IN WHAT SENSE IS THE NASH SOLUTION FAIR?

WALTER TROCKEL*
Universität Bielefeld

1. Introduction

An abstract two-person bargaining problem is a pair (T, d) where $T \subset \mathbb{R}^2$ and $d \in T$ with the following properties:

- T closed, convex, comprehensive (i.e. $x \in T \Longrightarrow \{x\} - \mathbb{R}^2_+ \subset T$)
- $T \cap \mathbb{R}^2_{++} \neq \emptyset$
- $d \in int(T \cap \mathbb{R}^2_+)$.

The interpretation is that two players have to agree on a joint payoff vector in T, the ith coordinate for the ith player, $i = 1, 2$ in order to receive these payoffs or, else, to fall back to the status quo point d.

Assuming, that this scenario results as the image under the two players' concave von Neumann-Morgenstern utility functions on an underlying economic or social scenario, this model is determined only up to affine transformations of both players' payoffs. Accordingly, d is sometimes assumed to be $0 \in \mathbb{R}^2$ (0-normalization), sometimes in addition it is assumed that $max_{x \in T}\ x_i = 1, i = 1, 2$ ($0-1-1$-normalization). Moreover, every part of T not in $\mathbb{R}^2_+$ is skipped representing the fact that the interest focusses only on individually rational payoff vectors. The resulting $S \subset \mathbb{R}^2_+$ is then a $0 - 1 - 1$-normalized bargaining situation, whose boundary is often assumed to be smooth.

The Nash bargaining solution has been introduced by John F. Nash (1953) as a solution for two person bargaining games. Nash already presented three approaches to the solution that are methodologically and in spirit quite different. One is the definition of the Nash solution as the maximizer of the Nash product, i.e. the product of the two players' payoffs. This might be seen as maximizing some social planners' preference relation on the set of players' utility allocations. So whatever fairness is represented by the Nash solution it should be hidden in this planners' preferences.

---

* I am happy to be able to contribute with this article to the honoring of Christian Seidl, a highly esteemed colleague.

U. Schmidt and S. Traub (eds.), Advances in Public Economics: Utility, Choice and Welfare, 17-30.

The second approach of Nash is the one via axioms for the bargaining solution. This approach became quite popular later on in the literature on cooperative games and, in particular bargaining games. It turned out that several important alternative bargaining solutions, like the Nash, Kalai-Smorodinsky, Perles-Maschler or Raiffa solution, coincide on hyperplane bargaining games where they may be characterized by the three axioms of cardinal invariance, Pareto-efficiency and symmetry (or anonymity) but differ by specific fourth axioms on general bargaining games. Nash's fourth axiom, the *Independence of Irrelevant Alternatives (IIA)* may be replaced by *consistency* due to Lensberg (1988). So any fairness specific to the Nash solution might be hidden in these alternative axioms.

The third approach of Nash was via his simple demand game and built the first attempt in the Nash program. The Nash Program is a research agenda whose goal it is to provide a non-cooperative equilibrium foundation for axiomatically defined solutions of cooperative games. This program was initiated by John Nash in his seminal papers *Non-cooperative Games* in the *Annals of Mathematics*, 1951, and *Two-Person Cooperative Games* in *Econometrica*, 1953. The term *Nash Program* was introduced by Binmore (1987). The original passages due to Nash that built the basis for this terming are in fact quite short.

The Nash program tries to link two different ways of solving games. The first one is non-cooperative. No agreements on outcomes are enforceable. Hence players are totally dependent on their own strategic actions. They try to find out what is best given, the other players are rational and do the same. In this context the Nash equilibrium describes a stable strategy profile where nobody would have an interest to unilaterally deviate. Nevertheless there is an implicit institutional context. The strategy sets define implicitly what choices are not allowed, those outside the strategy sets. The payoff functions reflect which strategies in the interplay with others' strategies are better or worse. It is not said explicitly who grants payoffs and how the physical process of paying them out is organized. But there is some juridical context with some enforcement power taken for granted. There is no interpersonal comparison of payoffs involved in the determination of good strategies. Each player only compares his different strategies contingent on the other players' different strategy choices. As applications in oligopoly show, institutional restrictions of social or economic scenarios are mapped into strategy sets and payoff functions, thereby lending them an institutional interpretation. Yet, totally different scenarios may considerably be modelled by the same non-cooperative game, say in strategic form. This demonstrates clearly the purely payoff based evaluation of games. Payoffs usually are interpreted as reflecting monetary or utility payments. Associated physical states or allocations occur only in applications and may be different in distinct applications of the same game.

The second way to solve a game is the cooperative one via axioms as first advocated by Nash (1953). Again the legal framework is only implicit. Yet, now not only obedience to the rules is assumed to be enforceable but even contracts. Mutual gains are in reach now as it becomes possible by signing a contract to commit himself to certain behavior. In this context it is the specific payoff configuration which is of interest rather than the strategy profile that would generate it. In this framework it is reasonable, therefore, to neglect the strategic options and concentrate on the feasible payoff configurations or utility allocations on which the players possibly could agree by signing a contract. Again the formal model does not specify the process by which physical execution of a contract is performed. Again it is the payoff space rather than some underlying social scenario on which the interest rests except in applications of game theory.

