Discussion Paper No. 321

LINEARITY, CONCAVITY, AND SCALE INVARIANCE IN SOCIAL CHOICE FUNCTIONS

by

Roger B. Myerson*

March 1978

Abstract. Three theorems are derived about social choice functions, which are defined on comprehensive convex subsets of utility-allocation space. Theorem 1 asserts that a linearity condition, together with Pareto-optimality, implies that a social choice function must be utilitarian. Theorem 2 asserts that a concavity condition, together with Pareto-optimality and independence of irrelevant alternatives, implies that a social choice function must be either utilitarian or colinear, where colinearity is a property closely related to the maximin criterion. Theorem 3 asserts that only dictatorships can satisfy scale invariance and independence of irrelevant alternatives.

*Graduate School of Management, Northwestern University, Evanston, IL 60201

LINEARITY, CONCAVITY, AND SCALE INVARIANCE IN SOCIAL CHOICE FUNCTIONS

by

Roger B. Myerson

1. Introduction

The utilitarian principle in ethical theory asserts that the best social policy is the one which gives the greatest total happiness to the individual members of society, as measured by summing utility numbers for all individuals. This principle has been advocated by philosophers going back to Bentham [2] and others, and more recently from the viewpoint of Bayesian decision theory by Harsanyi [5], [6].

The maximin principle in ethical theory asserts that the best social policy is the one which gives the greatest happiness to the most unfortunate individuals in society. Rawls [11] [12] has argued for this principle in his theory of justice. As long as there is any positive tradeoff between the welfare of different individuals, the maximin principle always leads to social choices in which all individuals are equally happy. So the maximin principle is (in most cases) also equivalent to the equity-constrained collective choice theories discussed in Kalai [7] and Myerson [9].

As Shapley [15] has pointed out, these two ethical principles both use interpersonal comparisons of utility, but in very different ways. Translated into the practical debates of daily life, the utilitarian principle asserts that "you should do something for me if it will hurt you less than it will help me", whereas the maximin principle asserts that "you should do something for me if you are better off than I am (or if you have gained more from our cooperation than I have)."

This paper will investigate some properties of social choice rules related to these two principles, with the goal of helping to explain why these two principles have been so important both in the development of ethical theories and in practical social decision-making. Another approach to this same question has recently been offered by Deschamps and Gevers [3], [4].

In Section 2, we develop the basic definitions relating to social choice problems and the choice functions which may be used to solve them. In Section 3, we show that certain Pareto-optimality and linearity conditions imply that social choices must be made according to the utilitarian principle, for some collection of vonNeumann-Morgenstern individual utility scales. This result is closely related to Harsanyi's Theorem V in [5], with our linearity condition playing a role analogous to the sure-thing principle in Harsanyi's work.

In Section 4, we investigate colinear choice functions, which satisfy a generalized version of the maximin or equity-constraint principle. We show that, if the linearity condition from Section 2 is replaced by a weaker concavity condition, and if an independence of irrelevant alternatives condition is added, then only the colinear and utilitarian choice functions are possible. This result suggests that it may be their shared concavity property which makes the utilitarian and maximin-type choice rules more appealing than other social choice rules. The concavity condition has a natural interpretation in terms of the timing of social decisions: it guarantees that all individuals should always prefer society to plan ahead.

Finally, in Section 5, we investigate the possibility of finding a social choice function which does not require the individuals' utility scales to be interpersonally comparable. This is desirable because interpersonal comparison of utility cannot be justified in the context of individual decision theory. We will show that, even without concavity, it is not possible to find a satisfactory choice function satisfying this scale invariance condition. This impossibility result can be compared to other impossibility results in social choice theory, such those of Arrow [1], Sen [14] (see especially his Theorem 8*2), and Kalai and Schmeidler [8]. It is also closely related to Nash's theory of bargaining [10], and shows the impossibility of deriving his solution concept without a reference point.

2. Basic Definitions

In this paper, social choice problems are represented by the sets of feasible utility allocations available to the society. We assume that there are n individuals, numbered 1,2,...,n, in the group or society. Thus any vector $x = (x_1, x_2, \ldots, x_n)$ in $\mathbb{R}^n$ can be interpreted as a utility allocation, so that the i^{th}-component x_i represents the payoff to individual i, measured in some vonNeumann-Morgenstern utility scale for individual i.

