Even when one accepts that players decide and behave otherwise than game theoretically predicted, one still learns a lot from reasonable repairs as long as they are based on obvious immaterial motives and emotional aspects. When a situation is rather simple even a boundedly rational participant can easily understand it. In such situations the rational choice approach in explaining unexpected game playing results will often reflect how participants derive their decisions. Its tradition of enriching game models and testing them experimentally will thus be continued.

On the other hand more and more ideas of other social sciences, especially of cognitive and social psychology, will be imported into behavioral game theory since they are often mentioned (in think aloud-studies or in experiments with team players) by participants when asked to explain their deliberation process. This line of research in experimental game theory will still borrow its terminology, its paradigms and its (normative) benchmarks from game theory but it will otherwise rely on fundamentally different cognitive deliberations of players: Players will rather satisfice than optimize, be obliged to social norms and conventions rather than opportunistically exploiting etc.

Experimental results suggesting that observed behavior reacts differently to subtle strategic aspects than game theoretically predicted could revive the former tradition of characteristic function experiments and of explanations, rooted in cooperative game theory. Along these lines the equal split in ultimatum bargaining would result from the fact that this is the Shapley-value, the kernel or the nucleolus of its characteristic function.

References

Allais, M. (1953). Fondements d'une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de L'Ecole Americaine, *Econométrie* **40**, 257-332. [This is an influential article challenging the validity of the theory of expected utility based on questionaire results.]

Anderhub, V. and W. Güth (1999). On intertemporal allocation behavior – A selective survey of saving experiments, *ifo-Studien* **3**, 303-333. [This article reviews the few experimental studies of intertemporal allocation, a problem typically discussed as saving behavior.]

Anderhub, V., D. Engelmann, and W. Güth (1999). An experimental study of the repeated trust game with incomplete information, *SFB Discussion Paper* No. 97/1999, Humboldt-University of Berlin. [Reputation equilibria are tested by experimentally controlling incomplete information, allows for mixed strategies so that theory can also be tested on the level of an individual.]

Anderhub, V., W. Güth, U. Kamecke, H.-T. Normann (2000). Capacity choices and price competition in experimental markets, *Working Paper*, Humboldt-University Berlin. [After experiencing price competition for given capacities participants repeatedly choose capacities which remain constant for 10 market periods.]

Andreoni, J. (1988). Privately provided public goods in a large economy: The limits of altruism, *Journal of Public Economics* **35**(1), 57-73. [Discusses altruism as an explanation of behavior in public goods games.]

Aumann, R. (1985). Repeated games, in: *Issues in Contemporary Microeconomics and Welfare*, G. R Feiwel (ed.), 209-242, Albany: State University of New York Press. [An influential discussion of what is usually discussed as Folk Theorems showing that in infinite horizon games many – noncooperative and cooperative – forms of behavior are possible.]

Aumann, R. (1995). Backward induction and common knowledge of rationality, *Games and Economic Behavior* **8**, 6-19. [Argues that common knowledge of rationality is not self-contradictory if properly interpreted.]

Aumann, R. J. and M. Maschler (1964). The bargaining set for cooperative games, in: *Advances in game theory*, M. Dresher, L. S. Shapley, A. W. Tucker (eds.), Princeton, N. J.: Princeton University Press. [This cooperative solution concept avoids the emptyness of the core by allowing for counterobjections to objections, i. e. not every dominance relation between payoff vectors is accepted.]

Axelrod, R. (1984). *The evolution of cooperation*, New York: Basic Books, Inc., Publishers. [Influential article leading to the literature on tournaments where all strategies are matched and can be ranked according to their success.]

Becker, G. S., M. H. de Groot, and J. Marshak (1963). An experimental study of some stochastic models for wagers, *Behavioral Science* **8**, 199-202. [This introduces the random price mechanism according to which it is the only dominant strategy to reveal one's value.]

Becker, G. S. and K. M. Murphy (1988). A theory of rational addiction, *Journal of Political Economy* **96**, 675-700. [Drug consumption is determined by intertemporal optimization assuming that all future consequences can be foreseen. This is no joke!]

Bicchieri, C. (1989). Self-refuting theories of strategic interaction: A paradox of common knowledge, *Erkenntnis* **30**, 69-85. [Tries to show that common knowledge of rationality is self-contradictory.]

Binmore, K., A. Shaked, and J. Sutton (1985). Testing noncooperative bargaining theory: A preliminary study, *American Economic Review* **75**, 1178-1180. [This experimental study claims that fairness matters more for inexperienced than for experienced participants.]

Bohm, P., J. Lindén, and J. Sonnegard (1997). Eliciting reservation prices: Becker-De Groot-Marschak mechanisms vs. markets, *Economic Journal* **107** (443), 1079-1089. [This experimental study questions that random price mechanisms actually lead to truthful bids.]