In contrast to the non-cooperative approach now players *are* interested in what other players receive. Although utilities or payoff units for different players are in general not considered comparable typically there are tradeoffs that count. The axioms that are fundamental in this context reflect ideas of fairness, equity, justness that do not play a role in the non-cooperative model. But a process of negotiation with the goal to find an agreement makes it necessary for each player to somehow judge the coplayers' payoffs. But the axioms are in a purely welfaristic context. If very different underlying models lead to the same cooperative game in coalitional form it is only the solution in terms of payoff vectors that is relevant. And this determines in any application what underlying social or physical state is distinguished. It becomes irrelevant in the axiomatic cooperative approach which are the institutional details. Important are only the feasible utility allocations.

Now, why could it be interesting to have a non-cooperative strategic game and a cooperative game in coalitional form distinguishing via its equilibrium or solution, respectively, the same payoff vector? According to Nash the answer is that each approach *"helps to justify and clarify the other"*.

The equality of payoffs in both approaches seems to indicate that the institutional specifities represented by the strategic model are not so restrictive as to prevent the cooperative solution. Also the payoff function appears then to reflect in an adequate way the different axioms. On the other hand payoff combinations not adequate under the solution concept cannot be strategically stable. So the equivalence of both approaches seems to indicate that the strategic model from the point of view of social desirability is restrictive enough but not too restrictive. This abstract relation has different consequences if one is in one of the two different enforceability contexts. If we cannot enforce contracts the equivalence of two approaches means that this is not a real drawback, as we can reach the same via rational strategic interaction (at least in situations of games with a unique equilibrium). If, on the other hand, we are in a world where contracts are enforceable, we may use the equivalence of a suitable strategic approach as additional arguments for the payoff vectors distinguished by the solution. Therefore, results in the Nash program give players valuable insights into the interrelation between institutionally determined non-cooperative strategic interaction and social desirability based on welfaristic evaluation.

There is not, however, any focus on *decentralization* in the context of the Nash program simply because there is no entity like a center or planner. There are just players.

Nash's own first contribution to the Nash Program (1953) consists in his analysis of a game, the *demand game* and the so called *smoothed demand game* where he looked at the limiting behavior of non-cooperative equilibria of a sequence of smoothed versions of the demand game. Here the amount of smoothing approaches zero, and, hence the sequence approximates the demand game. While the original “simple” demand game has a continuum of equilibria, a fact which makes it useless for a non-cooperative foundation of the Nash solution, Nash argued that the Nash solution was the only necessary limit of equilibria of the smoothed games. Rigorous analyzes for his procedure have been provided much later by Binmore (1987), van Damme (1986) and Osborne and Rubinstein (1990).

A second quite different approximate non-cooperative support for the Nash solution is provided by Rubinstein’s (1982) model of sequential alternate offers bargaining. Binmore, Rubinstein and Wolinsky (1986) showed in two different models with discounted time that the weaker the discounting is the more closely approximates the subgame perfect Nash equilibrium an asymmetric Nash bargaining solution. Only if subjective probabilities of breakdown of negotiations or the lengths of reaction times to the opponents’ proposals are symmetric it is the symmetric Nash solution which is approximately supported. Again, in the frictionless limit model one does not get support of the Nash solution by a unique equilibrium. Rather every individually rational payoff vector corresponds to some subgame perfect equilibrium.

An exact support rather than only an approximate one of the Nash solution is due to Howard (1992). He proposes a fairly complex 10 stages extensive form game whose unique subgame perfect equilibrium payoff vector coincides with the bargaining solution.

Like in Rubinstein’s model and in contrast to Nash framework Howard’s game is based on underlying outcome space. Here this is a set of lotteries over some finite set on which players have utility functions. Although the analysis of the game can be performed without explicit consideration of the outcome space it is this underlying structure that allows it to look at the outcome associated with a subgame perfect equilibrium and thereby interpret Howard’s support result as a mechanism theoretic implementation of some Nash social choice rule in subgame perfect equilibrium. Whatever non-cooperative support for the Nash solution we take, according to Nash himself it should provide to our understanding of the Nash solution, and, so we may hope, of its inherent fairness.

In what follows I shall try to relate Nash’s three approaches to inherent fairness properties of the Nash solution. I will start with the axiomatic approach, continue with a related market approach and will derive from the latter one a further non-cooperative foundation, that allows a conclusion as to specific fairness. In the last part I shall discuss the fairness hidden in the Nash product.

