Econometrica, Vol. 47, No. 2 (March, 1979)

THE NASH SOCIAL WELFARE FUNCTION

BY MAMORU KANEKO AND KENJIRO NAKAMURA

We investigate the notion of the Nash social welfare function and make a fundamental assumption that there exists a distinguished alternative called an origin, which represents one of the worst states for all individuals in the society. Under this assumption, in Sections 1 and 2, we formulate several rationality criteria that a reasonable social welfare function should satisfy. Then we introduce the Nash social welfare function and the Nash social welfare indices which are the images of the welfare function. The function is proved to satisfy the criteria. In Section 3 it is shown that the Nash social welfare function is the unique social welfare function that satisfies the criteria. Then, in Section 4, we examine two examples which display plausibility of the welfare function.

1. INTRODUCTION

IT IS REMARKABLE and well known that Nash [**20**] provided a unique arbitration scheme for two-person bargaining games. Luce and Raiffa [**16**] reconsidered Nash's work and discussed its relation to the problem of social choice by giving an n-person modification of the Nash axioms for the two-person case. Sen [**25**] also pointed out this relationship and critically discussed Nash's scheme as a solution of social choice.

In the present paper we would like to carry their idea further and formulate the rationality criteria from which we derive Nash's arbitration scheme as a unique possible social welfare function in view of the social choice theory that has been developed since the work of Arrow [**1**]. Under a different approach from that adopted here, DeMeyer and Plott [**4**] provided a social welfare function by using "relative intensity" of preferences, in which Nash's scheme appeared as a special case.

A nonempty *set of alternatives*, denoted by X, is a set of social states that could be obtained as the result of the social decision process. The *society of individuals*, denoted by N, is the set of the individuals in the social decision process. N is assumed to be finite with cardinality $|N| = n$, i.e., $N = \{1, \ldots, n\}$. Here we assume the existence of a distinguished alternative x_0 which represents one of the worst states for all individuals that we may imagine, e.g., we may imagine all the members of the society die. The existence of such an alternative may also be explained by Isbell's argument that the utility space of an individual is bounded.[1] We call this alternative an *origin* and evaluate the social welfare by considering relative increases of the individuals' welfare from this position. We note that we never assume a status-quo as an origin. Operationally an origin can be chosen arbitrarily according to cases. For example, in problems of consumer choice, the origin may be chosen as the n-tuple of consumers' initial endowments, which is often implicitly assumed when one considers the welfare effect of competitive equilibria. However, in general the origin should be regarded as one of the worst states that can occur for the individuals. We will discuss this point and provide examples in a subsequent section.

[1] See also Owen [**23**, pp. 133–134].

With the above definition in mind, we always associate the origin x_0 with X. Let

$$X^* = X \cup \{x_0\};$$

we call X^* the *basic space of alternatives.* In the following, however, we will consider the probability distributions over X^* and regard them as the possible alternatives for the social choice problem. We give two reasons for this supposition. The first is that elements of uncertainty undoubtedly play an important role in the economy and should be included even in welfare economic theory. But one may object to this view because the uncertainty of economy follows a specific probability distribution and it is unreasonable to assume all probability distributions over X^*. Though this observation may be right, we do not in general specify the nature of the probability distribution and once we derive a reasonable welfare function under this supposition, we may apply it to a social choice problem with a specific probability distribution according to cases. The second is that the basic space of alternatives X^* often may be determined by individuals' strategic behavior and they may well use mixed strategies in this case. We define some notations to introduce the elements of uncertainty in our problem.

Let a_i $(i = 1, \ldots, t)$ be any alternatives in X^* and $\alpha = (\alpha_1, \ldots, \alpha_t)$ be a probability distribution, i.e., $\Sigma_{i=1}^t \alpha_i = 1$ and $\alpha_i \geqslant 0$ $(i = 1, \ldots, t)$. Then, by $(\alpha_1 a_1 * \ldots * \alpha_t a_t)$, we mean the *lottery* which has t possible outcomes a_i with probabilities α_i, respectively. For a subset A of X^*, let $m(A)$ denote the set of lotteries whose outcomes are any finite number of alternatives belonging to A, and we call $m(A)$ the *mixed extension* of A. If A is finite, $m(A)$ is clearly the set of lotteries whose outcomes are alternatives belonging to A. We call a lottery in $m(X^*)$ a *mixed alternative,* and we call an original alternative in X^* a *pure alternative.* Then we call $m(X^*)$ the *space of alternatives.* It is assumed that each individual has a weak ordering R_i over $m(X^*)$ that satisfies the von Neumann and Morgenstern utility axioms.[2] We use the following notation:[3]

$$(1.1) \qquad \begin{aligned} &aP_ib \quad \text{for} \quad \sim bR_ia, \\ &aI_ib \quad \text{for} \quad aR_ib \text{ and } bR_ia. \end{aligned}$$

We have already introduced the origin x_0 and explained its interpretation. With that interpretation in mind, we assume

(1.2) for every $i \in N$, xR_ix_0 for all $x \in X$.