Bolton, G. (1991). A comparative model of bargaining: Theory and evidence, *American Economic Review* **81**, 1096-1136. [An early influential study incorporating a taste for fairness when modelling the preferences of bargainers.]

Bolton, G. and A. Ockenfels (2000). ERC - A theory of equity, reciprocity and competition, *American Economic Review*, in press. [By including inequality aversion many experimental results can be explained even when rationality is maintained.]

Bush, E. and F. Mosteller (1955). *Stochastic models for learning*, New York: Wiley. [Influential book introducing reinforcement learning, also referred to as stimulus-response learning.]

Camerer, C. (1995). Individual decision making, in: *Handbook of Experimental Economics*, J. H. Kagel and A. E. Roth (eds.), 587-703, Princeton, N.J.: Princeton University Press. [Surveys experimental studies of individual decision making under risk.]

Cournot, A. A. (1838). *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, Paris: Hachette; English translation: Researches into the mathematical principles of the theory of wealth, New York: Macmillan, 1897. [Very influential study in oligopoly theory whose solution idea later on led to the concept of strategy equilibria and whose dynamics became later on known as best reply dynamics.]

Chamberlin, E. H. (1948). An experimental imperfect market, *Journal of Political Economy* **56**, 95-108. [Seen as the first study in experimental economics which tests the competitive (market) equilibrium.]

Davis, D. D. and C. A. Holt (1993). *Experimental Economics*, Princeton, N. J.: Princeton University Press. [Introduction to Experimental Economics providing examples and discussing experimental results.]

Edgeworth, F. Y. (1881). *Mathematical psychics, An essay on the applicationof mathematics to the moral sciences*, London: Kegan. [An impressive book which became famous for its sophisticated analysis of (replica) exchange economies which already used the concept of a core and proved an early limit theorem saying that the core shrinks to the set of competitive equilibria when the number of traders goes to infinity.]

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition and cooperation, *Quarterly Journal of Economics* **CXIV**, 817-868. [By including inequality aversion many experimental results can be explained even when rationality is maintained.]

Festinger, L. (1957). *A theory of cognitive dissonance,* Stanford: Stanfort University Press. [This book introduces an influential psychological concept claiming that one avoids cognitive dissonance often by reevaluating concerns and evidence or by biased information acquisition.]

Fischbacher, U. (1999). Z-Tree: Zurich Toolbox for Readymade Economic Experiments, *Working paper* No. 21, Institute for Empirical Research in Economics, University of Zurich. [Provides an easily applied software for programming the software for experiments with strategic interaction.]

Fouraker, L. E. and S. Siegel (1963). *Bargaining behavior*, New York: McGraw Hill. [An early and influential book on experimental studies of bargaining behavior, including market behavior.]

Friedman, D. and S. Sunder (1994). *Experimental methods – A primer for economists,* Cambridge University Press. [Introduction to experimental economics]

Gale, J., K. G. Binmore, and L. Samuelson (1995). Learning to be imperfect: The ultimatum game, *Games and Economic Behavior* **8**(1), 56-90. [This study tries to justify the typical experimental findings of ultimatum experiments as stable results of adaptive processes.]

Geanakoplos, J., D. Pearce, and E. Stacchetti (1989). Psychological games and sequential rationality, *Games and Economic Behavior* **1**(1), 60-79. [The so-called psychological idea is that players do not want to disappoint the expectations of others which, via consistency, have to be in line with actual behavior; is void of (social) psychology.]

Güth, W., R. Schmittberger, and B. Schwarze (1982). An experimental analysis of ultimatum bargaining, *Journal of Economic Behavior and Organization* **3**(4), 367 - 388. [Introduces what later on was named the ultimatum game (experiment), led to many subsequent studies, essentially all confirming the first basic results.]

Güth,W., P. Ockenfels, and M. Wendel (1993). Efficiency by trust in fairness? – Multiperiod ultimatum bargaining experiments with an increasing cake, *International Journal of Game Theory* **22**, 51-73. [In this experimental study of alternating offer bargaining the pie increases over time and proposers command always ultimatum power which, however, is not used in early periods.]

Güth, W., R. Ivanova-Stenzel, M. Königstein, and M. Strobel (1999). Learning to bid - An experimental study of bid function adjustments in auctions and fair division games, *SFB Discussion Paper* 70, Humboldt-University Berlin. [Participants play four different games repeatedly, namely 1^{st} and 2^{nd}-price auctions and fair division games; data reveal anticipation and deliberation as well as path dependence.]