2. The Axiomatic Approach

Nash’s axiom IIA asserts that if one bargaining problem is contained in another one and contains the other one’s solution as a feasible point its own solution should coincide with that point. The IIA is formally closely related to rationality axioms like the weak or strong axiom of revealed preferences. As such it does not hint to any underlying fairness concept. One may however weaken IIA in such a way that together with the other axioms it still characterizes the Nash solution. This is done by restricting in the IIA the larger bargaining problem to be always a hyperplane game, whose boundary, the intersection of $\mathbb{R}^2_+$ with a hyperplane, is tangent to the boundary of the game it contains. In such hyperplane games all bargaining solutions pick the barycenter, i.e. every player gets the same share of his utopia point or, put differently, makes the same concession measured in his specific personal utility units. This also represents a solution of the smaller NTU-game without making use of transfers offered by the containing hyperplane game. So in this weakened version IIA has the spirit of a no-trade equilibrium in general equilibrium theory. Once a Walrasian relation is considered possible one finds immediately that Lensberg's consistency, the alternative to IIA, that characterizes the Nash solution is formally almost identical to the consistency of Walrasian equilibrium (cf. Young, 1994, p.153). The inherent fairness of the Walrasian equilibrium is known to go beyond its Pareto efficiency guaranteed by the First Welfare Theorem. It is represented by the Equivalence Principle, a group of results assuring the near or exact equality of Walrasian allocations and those allocations determined by various game theoretical solutions in large pure exchange economies. The most famous equivalence results, those for the Shapley value, the Core and the Mas-Colell Bargaining set, guarantee that in large competitive environments any kind of strategic arbitrage is prevented by the power of perfect competition. True, this context of pure exchange economies is totally different from our purely welfaristic bargaining situations. Nevertheless, there are more indications in the axiomatic approach that underlying fairness of the Nash solution is a "Walrasian" one.

Shapley (1969) showed that the simultaneous requirements of efficiency (maximal sum of utilities) and equity (equal amounts of utility) that are in general incompatible become compatible for a suitable affine transformation of the original bargaining situation. The preimage under this affine transformation of the efficient and equitable utility allocation in the transformed problem turns out to be the Nash solution of the original problem. For the status quo point being zero the affine transformation becomes linear and is uniquely described by the normal vector λ at the Nash solution. This λ, that may be interpreted as an efficiency price system, defines endogenously local rates of utility transfer. Shubik (1985) speculates that this λ reminds very much of a competitive price system. In fact, this conjecture has been proved in Trockel (1996), where the bargaining problem has been interpreted as an artificial Arrow-Debreu economy, whose unique Walrasian equilibrium allocation coincides with the Nash solution, while the normal vector λ is an equilibrium price system.

So the fairness of the Nash solution seems to be the immunity against undue exploitation by the opponent as guaranteed by perfect competition.

Interestingly enough, a similar message can be read off Rubinstein's approximate foundation of the Nash solution in his alternating offer game. The approximation is the better the less Rubinstein's cake shrinks when time passes. That means almost no shrinking creates arbitrary many future alternative options for finding an adequate bargaining outcome. These future alternative options correspond to "the many outside options" represented in a stylized way by the concept of a Walrasian equilibrium. That the equivalence principle holds also for our special construct of a bargaining economy is shown in the next section.

3. An Edgeworth-Debreu-Scarf Type Characterization of the Nash Solution

In the present section that is based on Trockel (2005) we relate the Nash solution with the Edgeworthian rather than the Walrasian version of perfect competition. To do so, we define an artificial coalition production economy (cf. Hildenbrand, 1974) representing a two person bargaining game.

In a similar way the Nash solution has been applied in Mayberry et al. (1953) to define a specific solution for a duopoly situation and comparing it with other solutions, among them the Edgeworth contract curve. The relation between these two solutions will be the object of our investigation in this paper.

Though it would not be necessary to be so restrictive we define a two person bargaining game as the closed subgraph of a continuously differentiable strictly decreasing concave function $f : [0,1] \longrightarrow [0,1]$ with $f(0) = 1$ and $f(1) = 0$.

$$S := subgraph f := \{(x_1, x_2) \in [0,1]^2 | x_2 \leq f(x_1)\}$$

The normalization reflects the fact that bargaining games are usually considered to be given only up to positive affine transformations. Smoothness makes life easier by admitting unique tangents.

The model S is general enough for our purpose of representation by a coalition production economy. In particular, S is the intersection of some strictly convex comprehensive set with the positive orthant of $\mathbb{R}^2$.

Define for any S as described above a two person coalition production economy $\mathcal{E}^S$ as follows:

$$\mathcal{E}^S := ((e_i, \succsim_i, Y_i)_{i=1,2}, (\vartheta_{ij})_{i,j=1,2})$$

such that

$$e_i = (0,0), x = (x_1, x_2) \succsim_i x' = (x'_1, x'_2) \Leftrightarrow x_i \geq x'_i, i = 1,2\ ,$$
$$\vartheta_{11} = \vartheta_{22} = 1, \vartheta_{12} = \vartheta_{21} = 0, Y_1 = Y_2 = (\frac{1}{2})S\ .$$

The zero initial endowments reflect the idea that all available income in this economy comes from shares in production profits. Each agent owns fully a production possibility set that is able to produce for any $x \in S$ the bundle $(\frac{1}{2})x$ without any input. Both agents are interested in only one of the two goods called "agent $i's$ utility", $i = 1,2$. Without any exchange agent i would maximize his preference by producing and consuming one half unit of commodity i and zero units of commodity $3-i, i = 1,2$. However, the agents would recognize immediately that they left some joint utility unused on the table.