We denote by $\mathscr{R}$ the set of all R_i that specifies the properties assumed above. A *profile* of individual preferences is an n-vector $p = (R_1, \ldots, R_n)$ representing the weak ordering R_i $(\in \mathscr{R})$ of each individual in the society.

We strengthen (1.2) for mathematical simplicity. For a given profile $p = (R_1, \ldots, R_n)$, let us define

$$\tilde{m}_p(X^*) = \{x \in m(X^*) | xP_ix_0 \text{ for all } i \in N\}.$$

[2] See, e.g., Owen [**23**, pp. 126–131]. We also assume the usual laws of linear algebra for the combination of alternatives by means of lotteries.

[3] We use $\sim$ for the symbol of negation.

We assume that

(1.3) $\tilde{m}_p(X^*)$ is nonempty.

To consider the meaning of this assumption, we give the following lemma. The proof is not difficult and omitted.

LEMMA 1.1: *$\tilde{m}_p(X^*)$ is empty if and only if there is some $j \in N$ such that xI_jx_0 for all $x \in m(X^*)$.*

Hence, when $\tilde{m}_p(X^*)$ is empty, we can delete individuals for whom all alternatives in $m(X^*)$ are indifferent to the origin x_0, and we may consider the social choice problem for the remaining individuals. In this case, we will also have a situation like (1.3) except for the case where all members in N are deleted. Then, we may assume (1.3) without loss of generality.

The basic assumption underlying the studies of social choice theory is that the social choice is related to the preferences of the individuals of the society. Then, as usual, we define a *social welfare function* W as a mapping which assigns to each profile of individual preferences $p = (R_1, \ldots, R_n)$ a social preference relation R over $\tilde{m}_p(X^*)$; i.e.,

(1.4) $R = W(R_1, \ldots, R_n)$.

As with individuals, we suppose that a social preference relation R is a weak ordering on $\tilde{m}_p(X^*)$. We use the following notation:

(1.5) aPb for $\sim bRa$,
aIb for aRb and bRa.

Here we note that a social preference relation R is *not* defined on $m(X^*)$, *but* on $\tilde{m}_p(X^*)$ depending on a given profile p. This means that, under assumption (1.3), we presuppose that no alternatives in $m(X^*) - \tilde{m}_p(X^*)$ are selected as a socially best alternative and that any such alternatives are not compared to each other.[4]

Now one may wonder why the social preference relation R does not satisfy the von Neumann and Morgenstern utility axioms, since we assumed individual preferences do. We give a simple example explaining the reason.[5]

EXAMPLE 1.1: Let R be a social preference relation satisfying von Neumann and Morgenstern utility axioms. Particularly R must satisfy the following axiom:

(1.6) If aIc, then for any $0 \leq \alpha \leq 1$, b,
$(\alpha a * (1-\alpha)b)I(\alpha c * (1-\alpha)b)$.[6]

Let a fixed amount of money M be distributed to two homogeneous individuals. Alternatives $a = (M, 0)$ and $c = (0, M)$ mean the allocations that distribute M

[4] In fact, we cannot construct any social welfare function which takes a value in $m(X^*)$ and satisfies the rationality criteria given in the next section.

[5] This example seems to be similar to Diamond's example [**5**].

[6] See, e.g., Owen [**23**, p. 127, Axiom 6.2.4].

only to individual i ($i = 1, 2$), respectively. If we treat the two individuals symmetrically, it is reasonable to assume that aIc. Let us put $b = c$ and $\alpha = 1/2$ in (1.6). Then we get

$$((1/2)a * (1/2)c)I((1/2)c * (1/2)c), \qquad \text{i.e.,}$$
$$((1/2)a * (1/2)c)Ic.$$

Hence the society is indifferent between the case in which every individual has a chance to get the money and the case in which one of the individuals has no chance to get it. This seems to violate an intuitive understanding of the principle of equity.

We regard, in general, the rationality criteria for the society as a whole as different from those for the individuals themselves. This is the reason we do not postulate that the social preference relation should satisfy the von Neumann and Morgenstern utility axioms.[7]

2. THE RATIONALITY CRITERIA AND THE POSSIBILITY THEOREM

From many social welfare functions defined in Section 1, we would like to select only one function as a possible candidate for a reasonable social welfare function. For this purpose, we will postulate several plausible rationality criteria that social welfare functions should satisfy. The class of rationality criteria with which we will be concerned is the one specifying how social preference should vary in response to variations in individual preferences.

The first condition is the most fundamental property called Pareto optimality.

CONDITION I (Pareto Optimality): Let p be any profile and let a and b be in $\tilde{m}_p(X^*)$. If aR_ib for all $i \in N$ and aP_jb for some $j \in N$, then aPb.