Hammerstein, P. and R. Selten (1994). Game theory and evolutionary biology, in: *Handbook of Game Theory*, Volume 2, Chapter 28, R. J. Aumann and S. Hart (eds.), 928-993, Amsterdam, Netherlands: Elsevier Science B.V. [Surveys the concepts of evolutionary stability in biology and game theory.]

Harrison, G. W. (1994). Expected utility theory and the experimentalists, *Empirical Economics* **19**, 223-253. [Claims that most deviations from expected utility maximization are non-salient.]

Henning-Schmidt, H. (1999). *Bargaining in a video experiment – Determinants of Boundedly Rational Behavior*, Lecture Notes in Economics and Mathematical Systems 467, Berlin, Heidelberg, New York: Springer-Verlag. [Reports on video-taped bargaining experiments.]

Hey, J. D. (1991). *Experiments in Economics*, Cambridge, Mass.: Blackwell. [Introduction to experimental economics.]

Huck, S. and J. Oechssler (1999). The indirect evolutionary approach to explaining fair allocations, *Games and Economic Behavior* **28**, 13-24. [This study tries to justify the typical experimental findings of ultimatum experiments as stable results of adaptive processes.]

Huck, S., Normann, H.-T. and Oechssler, J. (1999). Learning in Cournot oligopoly - An experiment, *Economic Journal* **109**, C80-C95. [An experimental study of oligopoly markets with varying degrees of information.]

Hoggatt, A. C. (1959). An experimental business game, *Behavioral Science* **4**, 192-203. [An early but (at hindsight) less influential study in experimental economics.]

Kagel, J. H. and A. E. Roth (1995). *Handbook of Experimental Economics*, Princeton, N.J.: Princeton University Press. [Contains many survey articles on various topics in experimental economics as well as an account of how experimental economics developed.]

Kalai, E. and M. Smordinski (1975). Other solutions to Nash's bargaining problem, *Econometrica* **43**, 513-518. [An axiomatic bargaining theory, based on an intuitive monotonicity axiom rather than on independence of irrelevant alternatives, which allows for a simple geometric derivation.]

Kalish, G., J. W. Milnor, J. F. Nash, and E. D. Nering (1954). Some experimental n-person games, in: *Decision Processes*, R. M. Thrall, C. H. Coombs and R. L. Davis (eds.), 301-328, New York: Wiley. [An early experimental study of game playing behavior.]

Kirchsteiger, G. (1994). The role of envy in ultimatum games, *Journal of Economic Behavior and Organization* **25**(3), 373-389. [Explains ultimatum bargaining results by envy.]

Komorita, S. S. and J. M. Chertkoff. A bargaining theory of coalition formation, *Psychological Review* **80**, 149-162. [An experimental study of weighted majority games.]

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982). Rational cooperation in the finitely repeated prisoners' dilemma, *Journal of Economic Theory* **27**, 245-252. [Introduces the reputation equilibrium by claiming incomplete information even if it is not experimentally controlled.]

Maynard Smith, J. and G. R. Price (1973). The logic of animal conflict, *Nature* **246**, 15-18. [Influential article showing the similarity of evolutionary stability and the (symmetric) equilibrium concept in game theory.]

McKelvey, R. D. and T. Palfrey (1992). An experimental study of the centipede game, *Econometrica* **60**, 803-836. [Experiment of the centipede game, explains initial cooperation via reputation formation.]

Modigliani, F. and R. Brumberg (1954). Utility analysis and the consumption function: An interpretation of cross-section data, in: *Post-Keysian Economics*, K. K. Kurihara (ed.), New Brunswick, N. J.: Rudgers University. [Related to intertemporal allocation and the life cycle-hypothesis.]

Nash, J. F. (1953). Two-person cooperative games, *Econometrica* **21**, 128-140. [Presents the equilibrium concept and proves its existence.]

Nydegger, R. V. and G. Owen (1975). Two person bargaining: An experimental test of the Nash axioms, *International Journal of Game Theory* **3**, 239-349. [In this experiment two players can bargain how to allocate two stacks of chips with different colors and different values for each; all pairs distributed such that both received the same monetary payoff.]

Peleg, B. and S. Tijs (1996). The consistency principle for games in strategic forms, *International Journal of Game Theory* **25**(1), 13-34. [Provides a new axiomatic justification of the equilibrium by (converse) consistency.]

Pruitt, D. G. (1967). Reward structure and cooperation – The decomposed prisoner's dilemma game, *Journal of Personality and Social Psychology* **7**, 21-27. [The experimental treatments are various decompositions of the same prisoners' dilemma game which, however, yield widely differing degrees of cooperation.]

Rabin, M. (1993). Incorporating fairness into game theory and economics, *American Economic Review* **83**(5), 1281-1302. [Applies psychological game theory to explain experimental results.]