Given exchange possibilities for the two commodities they would see that improvement would require exchange or, to put it differently, coordinated production

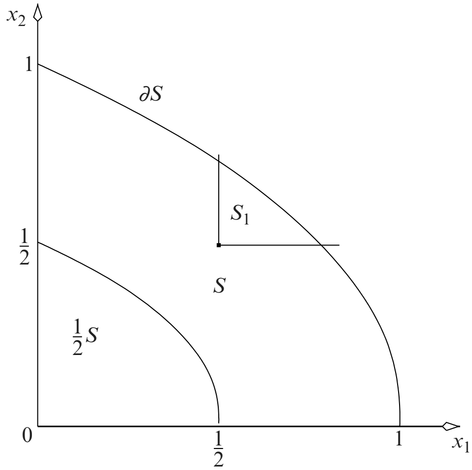

Figure 1.

(see Figure 1). The point $(\frac{1}{2}, \frac{1}{2})$ corresponds to the vector of initial endowments, the set $S_1 := S \cap (\{(\frac{1}{2}, \frac{1}{2})\} + \mathbb{R}^2_+)$ to the famous lens and the intersection of S_1 with the efficient boundary of S, i.e. $S_1 \cap \partial S$, to the core in the Edgeworth Box. This is exactly what Mayberry and al. (1953, p. 144) call the *Edgeworth contract curve* in their similar setting.

The according notions of improvement and of the core are analogous to the ones used for *Coalitional Production Economies* by Hildenbrand (1974, p. 211). $\tilde{Y} : \{\{1\},\{2\},\{1,2\}\} \Longrightarrow \mathbb{R}^2$ with $\tilde{Y}(\{1\}) = Y_1, \tilde{Y}(\{2\}) = Y_2, \tilde{Y}(\{1,2\}) = S$, is the production correspondence, which is additive, as $Y(\{1\} \cup \{2\}) = Y_1 + Y_2 = S$. An allocation $x^i = ((x^i_1, x^i_2))_{i=1,2}$ for $\mathcal{E}^S$ is T-*attainable* for $T \in \{\{1\},\{2\},\{1,2\}\}$ if $\sum_{i \in T} x^i \in \tilde{Y}(T)$; it is called *attainable* if it is $\{1,2\}$-attainable.

An allocation (x^1, x^2) can be *improved upon* by a coalition $T \in \{\{1\},\{2\},\{1,2\}\}$ if there is a T-*attainable* allocation (y^1, y^2) such that $\forall i \in T : y^i \succ_i x^i$. The *core* of $\mathcal{E}^S$ is the set of $\{1,2\}$-attainable allocations that cannot be improved upon. The analogous definitions hold for all n-replicas $\mathcal{E}^S_n$ of $\mathcal{E}^S$, $n \in \mathbb{N}$.

Notice that our choice of $Y_i = (\frac{1}{2})S, i = 1,2$ ensures the utility allocation $(\frac{1}{2}, \frac{1}{2})$ for the two players in case of non-agreement. This differs from Nash's status quo or threat point $(0,0)$.

Formalizing an n-replica economy $\mathcal{E}^S_n$ is standard. All characteristics are replaced by n-tupels of identical copies of these characteristics. In particular $\mathcal{E}^S_n$ has $2n$ agents, n of each of the two types 1 and 2. And the total production possibility set for the grand coalition of all $2n$ agents is nS.

Although the use of strict convex preferences as in Debreu and Scarf (1963) is not available here a short moment of reflection shows that a major part of their arguments can be used in our case as well.

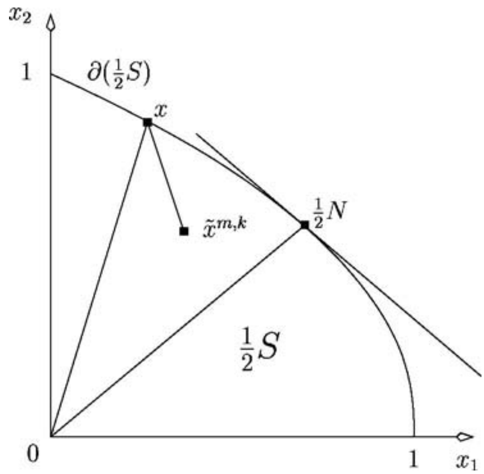

Figure 2.