Given any vectors x and y in $\mathbb{R}^n$, we write $x \geq y$ (or $y \leq x$) iff $x_i \geq y_i$ for every $i = 1,2,\ldots,n$. Similarly $x > y$ (or $y < x$) means that $x_i > y_i$ for every i.

The usual dot product is used for vectors in $\mathbb{R}^n$, that is:

$$x \cdot y = \sum_{i=1}^{n} x_i y_i$$

A set $S \subseteq \mathbb{R}^n$ is a comprehensive iff:

$$y \leq x \text{ and } x \in S \text{ together imply that } y \in S.$$

That is, a comprehensive set describes a choice situation in which free disposal of any individual's utility is always possible.

A set $S \subseteq \mathbb{R}^n$ is convex iff:

$$x \in S \text{ and } y \in S \text{ and } 0 \leq \lambda \leq 1 \text{ together imply that}$$
$$\lambda x + (1-\lambda)y \in S,$$

where $\lambda x + (1-\lambda)y$ is the vector whose i^{th}-component is $\lambda x_i + (1-\lambda)y_i$. Since we are measuring utilities in von Neumann-Morgenstern scales, if the group can always plan to randomize between any two collective choice options, then the set of feasible expected utility allocations will be convex.

Given any finite collection of vectors $\{x^1,\ldots,x^k\} \subseteq \mathbb{R}^n$ (so that each $x^j = (x_1^j, x_2^j, \ldots, x_n^j)$ is a vector in $\mathbb{R}^n$), we define $H(x^1,\ldots,x^k)$ to be the smallest convex and comprehensive set containing the set $\{x^1,\ldots,x^k\}$. That is:

$$H(x^1,\ldots,x^k) = \left\{ y \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \text{there exist numbers } \lambda_1,\ldots,\lambda_k \text{ such that} \\ \text{every } \lambda_j \geq 0,\ \sum_{j=1}^{k} \lambda_j = 1, \text{ and} \\ y \leq \sum_{j=1}^{k} \lambda_j x^j \end{array} \right\}$$

We call $H(x^1,\ldots,x^k)$ the comprehensive-convex hull of $\{x^1,\ldots,x^k\}$.

In this paper, a choice problem is formally defined to be a nonempty, closed, convex, and comprehensive subset of $\mathbb{R}^n$, representing the set of feasible utility allocations. That is, we shall always assume that randomized strategies and free disposable are allowed in every choice problem.

Throughout, we let CP denote the set of choice problems to be studied. Also, we define CP^o to be the class of all choice problems which can be generated as the comprehensive-convex hulls of nonempty finite sets of allocations. That is:

$$CP^o = \{H(x^1,\ldots,x^k) \mid \{x^1,\ldots,x^k\} \text{ is a finite subset of } \mathbb{R}^n\}.$$

All the results which we derive will hold for the case of $CP = CP^o$, but we may allow CP to represent more general classes of convex comprehensive sets as well.

For any two sets $S \subseteq \mathbb{R}^n$ and $T \subseteq \mathbb{R}^n$ and any number λ, we define $\lambda S + (1-\lambda)T$ to be the set:

$$\lambda S + (1-\lambda)T = \{\lambda x + (1-\lambda)y \mid x \in S \text{ and } y \in T\}.$$

If $0 \leq \lambda \leq 1$, then we can interpret this set as follows. Suppose there is some random variable which may take the value 0, with probability λ, or may take the value 1, with probability $1-\lambda$. Suppose further that the individuals know that they will learn the random variable's true value tomorrow; if the value is 0 then the group will get a choice problem with feasible set S, and if the value is 1 then the feasible set will be T. Now suppose the group decides to plan its choices today, before learning the random variable. Then $\lambda S + (1-\lambda)T$ is the set of expected utility allocations which can be generated by making such conditional plans today.