The second condition is a modified version of the so called independence of irrelevant alternatives condition. The condition states that the social preference between two alternatives depends only on individuals' preferences between these alternatives, regardless of individuals' preferences relating to other alternatives. Our modification comes from the use of the von Neumann and Morgenstern utilities and the existence of the origin. The latter should be noted because we have made a fundamental assumption that we evaluate the social welfare by considering relative increases of individuals' welfare from the origin. We state this second condition together with the neutral property for the alternatives. This neutral property means that the social welfare function does not depend on the labeling of the alternatives.

[7] Then our social preference relation may be different from that of Harsanyi [**8**], only because the underlying situation considered here is different from that of Harsanyi. This example is not intended to support Diamond's criticism against Harsanyi's social welfare function. Anyway we do not exclude the possibility that the von Neumann and Morgenstern utility axioms are satisfied.

CONDITION II (Independence of Irrelevant Alternatives with Neutral Property): Let $p=(R_1,\ldots,R_n)$ be a profile with $R=W(p)$, and let $p'=(R'_1,\ldots,R'_n)$ be any other profile with $R'=W(p')$. Let $a_1, a_2 \in \tilde{m}_p(X^*)$ and let $b_1, b_2 \in \tilde{m}_{p'}(X^*)$. Suppose, for all $i \in N$,

$$(\alpha_1 a_1 * \alpha_2 a_2 * \alpha_3 x_0) R_i (\beta_1 a_1 * \beta_2 a_2 * \beta_3 x_0)$$

if and only if

$$(\alpha_1 b_1 * \alpha_2 b_2 * \alpha_3 x_0) R'_i (\beta_1 b_1 * \beta_2 b_2 * \beta_3 x_0) \tag{2.1}$$

for all probability distributions α and β. Then, $a_1 R a_2$ if and only if $b_1 R' b_2$.

For a given social choice problem, if we exchange the roles of individuals, e.g., if two individuals exchange the commodity bundles they receive and the preference relations they possess, then the judgement of the society is invariant. That is, the social welfare function does not depend on the naming of individuals. The following condition states that the social welfare function satisfies the principle of equity.

CONDITION III (Anonymity): Let π be any permutation of the individuals $(1,\ldots,n)$. Let $p=(R_1,\ldots,R_n)$ and let $p_\pi=(R_{\pi(1)},\ldots,R_{\pi(n)})$. Then

$$W(p)=W(p_\pi). \tag{2.2}$$

Here note that $\tilde{m}_p(X^*)=\tilde{m}_{p_\pi}(X^*)$.

The final condition is the following. This is the same as the continuity axiom assumed for the von Neumann and Morgenstern utilities.

CONDITION IV (Continuity): Let p be any profile and let a, b, and c in $\tilde{m}_p(X^*)$ satisfy $aPcPb$. Then there exists some α, $0 \leqslant \alpha \leqslant 1$, such that

$$(\alpha a * (1-\alpha) b) I c.$$

In the following we will show that Conditions I, II, III, and IV are consistent and yield a unique social welfare function. We now define this social welfare function, which we call the *Nash social welfare function*.

Let $U(R_i)$ be the set of von Neumann and Morgenstern utility functions representing R_i ($\in \mathcal{R}$). $U(R_i)$ is nonempty. If $u_i \in U(R_i)$ and p, q are real numbers with $p>0$, then $pu_i+q \in U(R_i)$, and if u_i and $v_i \in U(R_i)$, then there are real numbers p, q with $p>0$, such that $v_i = pu_i + q$.

Let $p=(R_1,\ldots,R_n)$ be a profile of individual preferences and let $u_i \in U(R_i)$ $(i=1,\ldots,n)$ be arbitrarily chosen utility functions. We denote $u=(u_1,\ldots,u_n)$. We define a function on $\tilde{m}_p(X^*)$ as

$$w_0(u(x))=\sum_{i=1}^{n} \log\,(u_i(x)-u_i(x_0)). \tag{2.3}$$

We put a social preference relation R_0 over $\tilde{m}_p(X^*)$ as

(2.4) $\quad xR_0y \quad$ if and only if $\quad w_0(u(x)) \geqslant w_0(u(y))$.

By noting that R_0 does not depend on the selection of $u_i \in U(R_i)$ $(i=1, \ldots, n)$, we define the Nash social welfare function W_0 as $W_0(p)=R_0$. Each $w_0(u(x))$ representing R_0 is called a *Nash social welfare index.*

For $u=(u_1, \ldots, u_n)$, $u_i \in U(R_i)$ $(i=1, \ldots, n)$, and a subset A of $m(X^*)$, we put $u(A)=\{u(a)|a \in A\}$, which is a subset of Euclidean space of n dimensions, E^n. For u and v in E^n, we define order relation as usual: $u \geqslant v$ if and only if $u_i \geqslant v_i$ for $i=1, \ldots, n$. $u>v$ if and only if $u_i>v_i$ for $i=1, \ldots, n$. $u \geqslant v$ if and only if $u \geqslant v$ and $u \neq v$.