Next we are looking at the core of n-replicas $\mathcal{E}_n^S$ of the economy $\mathcal{E}^S$. It suffices to look at S. Notice that it does not make any difference whether in an n-replica economy every agent has the technology $Y = \frac{1}{2n}S$ and the total production set is S or wether each agent has $Y = (\frac{1}{2})S$ and total production is nS. We will assume that each agent in $\mathcal{E}_n^S$ owns a production possibility set $Y := \frac{1}{2}S$ as illustrated in Figure 2.

We assume w.l.o.g. that $x \in \partial(\frac{1}{2}S)$ and $x_1 < \frac{1}{2}N_1, x_2 > \frac{1}{2}N_2$. By choosing $n, m, k \in \mathbb{N}, k < m \leq n$ sufficiently large we can make the vector $\frac{m-k}{m+k}(x_1, -x_2)$ arbitrarily small and, thereby, position the point $\tilde{x}^{m,k} := (x_1, x_2) + \frac{m-k}{m+k}(x_1, -x_2)$ in $int(\frac{1}{2}S)$.

A coalition C_n^x in the n-replica economy $\mathcal{E}_n^S$ of $\mathcal{E}^S$ consisting of m agents of type 1 and k agents of type 2 can realize the allocation $(m+k)\tilde{x}^{m,k} = ((m+k)x_1 + (m-k)x_1, (m+k)x_2 - (m-k)x_2) = (2mx_1, 2kx_2)$.

This bundle can be reallocated to the members of C_n^x by giving to each of the m type 1 agents $(2x_1, 0)$ and to each of the k type 2 agents $(0, 2x_2)$. Clearly, everybody gets thereby the same as he received in the beginning when everybody produced x. Therefore, nobody improves! However, for $\eta > 0$ sufficiently small $\tilde{x}^{m,k} \in int\ \frac{1}{2}S$ implies that $\tilde{x}^{m,k} + \eta N \in int \frac{1}{2}S$. Now reallocation of that bundle among the members of C_n^x can be performed in such a way that each type 1 agent receives $(2x_1 + \frac{m+k}{m}\eta N_1, 0)$ and each type 2 agent gets $(0, 2x_2 + \frac{m+k}{k}\eta N_2)$. Therefore x for every agent can be improved upon by C_n^x via production of $\tilde{x}^{m,k} + \eta N$ by each of its members. Again, the only element of $\partial(\frac{1}{2}S)$ remaining in the core for all $n-$replications of $\mathcal{E}^S$ is the point $\frac{1}{2}N$, i.e. the Nash solution for $\frac{1}{2}S$.

Notice that any point $y \in \partial(\frac{1}{2}S)$ with $y_1 < x_1 < N_1$ can be improved upon by the

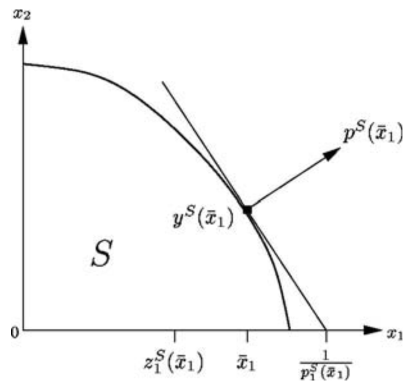

Figure 3.

same coalition C_n^x via $\tilde{y}^{m,n} + \eta N$ with the same η by a totally identical construction of $\tilde{y}^{m,n}$ from y. The same is not true for $z \in \partial(\frac{1}{2}S)$ with $x_1 < z_1 < N_1$. Here the $\frac{m-k}{m+k}(z_1, -z_2)$ may require a larger m and k to make $\frac{m-k}{m+k}(z_1, -z_2)$ small enough. We may for any $x \in \partial(\frac{1}{2}S), x_1 < N_1$ choose the m, k in the construction of $\tilde{x}^{m,k}$ in such a way that $\tilde{x}^{m,k}$ is on or arbitrary close to the segment $[0, \frac{1}{2}N]$.

This section continues the idea of Trockel (1996) to approach cooperative games with methods from microeconomic theory. Considering sets of feasible utility allocations as production possibility sets representing the possible jointly "producable" utility allocations and transformation rates as prices goes back to Shapley (cf. Shapley, 1969). See also Mayberry et al. (1953). The identity of the Walrasian equilibrium of a finite bargaining economy $\mathcal{E}^S$ with the Nash solution of its underlying bargaining game S stresses the competitive feature of the Nash solution.

Moreover the Nash solution's coincidence with the Core of a large bargaining coalitional production economy with equal production possibilities for all agents reflects a different fairness aspect in addition to those represented by the axioms.