Rapoport, A. and A. M. Chammah (1965). *Prisoner's dilemma: A study in conflict and cooperation*, Ann Abor: University of Michigan Press. [An influential book leading to many subsequent studies of prisoners' dilemma games.]

Roth, A. E. (1995a). Introduction to experimental economics, in: *Handbook of Experimental Economics*, J. H. Kagel and A. E. Roth (eds.), 3-109, Princeton, N.J.: Princeton University Press. [Reviews the early development of experimental economics.]

Roth, A. E. (1995b). Bargaining experiments, in: *Handbook of Experimental Economics*, J. H. Kagel and A. E. Roth (eds.), 253-348, Princeton, N.J.: Princeton University Press. [Reviews bargaining experiments.]

Roth, A. E. and M. W. K. Malouf (1979). Game theoretic models and the role of information in bargaining, *Psychological Review* **86**, 574-594. [In this experiment two parties bargain how the unit probability for receiving a high versus a low premium should be allocated among them, varies information about other's premia, an early application of the binary lottery technique.]

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model, *Econometrica* **50**, 97-110. [An elegant alternating offer bargaining model for a shrinking pie which avoids multiple (perfect) equilibria.]

Rubinstein, A. (1991). Comments on the interpretation of game theory, *Econometrica* **59**, 909-924. [A claim that game theory's assumptions and implications can be interpreted rather freely.]

Sauermann, H. (1978a*). Bargaining Behavior, Contributions to experimental economics*, volume 7, Tübingen: Mohr. [A collection of experimental studies of bargaining.]

Sauermann, H. (1978b). *Coalition former behavior, Contributions to experimental economics*, volume 8, Tübingen: Mohr. [A collection of experimental studies of bargaining and coalition formation.]

Sauermann, H. and R. Selten (1959). Ein Oligopolexperiment, *Zeitschrift für die gesamte Staatswissenschaft* **155**, 427-471; Translation: An Experiment in Oligopoly, in: *General Systems Yearbook of the Society for General Systems Research* 5, L. von Bartalanffy and A. Rapoport (eds.), Michigan 1960, 85-114. [An early study in experimental economics.]

Schmeidler, D. (1969). The nucleolus of a characteristic function game, *SIAM Journal of Applied Mathematics* **17**, 1163-1170. [Introduces the nucleolus and proves its uniqueness.]

Shapiro, E. G. (1975). Effects of future interaction in reward allocation in dyads: Equity or equality, *Journal of Personality and Social Psychology* **31**, 873-880. [Participants first work and get a joint reward which is distributed by a dictator, varies the expectation of future interaction.]

Shapley, L. S. (1953). A value for n-person games, in: Contributions to the theory of games II, H. W. Kuhn and A. W. Tucker (eds.), *Annals of Mathematical Studies Series* **28**, 307-317, Princeton, N.J.: Princeton University Press. [Introduces the Shapley-value both, by an axiomatic characterization and by an algorythmic procedure for forming the grand coalition.]

Selten, R. and R. Stöcker (1986). End behavior in sequences of finite prisoner's dilemma supergames: A learning theory approach, *Journal of Economic Behavior and Organization* **7**(1), 47-70. [An experimental study of repeating repeated games.]

Selten, R. and G. R. Uhlich (1988). Order of strength and exhaustivity as additional hypotheses in theories for 3-person characteristic function games, in: Bounded rational behavior in experimental games and markets: Proceedings of the Fourth Conference on Experimental Economics (Bielefeld, West Germany, September 21-25, 1986), R. Tietz, W. Albers, R. Selten (eds.), *Lecture Notes in Economics and Mathematical Systems* **314**, Berlin: Springer-Verlag, 235-250. [A behavioral theory for characteristic function experiments explaining many results.]

Taylor, P. and L. Jonker (1978). Evolutionary stable strategies and game dynamics, *Mathematical Biosciences* **40**, 145-156. [Studies adaptive dynamics which are known as replicator dynamics.]

Tirole, J. (1988). *The theory of industrial organization*, Cambridge, Mass.: M.I.T. Press. [An introduction to what has been called oligopoly theory and is now named industrial economics.]

von Neumann, J. and O. Morgenstern (1944). *Theory of games and economic behavior*, Princeton, NJ.: Princeton University Press. [The book which has established game theory as a field of active research.]

Weg, E, A. Rapoport, and D. S. Felsenthal (1990). Two-person bargaining behavior in fixed discounting factors games with infinite horizon, *Games and Economic Behavior* **2**, 76-95. [An experimental study of alternating offer bargaining with a decreasing pie and no definite time horizon.]

Weibull, J. W. (1995). *Evolutionary Game Theory*, Cambridge and London: M.I.T. Press. [Surveys the theoretical results for various static and dynamic concepts of evolutionary stability.]