THEOREM 2.1 (The Possibility Theorem): *The Nash social welfare function satisfies conditions I, II, III, and IV.*

PROOF: For Conditions I, III, and IV, the proof is straightforward and omitted. Suppose the assumptions in Condition II hold. Let $u_i \in U(R_i)$ and $u_i' \in U(R_i')$ $(i=1, \ldots, n)$. We transform each u_i' by the following positive linear transformation:

(2.5) $$v_i'(x)=\frac{u_i(a_1)-u_i(x_0)}{u_i'(b_1)-u_i'(x_0)}[u_i'(x)-u_i'(x_0)]+u_i(x_0) \qquad (i=1, \ldots, n).$$

Then $v_i' \in U(R_i')$ for all $i \in N$. Clearly

(2.6) $\quad v_i'(x_0)=u_i(x_0) \quad$ and $\quad v_i'(b_1)=u_i(a_1) \quad$ for all $i \in N$.

Moreover it is not difficult to see $v_i'(b_2)=u_i(a_2)$ for all $i \in N$. Hence it is known that $u(m(\{x_0, a_1, a_2\}))=v'(m(\{x_0, b_1, b_2\}))$ and $u(m(\{x_0, a_1, a_2\}))$ is a similar figure of $v(m(\{x_0, b_1, b_2\}))$. $a_1R_0a_2$ means $w_0(u(a_1)) \geqslant w_0(u(a_2))$. By the above fact, this is equivalent to $w_0(v'(b_1)) \geqslant w_0(v'(b_2))$, which means $b_1R_0b_2$. Hence Condition III is satisfied. *Q.E.D.*

3. THE UNIQUENESS OF THE SOCIAL WELFARE FUNCTION

In this section we will show that the Nash social welfare function is uniquely determined by Conditions I, II, III, and IV.

Let a social welfare function $R=W(p)$ be given and let X have at least 3 alternatives other than x_0. As we will use a strong lemma due to Osborne [**22**], a weak ordering on the positive orthant of E^n, denoted by E_+^n, is introduced. Let x and y be any elements in E_+^n. We choose arbitrarily two alternatives a and b $(\neq x_0)$ from X^* and define utility functions u_i so that they satisfy

$$\begin{aligned} &u_i(a)=x_i && \text{for all } i \in N,\\ (3.1)\qquad &u_i(b)=y_i && \text{for all } i \in N,\\ &u_i(x_0)=0 && \text{for all } i \in N. \end{aligned}$$

Let $u_i \in U(R_i)$ $(i=1,\ldots,n)$. Then, by $R = W(R_1,\ldots,R_n)$, we can define a binary relation between x and y as

(3.2) $\quad x \gtrsim y \quad$ if and only if $\quad aRb.$

Here if we choose another $c, d \in X^*$ and $v_i \in U(R_i')$ $(i=1,\ldots,n)$ with (3.1), then, by Condition II, aRb if and only if cRd. Hence the relation is well defined.

LEMMA 3.1: *Let X have at least 3 alternatives other than x_0 and let Condition II be assumed. Then the binary relation $\gtrsim$ is a weak ordering.*

PROOF: Reflexivity and connectedness are trivial. We prove transitivity. Let x, y, and z be in E_+^n such that $x \gtrsim y$ and $y \gtrsim z$. Since X has at least 3 alternatives other than x_0, we can choose utility functions $u_1,\ldots,u_n$ such that

$$u(x_0)=0, \qquad u(a)=x, \qquad u(b)=y, \qquad u(c)=z,$$

for some $a, b, c \in X$. Let $R = W(R_1,\ldots,R_n)$, where $u_i \in U(R_i)$ $(i=1,\ldots,n)$. Then, by definition, aRb and bRc. By the transitivity of R, we have aRc. This shows that $x \gtrsim z$. *Q.E.D.*

It is easy to have the following lemma.

LEMMA 3.2: *Let X have at least 3 alternatives other than x_0 and let Conditions I and II be satisfied. Let x and y be in E^n such that $x \geqslant y$. Then $x \succ y$.*

LEMMA 3.3: *Let X have at least 3 alternatives other than x_0 and let Conditions I, II, and IV be assumed. Then there exists a real-valued function $F(x)$ over E_+^n such that*

$$x \gtrsim y \qquad \textit{if and only if} \qquad F(x) \geqslant F(y).$$

PROOF: Let e denote the vector of E^n, all components of which are 1. If $x \in E_+^n$, there are positive real numbers λ_0 and λ_1 such that $\lambda_0 e < x < \lambda_1 e$. Then, by Lemma 3.2, $\lambda_0 e \prec x \prec \lambda_1 e$. Let alternatives a, b, c and utility functions u_i be chosen so that they satisfy: $u(a)=\lambda_1 e$, $u(b)=\lambda_0 e$, $u(c)=x$, and $u(x_0)=0$. Let $u_i \in U(R_i)$ $(i=1,\ldots,n)$ and $R = W(R_1,\ldots,R_n)$. Then $aPcPb$. From Condition IV, there is an α $(0 \leqslant \alpha \leqslant 1)$ such that