4. A Walrasian Demand Game

Consider a two person bargaining situation S as illustrated in Figure 3. The compact strictly convex set $S \subset \mathbb{R}^2$ represents all feasible utility allocations for two players. For simplicity assume that the efficient boundary ∂S of S is the graph of some smooth decreasing concave function from $[0, 1]$ to $[0, 1]$. Such a bargaining situation can be looked at as a two-person NTU-game, where S is the set of payoff vectors feasible for the grand coalition $\{1, 2\}$, while $\{0\}$ represents the payoffs for the one player coalitions. The normalization to $(0; 1, 1)$ is standard and reflects the idea that S arose as the image under the two players' cardinal utility functions of some underlying set of outcomes or allocations. Cardinality determines utility functions only up to positive affine transformations and therefore justifies our normalization. Now, consider the following modification of Nash's simple demand game due to Trockel (2000)

$$\Gamma^S = (\Sigma_1, \Sigma_2; \pi_1^S, \pi_2^S) \ .$$

$\Sigma_1 = \Sigma_2 = [0,1]$ are the players' sets of (pure) strategies. The payoff functions are defined by $\pi_i^S(x_1, x_2) := x_i \mathbf{1}_S(x_1, x_2) + z_i^S(x_i)\mathbf{1}_{S^C}(x_1, x_2)$. Here S^C is the complement of S in $[0,1]^2$ and $\mathbf{1}_S$ is the indicator function for the set S. Finally $z_i^S(x_i)$ is defined as follows: For each $x_i \in [0,1]$ the point $y^S(x_i)$ is the unique point on ∂S with $y_i^S(x_i) = x_i$. By $p^S(x_i)$ we denote the normal vector to ∂S at $y^S(x_i)$ normalized by $p^S(x_i) \cdot y^S(x_i) = 1$. Now $z_i^S(x_i)$ is defined by $z_i^S(x_i) = min(x_i, \frac{1}{2p_i^S(x_i)})$, $i = 1, 2$.

This game has a unique Nash equilibrium (x_1^*, x_2^*) that is strict, has the maxmin-property and coincides with the Nash solution of S, i.e. $\{(x_1^*, x_2^*)\} = N(S)$. The idea behind the payoff functions is it to consider for any efficient utility allocation y its value under the efficiency price vector $p(y)$. If the utility allocation could be sold at $p(y)$ on a hypothetical market and the revenue would be split equally among the players there is only *one* utility allocation such that both players could buy back their own utility with their incomes without the need of any transfer of revenue. This equal split of revenue in the payoff function corresponds to *equity* in Shapley's (1969) cooperative characterization of the λ-transfer value via equity and efficiency. As for our two-person bargaining games the λ-transfer value just singles out the Nash solution this result does not come as a big surprise. By supplementing efficiency, which characterizes the infinitely many equilibria in Nash's demand game, by the additional equity, embodied in the payoff functions $\pi_i^S, i = 1, 2$, one gets the Nash solution as the unique equilibrium of the modified demand game. This result provides obviously a non-cooperative foundation of the Nash solution in the sense of the Nash program.

The fairness concept behind the rules of this game is the equity coming from the Walrasian approach in Trockel (1996) mentioned above, where "equity" means equal shares in the production possibility set used to produce utility allocations.

5. On the Meaning of the Nash Product

One possible way to try to find out any fairness concept behind the Nash product is it to derive the Nash product as a social planner's welfare function based on certain axioms on his preference relation on the set of feasible utility allocations. This route had been followed by Trockel (1999). For 0-normalized two-person bargaining situations it is shown that a preference relation on S is representable by the Nash product if it is a binary relation on S that satisfies the following properties: lower continuity, neutrality, monotonicity, unit-invariance and indifference-invariance. Continuity is a technical assumption, monotonicity reflects the planner's benevolence by liking higher utilities of the player's move. Neutrality is certainly a fairness property. What about the remaining two properties? *Indifference-invariance* is defined by:

$$x \succ y, x' \sim x, y' \sim y \Longrightarrow x' \succ y' \qquad \forall\, x, y, x', y' \in S \ .$$

It says that equivalent utility allocations for the planner are perfectly substitutable for each other in any strict preference. It is a weak consistency property.

Unit invariance is defined by:

$$x \succ y \Longleftrightarrow z * x \succ z * y \qquad \forall\ x, y, z \in S, * \text{ denoting pairwise multiplication.}$$

This property reflects the fact that the planner's preference is not influenced by the choices of units of the players.

Interestingly enough these properties not containing the standard rationality properties of transitivity and completeness suffice to yield a complete, transitive, continuous preordering on S representable by the Nash product. The only obvious fairness property is neutrality.

The Nash product itself is not seen in the literature as an easily interpretable function, not to speak of one reflecting any kind of fairness. This approach to the Nash solution is based on Trockel (2003). Concerning its direct interpretation the situation is best described by the quotation of Osborne and Rubinstein (1994, p. 303):

> Although the maximization of a product of utilities is a simple mathematical operation it lacks a straightforward interpretation; we view it simply as a technical device.

It is the purpose of the remaining part of this section to provide one straightforward, in fact surprising interpretation. Maximizing the Nash product is equivalent to finding maximal elements of one natural completion of the Pareto ordering. The maximal elements for the other natural completion are just the Pareto optimal points.