$$(\alpha a * (1-\alpha)b)Ic.$$

This implies that $\alpha[\lambda_1 e]+(1-\alpha)[\lambda_0 e] \sim x$, i.e., $[\lambda_0 - \alpha(\lambda_0-\lambda_1)]e \sim x$. We put $\mu_x = \lambda_0 - \alpha(\lambda_0 - \lambda_1)$. Here μ_x will be known to be independent of λ_0 and λ_1, and we can define $F(x)=\mu_x$. We let x and y in E_+^n be such that $x \gtrsim y$. This is clearly equivalent to $\mu_x \geqslant \mu_y$, i.e., $F(x) \geqslant F(y)$. *Q.E.D.*

By Lemma 3.3, we have a real-valued function representing $\gtrsim$. Note that any monotone increasing transformation of F has this property.

LEMMA 3.4: *Let X have at least 3 alternatives other than x_0. Let Conditions I, II, III, and IV be assumed. Then the following properties hold. Let x and $y \in E^n_+$.*

(i) *$F(x) \geqslant F(y)$ if and only if*

$$F(\lambda_1 x_1, \ldots, \lambda_n x_n) \geqslant F(\lambda_1 y_1, \ldots, \lambda_n y_n),$$

where λ_i are any positive real numbers.

(ii) *If $x \geqslant y$, then $F(x) > F(y)$.*

(iii) *$F(x) = F(x_\pi)$, where π is a permutation of $(1, \ldots, n)$ and $x_\pi = (x_{\pi(1)}, \ldots, x_{\pi(n)})$.*

PROOF: (i) Let alternatives a, b and utility functions u_i be chosen so that they satisfy (3.1). Let $u_i \in U(R_i)$ $(i = 1, \ldots, n)$ and let $R = W(R_1, \ldots, R_n)$. We have $\lambda_i u_i(a) = \lambda_i x_i$, $\lambda_i u_i(b) = \lambda_i y_i$, $\lambda_i u_i(x_0) = 0$, and $\lambda_i u_i \in U(R_i)$ for all $i \in N$. Then $F(\lambda_1 x_1, \ldots, \lambda_n x_n) \geqslant F(\lambda_1 y_1, \ldots, \lambda_n y_n)$ if and only if $(\lambda_1 x_1, \ldots, \lambda_n x_n) \gtrsim (\lambda_1 y_1, \ldots, \lambda_n y_n)$, which is equivalent to aRb. This is also equivalent to $F(x) \geqslant F(y)$.

(ii) If follows from Lemmas 3.2 and 3.3.

(iii) To prove the symmetry of $F(x)$, it is sufficient to show its symmetry under any permutation that exchanges only two components. Let i and j be any elements of N and let π be the permutation such that $\pi(i) = j$, $\pi(j) = i$, and $\pi(k) = k$ for all $k \neq i, j$. We assume without loss of generality that $i = 1$ and $j = 2$. Let alternatives a, b and utility functions u_i be chosen so that $u(a) = x$, $u(b) = x_\pi$, and $u(x_0) = 0$, where $u_i \in U(R_i)$ $(i = 1, \ldots, n)$. Let $p = (R_1, \ldots, R_n)$ and $R = W(p)$. If utility functions v_i are defined by $v_i = u_{\pi(i)}$ for all $i \in N$, then $v_i \in U(R_{\pi(i)})$ for all $i \in N$ and

$$\begin{aligned} v_1(a) &= u_2(a) = x_2 = u_1(b), \\ v_2(a) &= u_1(a) = x_1 = u_2(b), \\ v_k(a) &= u_k(a) = x_k = v_k(b) \qquad \text{for all } k \neq 1, 2, \end{aligned}$$

and

$$\begin{aligned} v_1(b) &= u_2(b) = x_1 = u_1(a), \\ v_2(b) &= u_1(b) = x_2 = u_2(a), \\ v_k(b) &= u_k(b) = x_k = u_k(a) \qquad \text{for all } k \neq 1, 2. \end{aligned}$$

Then, by Condition II, aRb if and only if $bR'a$, where $R' = W(p_\pi)$. But, by Condition III, $R' = R$; i.e., aRb means bRa. If $F(x) \geqslant F(x_\pi)$, then $x \gtrsim x_\pi$ or aRb, which means bRa or $x_\pi \gtrsim x$. Then $F(x_\pi) \geqslant F(x)$. Hence $F(x) = F(x_\pi)$. *Q.E.D.*

LEMMA 3.5 (Osborne): *Let G be a real-valued function on E^n_+. Suppose the following properties:*

(A) *If $x_i \geqslant y_i$, then $G(x_1, \ldots, x_n) \geqslant G(x_i, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n)$ for all $i \in N$.*

(B) *$G(x) \geqslant G(y)$ if and only if $G(\lambda_1 x_1, \ldots, \lambda_n x_n) \geqslant G(\lambda_1 y_1, \ldots, \lambda_n y_n)$ for all x, $y \in E^n_+$ and positive real numbers λ_i $(i = 1, \ldots, n)$.*