We shall look at the vector ordering and complete preorderings on compact subsets of $\mathbb{R}^n$ but restrict the analysis without loss of generality to the case $n = 2$. A complete preordering $\succsim$ on a compact set S is a complete, transitive (hence reflexive) binary relation on S. The weak vector ordering $\geq$ in contrast fails to be complete. It is, however, transitive, too. To make things simple assume $\succsim$ on S to be continuous, hence representable by a continuous utility function $u : S \longrightarrow \mathbb{R}$. The $\succsim$-maximal elements are given by the set of maximizers of u on S, i.e. $argmax_{x \in\ S}\ u(x)$.

Let $B_{\succsim}(x)$ be the set $\{x' \in S | x' \succsim x\}$ and $W_{\succsim}(x)$ the set $\{x' \in S | x \succsim x'\}$. For any $x', x \in S$ we obviously have:

$$x' \sim x \Leftrightarrow \lambda(B_{\succsim}(x)) = \lambda(B_{\succsim}(x')) \Leftrightarrow \lambda(W_{\succsim}(x)) = \lambda(W_{\succsim}(x'))\ .$$

Deviating from earlier notation λ now denotes the Lebesgue measure on $\mathbb{R}^2$ the extension of the natural measure of area in $\mathbb{R}^2$ to all Lebesgue measurable sets.

The correspondences $B_{\succsim} : S \Longrightarrow S$ and $W_{\succsim} : S \Longrightarrow S$ composed with the Lebesgue measure define alternative utility functions $\lambda \circ B_{\succsim}$ and $\lambda \circ W_{\succsim}$ representing $\succsim$ as well as u.

Now consider for the vector ordering $\geq$ the analogous sets $B_{\geq}(x)$, $W_{\geq}(x)$ for arbitrary $x \in S$:

$$B_{\geq}(x) = \{x' \in S | x' \geq x\},\ W_{\geq}(x) = \{x' \in S | x \geq x'\}\ .$$

Next, introduce the mappings $\lambda \circ B_{\geq}$ and $\lambda \circ W_{\geq}$ defined by:

$$\lambda \circ B_{\geq}(x) := \lambda(B_{\geq}(x)) \quad \text{and} \quad \lambda \circ W_{\geq}(x) := \lambda(W_{\geq}(x)) \ .$$

Both are mappings from S to $\mathbb{R}$ and define therefore preference relations that are completions of $\geq$.

We have $x \geq x' \implies \lambda(B_{\geq}(x)) \leq \lambda(B_{\geq}(x'))$ and $x \geq x' \implies \lambda(W_{\geq}(x)) \geq \lambda(W_{\geq}(x'))$. The two dual completions of $\geq$ are different in general:

$$x \succsim_1 x' :\Longleftrightarrow \lambda(B_{\geq}(x)) \leq \lambda(B_{\geq}(x')) \ ,$$
$$x \succsim_2 x' :\Longleftrightarrow \lambda(W_{\geq}(x)) \geq \lambda(W_{\geq}(x')) \ .$$

They only coincide when the binary relation one starts with is already a complete preordering. Notice, that $B_{\geq}(x)$ and $W_{\geq}(x)$ are in general proper subsets of $B_{\succsim_1}(x)$ and $W_{\succsim_2}(x)$, respectively.

Now we apply our gained insight to bargaining games. To keep things simple we define again a normalized two-person bargaining game S as the subgraph of a concave strictly decreasing function f from $[0,1]$ onto $[0,\,1]$. The two axes represent the players' utilities, S the feasible set of utility allocations. The vector ordering on S represents in this framework the Pareto ordering. The efficient boundary $graph f$ of S is the set of Pareto optimal points or vector maxima. Obviously each point x in $graph f$ minimizes the value of $\lambda \circ B_{\geq}$. In fact, for $x \in graph f$ we have $\lambda(B_{\geq}(x)) = 0$ Notice that $\lambda(W_{\geq}(x))$ takes different values when x varies in $graph f$.

Now, consider the set $argmax_{x\in S}\ \lambda(W_{\geq}(x))$ of maximizers of $\lambda \circ W_{\geq}$. This set is exactly the set $\{N(S)\}$ where $N(S)$ is the Nash solution of S. Maximizing the Nash product $x_1 x_2$ for $x \in S$ means maximizing the measure of points in S *Pareto dominated* by x'. Hence the two completions $\succsim_1, \succsim_2$ of the Pareto ordering $\geq$ on S have as their sets of maximizers the Pareto efficient boundary and the Nash solution, respectively. Thus we have shown that two different methods of representing complete preorderings via the measure of better sets versus worse sets may be applied as well to incomplete binary relations. Here they lead to two different functions inducing two different complete preorderings.

Applied to the non-complete Pareto ordering on a compact set S representing a bargaining situation the two completions have as their respective sets of maximizers the Pareto efficient boundary and the Nash solution of S. This result provides a straightforward interesting interpretation of the Nash solution as a dual version of Pareto optimality. In contrast to the latter it has the advantage to single out a unique point in the efficient boundary.