Then there are nonnegative real constants $c_1, \ldots, c_n$ and monotone increasing function V over E^1 such that

$$G(x) = V\Big(\prod_{i \in N} x_i^{c_i}\Big).$$

THEOREM 3.6 (The Uniqueness Theorem): *Let X have at least 3 alternatives other than x_0. Then a social welfare function W satisfies Conditions I, II, III, and IV if and only if W is the Nash social welfare function W_0.*

PROOF: Let $p = (R_1, \ldots, R_n)$ be any profile and let $R = W(p)$. Let a and b be any alternatives in $\tilde{m}_p(X^*)$. Let $u_i \in U(R_i)$ for all $i \in N$. Then aRb if and only if $(u(a) - u(x_0)) \gtrsim (u(b) - u(x_0))$, which means

$$F(u(a) - u(x_0)) \geqslant F(u(b) - u(x_0)). \tag{3.3}$$

Since $F(x)$ satisfies the hypotheses of Lemma 3.5 from (i) and (ii) of Lemma 3.4, there are nonnegative real numbers $c_1, \ldots, c_n$ and a monotone increasing function V over E^1 such that $F(x) = V\,(\prod_{i \in N} x_i^{c_i})$.

Let i and $j \in N$, and let π be a permutation with $\pi(i) = j$ and $\pi(j) = i$ and $\pi(k) = k$ for all $k \neq i, j$. By (iii) of Lemma 3.4, $F(x) = F(x_\pi)$, which implies

$$x_i^{c_i} x_j^{c_j} = x_j^{c_i} x_i^{c_j} \qquad \text{for all } x \in E^n_+.$$

Then $c_i(\log x_i - \log x_j) = c_j(\log x_i - \log x_j)$. If we put $x_i \neq x_j$, then $c_i = c_j$. By the arbitrariness of i and j, all c_i are identical with the same $c \geqslant 0$. If $c = 0$, this contradicts (ii) of Lemma 3.4. Then $c > 0$. Hence $F(x) = V((\prod_{i \in N} x_i)^c)$. Hence (3.3) is equivalent to

$$\prod_{i \in N} (u_i(a) - u_i(x_0)) \geqslant \prod_{i \in N} (u_i(b) - u_i(x_0))$$

or

$$\sum_{i \in N} \log\,(u_i(a) - u_i(x_0)) \geqslant \sum_{i \in N} \log\,(u_i(b) - u_i(x_0)).$$

Then $w_0(u(a)) \geqslant w_0(u(b))$. This means that aR_0b, where $R_0 = W(p)$. This, together with Theorem 2.1, proves the theorem. *Q.E.D.*

Thus we have demonstrated that our social welfare function is uniquely determined as the Nash social welfare function W_0. Then for given $p = (R_1, \ldots, R_n)$, we can compare the relative ranking of all alternatives in $\tilde{m}_p(X^*)$ by the values of the Nash social welfare index $w_0(u(x))$, where $u_i \in U(R_i)$ $(i = 1, \ldots, n)$.

4. EXAMPLES

We are now in a position to exemplify the usefulness of the introduction of the origin to our purpose, and examine whether the Nash social welfare function fits well our intuitive understanding of social choices.

EXAMPLE 4.1. Let an amount of money M ($M > 0$) be distributed to two individuals 1 and 2. Let m_1 and m_2 ($m_1, m_2 > 0$) be 1's and 2's initial endowments of money, respectively. The set of alternatives X is given by

$$X = \{(x, M-x) | 0 \leq x \leq M\}. \tag{4.1}$$

Let individuals 1 and 2 have the same monotone increasing and concave utility functions $u(m)$ for money.

In Owen [**23**], $M = \$100$, $u(m) = \log m$, and individual 1 is assumed to be very rich with a very large m_1, while individual 2 is very poor with $m_2 = \$100$. Moreover, the status-quo is set as the origin, i.e., $x_0 = (m_1, m_2)$.[8] Then the socially most preferred alternative given by the Nash social welfare function is approximately (54.4, 45.6). The result is favorable to the rich and seems to violate an intuitive understanding of the principle of equity.[9] Does this violation come from the use of the Nash social welfare function? We think that the selection of the origin x_0 does not follow our understanding.

As we supposed in Section 1, the origin must be defined as $x_0 = (0, 0)$. We can put $u(0) = 0$ without loss of generality.[10] Then the socially most preferred alternative $(a, M-a)$ is given by solving the following:

$$\max\,(\log u(m_1 + x) + \log u(m_2 + M - x)) \tag{4.2}$$

subject to $0 \leq x \leq M$.

Let $u(m)$ be differentiable and let $u'(m) = du(m)/dm$.
If

$$\frac{u'(m_1+x)}{u(m_1+x)} = \frac{u'(m_2+M-x)}{u(m_2+M-x)},$$

then $m_1 + x = m_2 + M - x$, since $u'(m)/u(m)$ is monotone decreasing. If we put m_0 as $m_1 + m_0 = m_2 + M - m_0$, then (4.2) is solved as follows:

$$\begin{aligned} a &= M && \text{if} \quad m_0 \geq M, \\ a &= m_0 && \text{if} \quad 0 \leq m_0 \leq M, \\ a &= 0 && \text{if} \quad m_0 \leq 0. \end{aligned} \tag{4.3}$$

Hence the Nash social welfare function makes the endowments of money equal, if possible. If impossible, all the amount of money M is distributed only to the poor,

[8] In Owen [**23**, pp. 146–147], the marginal utility of the rich man's utility function $1/(m_1 + x)$ is assumed to be constant over $0 \leq x \leq 100$, since m_1 is very large.

[9] Of course, Owen gives a correct interpretation of the solution to the bargaining game.

[10] In the above example, $u(m) = \log m$ will be appropriately arranged so as to satisfy this requirement without any essential change.

which seems to be the correct answer to Owen's example, if one views the example as an income distribution problem. The result (4.3) seems to fit our intuition. Therefore, when one considers a general income distribution problem, one may use the Nash social welfare function by selecting the origin as the n-tuple of zero commodity bundles.

In the second example, we consider a situation which is similar to that in a counterexample of Harsanyi [**11**] directed against the maximin principle of Rawls [**24**], since this counterexample also seems to work against the Nash social welfare function at first glance.

EXAMPLE 4.2: Let individuals 1 and 2 be taken seriously ill. There is medicine, but only for one. Both surely die without the medicine. If individual 1 takes the medicine, he surely recovers. If individual 2 takes it, he recovers with probability 1/2 and dies with probability 1/2. Let x_1 and x_2 be the alternatives that 1 and 2 take the medicine, respectively. We assume that the utility function of individual i ($i = 1, 2$) satisfies $u_i(L) > u_i(D) = 0$, where L means that individual i is alive and well and D that he dies. Of course, the origin is $x_0 = (D, D)$, i.e., the state in which both die. Then

$$u_1(x_1) = u_1(L), \qquad u_1(x_2) = u_1(D) = u_1(x_0),$$

and

$$u_2(x_2) = (1/2)u_2(L), \qquad u_2(D) = u_2(x_0).$$

The socially most preferred alternative is $((1/2)x_1 * (1/2)x_2)$. This does not violate our intuition.

Suppose next that if individual 2 takes the medicine, then he recovers with only probability 1/100. Even in this case, the socially most preferred alternative is still $((1/2)x_1 * (1/2)x_2)$. This may contradict our intuition. One may well insist that, in this case, the appropriate probability of the choice of x_2 should be very small or zero. This objection is similar to that of Harsanyi against the maximin welfare function. The objection will be answered by a more complete description of the situation in our case.

When one raises this objection, he thinks of the situation as if he replaces the individuals by himself. Then the utility function of each individual cannot necessarily be independent of the state of another. But, in the above definition of utility function, each individual is interested in his own state and is not affected by the state of another. This is the reason why one feels that our result may violate his intuition. If one agrees that the utility function of each individual is independent of the state of another, he would not find the reason to insist that the appropriate probability of choice of x_2 should be very small or zero. Then, if one still has this insistence, one should change the definition of utility function u_1 and u_2 so that they satisfy:

$$u_1(L, D) > u_1(D, L) > u_1(x_0) = 0$$

and

$$u_2(D, L) > u_2(L, D) > u_2(x_0) = 0$$

where (L, D) (or (D, L)) means that individual 1 is alive and well (or dies) and individual 2 dies (or is alive and well). Here $u_1(D, L)$ and $u_2(L, D)$ may be very small and

$$u_1(x_1) = u_1(L, D) > u_1(x_2) = (1/100)u_1(D, L) > u_1(x_0)$$

and

$$u_2(x_2) = (1/100)u_2(D, L), \; u_2(x_1) = u_2(L, D) > u_1(x_0).$$

This formulation will represent one's observation and the Nash social welfare function suggests that the appropriate probability of choice of x_2 should be very small or zero. This result is consistent with one's intuition. Furthermore since $u_1(D, L)$ and $u_2(L, D)$ are very small, in the preceding case where individual 2 recovers with probability 1/2 by taking the medicine, the socially most preferred alternative $(p_1^0 x_1 * p_2^0 x_2)$ will be seen to be very close to $((1/2)x_1 * (1/2)x_2)$ but still $p_1^0 > p_2^0$.