The idea of defining rankings by counting the less preferred alternatives has an old tradition in social choice theory as the famous Borda Count (cf. Borda, 1781) shows. In our context with a continuum of social alternatives counting is replaced by measuring. The level sets of the Nash product collect those utility allocations Pareto dominating equally large (in terms of Lebesgue measure) sets of alternatives.

6. Concluding Remarks

The Nash solution is the most popular and most frequently used bargaining solution in the economic and game theoretic literature. Authors working on efficient bargaining on labour markets predominantly use the Nash solution. Experiments on bargaining have been numerous and in various frameworks. Altogether they do not provide unanimous support for the Nash solution. But Binmore et al. (1993) provide empirical evidence for the Nash solution in laboratory experiments. Young (1993) presents an evolutionary model of bargaining supporting the Nash solution. And Skyrms (1996, p.107) writes:

> The evolutionary dynamics of distributive justice in discrete bargaining games is evidently more complicated than any one axiomatic bargaining theory. But our results reveal the considerable robustness of the Nash solution.

Despite the popularity of the Nash solution in the economic literature mentioned above Skyrms continues:

> Perhaps philosophers who have spent so much time discussing the utilitarian and Kalai-Smorodinsky schemes should pay a little more attention to the Nash bargaining solution.

Even if not a philosopher, in the present article I followed this advice by trying to find traces of fairness in different representations of the Nash solution available in the literature.

References

Binmore, K. 1987. "Nash Bargaining Theory I, II", in: K. Binmore and P. Dasgupta (eds.): *The Economics of Bargaining*, Cambridge: Basic Blackwell.

Binmore, K. 1997. "Introduction", in: J.F. Nash (ed.): *Essays on Game Theory*, Cheltenham: Edward Elgar.

Binmore, K., A. Rubinstein, and A. Wolinsky. 1986. "The Nash Bargaining Solution in Economic Modelling", *Rand Journal of Economics* 17, 176–188.

Binmore, K., J. Swierzbinski, S. Hsu, and C. Proulx. 1993. "Focal Points and Bargaining", *International Journal of Game Theory* 22, 381–409.

Borda, J.C. 1781. *Mémoire sur les élections au scrutin*, Paris: Histoire de L'Académie Royale des Sciences.

Debreu, G., and H. Scarf. 1963. "A Limit Theorem on the Core", *International Economic Review* 4, 235–246.

Hildenrand, W. 1974. *Core and Equilibria of a Large Economy*, Princeton: Princeton University Press.

Howard, J.V. 1992. "A Social Choice Rule and Its Implementation in Perfect Equilibrium", *Journal of Economic Theory* 56, 142–159.

Lensberg, T. 1988. "Stability and Collective Rationality", *Journal of Economic Theory* 45, 330–341.

Mayberry, J.P., J.F. Nash, and M. Shubik. 1953. "A Comparison of Treatments of a Duopoly Situation", *Econometrica* 21, 128–140.

Nash, J.F. 1951. "Non-cooperative Games", *Annals of Mathematica* 54(2), 286–295.

Nash, J.F. 1953. "Two-person Cooperative Games", *Econometrica* 21, 128–140.

Osborne, M.J., and A. Rubinstein. 1990. *Bargaining and Markets*, Academic Press, New York.

Osborne, M.J., and A. Rubinstein. 1994. *A Course in Game Theory*, Cambridge: MIT Press.

Rubinstein, A. 1982. "Perfect Equilibrium in a Bargaining Model", *Econometrica* 50, 97–109.

Shapley, L. S. 1969. "Utility Comparison and the Theory of Games", in: *La decision: aggregation et dynamique des ordres de preference*, Paris, 251–263.

Shubik, M. 1985. *Game Theory in the Social Sciences*, Cambridge: MIT Press.

Skyrms, B. 1996. *Evolution of the Social Contract*, Cambridge: Cambridge University Press.
Trockel, W. 1996. "A Walrasian Approach to Bargaining Games", *Economics Letters* 51, 295–301.
Trockel, W. 1999. "Rationalizability of the Nash Bargaining Solution", *Journal of Economics* 8, 159–165.
Trockel, W. 2000. "Implementation of the Nash Solution Based on Its Walrasian Characterization", *Economic Theory* 16, 277–294.
Trockel, W. 2003. *On the Meaning of the Nash Product*, Bielefed: IMW-working paper Nr. 354.
Trockel, W. 2005. "Core-Equivalence for the Nash Bargaining Solution", *Economic Theory* 25, 255–263.
Van Damme, E. 1986. "The Nash Bargaining Solution Is Optimal", *Journal of Economic Theory* 38, 78–100.
Young, H.P. 1993. "An Evolutionary Model of Bargaining", *Journal of Economic Theory* 59, 145–168.
Young, H.P. 1994. *Equity in Theory and Practice*, Princeton: Princeton University Press.

Walter Trockel
Institut für Mathematische Wirtschaftsforschung
Universität Bielefeld
D-33501 Bielefeld
Germany
wtrockel@uni-bielefeld.de