5. CONCLUDING REMARKS

We have established the notion of the Nash social welfare function and Nash social welfare indices. As the indices have a simple analytical form, some positive analysis based on the indices would be of interest. We would like to stress the importance of conceptual connections between game theory and social choice theory. This general recognition has been growing, as the pioneering work of Farquharson [**7**], Luce and Raiffa [**16**], Dummett and Farquharson [**6**], Wilson [**28, 29**], Bloomfield [**3**], Nakamura [**17, 18**], Kaneko [**14**], and others have shown. The socially most preferred alternative with respect to the Nash social welfare function is generated from the corresponding bargaining process which will be a special case of n-person bargaining games. Then it would be most interesting to extend the analysis in order to relate some theory of general n-person bargaining games to social choice theory. This seems to shed new light on fields of game theory, social choice theory, and the theory of justice.

Finally the authors are grateful to Professor Mitsuo Suzuki for valuable discussions. They are also indebted to Professor Mikio Nakayama and Mr. Shin Ishikawa for helpful comments and referees of this journal for kind suggestions.

University of Tsukuba
and
Tokyo Institute of Technology

Manuscript received March, 1977; final revision received February, 1978.

REFERENCES

[1] ARROW, K. J.: *Social Choice and Individual Values*, 2nd Edition. New York: John Wiley and Sons, 1963.

[2] ———: "Some Ordinalist-Utilitarian Notes on Rawls' Theory of Justice," *Journal of Philosophy*, 70 (1973), 245-263.

[3] BLOOMFIELD, S. D.: "Social Choice Interpretation of the von Neumann-Morgenstern Game," *Econometrica*, 44 (1976), 105–113.

[4] DEMEYER, F., AND C. R. PLOTT: "A Welfare Function Using 'Relative Intensity' of Preferences," *Quarterly Journal of Economics*, 85 (1971), 179–186.

[5] DIAMOND, P. A.: "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: Comment," *Journal of Political Economy*, 75 (1967), 765–766.

[6] DUMMETT, M., AND R. FARQUHARSON: "Stability in Voting," *Econometrica*, 29 (1961), 33–43.

[7] FARQUHARSON, R. *Theory of Voting*. New Haven: Yale University Press, 1969.

[8] HARSANYI, J. C.: "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility," *Journal of Political Economy*, 63 (1955), 309–321.

[9] ———: "Cardinal Utility in Welfare Economics in the Theory of Risk-Taking," *Journal of Political Economy*, 61 (1953), 434–435.

[10] ———: "A Bargaining Model for the Cooperative n-person Game," in *Contributions to the Theory of Games IV*, ed. by A. W. Tucker & R. D. Luce. Princeton: Princeton University Press, 1959.

[11] ———: "Can the Maximin Principle Serve as a Basis for Morality?—A Critique of John Rawls' Theory," *The American Political Science Review*, 69 (1975), 594–606.

[12] ———: "Nonlinear Social Welfare Functions," *Theory and Decision*, 6 (1975), 311–332.

[13] ISBELL, J. R.: "Absolute Games," in *Contributions to the Theory of Games IV*, ed. by A. W. Tucker and R. D. Luce. Princeton: Princeton University Press, 1959.

[14] KANEKO, M.: "Necessary and Sufficient Conditions for Transitivity in Voting Theory," *Journal of Economic Theory*, 11 (1975), 385–393.

[15] ———: "Necessary and Sufficient Conditions for the Existence of a Nonempty Core of a Majority Game," *International Journal of Game Theory*, 4 (1975), 215–219.

[16] LUCE, R. D., AND H. RAIFFA.: *Games and Decisions*. New York: John Wiley and Sons, 1957.

[17] NAKAMURA, K.: "The Core of a Simple Game with Ordinal Preferences," *International Journal of Game Theory*, 4 (1975), 95–104.

[18] ———: "The Vetoers in a Simple Game with Ordinal Preferences," to appear in *International Journal of Game Theory*.

[19] ———: "A Composite Objective Function for the Multi-objective Programmings," *Research Report* B-37, Department of Information Sciences, Tokyo Institute of Technology, 1976.

[20] NASH, J. F.: "The Bargaining Problem," *Econometrica*, 18 (1950), 155–162.

[21] ———: "Two-person Cooperative Games," *Econometrica*, 21 (1953), 128–140.

[22] OSBORNE, D. K.: "Irrelevant Alternatives and Social Welfare," *Econometrica*, 44 (1976), 1001-1015.

[23] OWEN, G.: *Game Theory*. Philadelphia: W. B. Saunders, 1968.

[24] RAWLS, J.: *A Theory of Justice*. Cambridge, Massachusetts: Harvard University Press, 1971.

[25] SEN, A. K.: *Collective Choice and Social Welfare*. San Francisco: Holden-Day, 1970.

[26] SUZUKI, M., AND K. NAKAMURA: *Social Systems—A Game Theoretic Approach* (Japanese). Tokyo: Kyōritsu Shuppan, 1976.

[27] VON NEUMANN J., AND O. MORGENSTERN: *Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1953.

[28] WILSON, R.: "The Game-Theoretic Structure of Arrow's General Possibility Theorem," *Journal of Economic Theory*, 5 (1972), 14–20.

[29] ———: "Stable Coalition Proposals in Majority-rule Voting," *Journal of Economic Theory*, 3 (1971), 264–271.