Given any collection of choice problems CP, we define a <u>choice function</u> to be a mapping $F: CP \to \mathbb{R}^n$ such that, for every S in CP:

$$F(S) = (F_1(S),\ldots,F_n(S)) \in S.$$

That is, a choice function should select a feasible utility allocation vector for every choice problem. In the rest of this paper, we will study various properties which we might want a choice function to satisfy, and we will characterize the classes of choice functions which can satisfy these properties.

3. Linear and utilitarian choice functions

In this section, we consider choice functions satisfying weak Pareto-optimality and linearity properties.

A function $F:CP \to \mathbb{R}^n$ is weakly Pareto-optimal (WPO) iff:

(1) $F(S) \in S$, and

(2) $x > F(S)$ implies $x \notin S$,

for every S in CP and every x in $\mathbb{R}^n$. That is, a choice function is weakly Pareto-optimal if it always chooses a feasible utility allocation such that there is no other feasible allocation making everyone strictly better off. If a choice function were not weakly Pareto-optimal, then it could be criticized as inefficient, in those cases where it chose a strictly dominated point.

A function $F:CP \to \mathbb{R}^n$ is linear iff:

$$F(\lambda S + (1-\lambda)T) = \lambda F(S) + (1-\lambda)F(T)$$

for every pair of choice problems S and T in CP, and for every number λ such that $0 \leq \lambda \leq 1$ and $\lambda S + (1-\lambda)T \in CP$.

The linearity property has a natural interpretation in terms of timing of social choices. Suppose that a social decision will have to be made tomorrow, at which time the set of feasible utility allocations either will be S, with probability λ, or will be T, with probability $1-\lambda$. If F will be applied to the choice problem tomorrow, then the chosen utility allocation will either be F(S), with probability λ, or F(T), with probability $(1-\lambda)$. So the expected utility allocation (as assessed today) is $\lambda F(S) + (1-\lambda)F(T)$, if the social choices are to be made tomorrow. On the other hand, $\lambda S + (1-\lambda)T$ is the set of all expected utility allocations which are now feasible by planning tomorrow's decisions today (using contingency plans which may depend on the information to be learned tomorrow). If F is applied to the choice problem on the planning level today, then $F(\lambda S + (1-\lambda)T)$ should be the chosen utility allocation. Thus, a linear choice function is one for which every individual can expect the same utility from a social choice whether it is planned ahead today or made on a situational basis tomorrow.

A choice function $F:CP \to \mathbb{R}^n$ is <u>utilitarian</u> iff there exists some vector $p = (p_1,\ldots,p_n)$ in $\mathbb{R}^n$ such that:

(1) $\sum_{i=1}^{n} p_i = 1$ and every $p_i \geq 0$, and

(2) $p \cdot F(S) = \underset{x \in S}{\text{maximum}}\ p \cdot x$, for every $S \in CP$.

(Recall $p \cdot x = \sum_{i=1}^{n} p_i x_i$.) That is, a utilitarian choice function is one which always maximizes some weighted average of the individuals' utilities.

We say that CP is a <u>convex collection</u> of choice problems iff $\lambda S + (1-\lambda)T \in CP$ whenever $S \in CP$, $T \in CP$, and $0 \leq \lambda \leq 1$. With this regularity condition on the set of choice problems considered, we can state our first main result.

<u>Theorem 1</u>. Suppose CP is a convex collection of choice problems, and suppose $F:CP \to \mathbb{R}^n$ is a linear and weakly Pareto-optimal choice function. Then F is utilitarian.

<u>Proof</u>. Let $\Delta = \{p \in \mathbb{R}^n \mid \sum_{i=1}^{n} p_i = 1, \text{ all } p_i \geq 0\}$. For any S in CP, define the set

$$Q(S) = \{q \in \mathbb{R}^n \mid \text{for some } x \in S,\ q \cdot x > q \cdot F(S)\}.$$

Notice that $Q(S)$ must be an open subset of $\mathbb{R}^n$.

If F is not utilitarian, then we must have $\Delta \subseteq \bigcup_{S \in CP} Q(S)$, so the $Q(S)$ sets must form an open cover on Δ. Since Δ is compact, there must exist some finite collection

$$\{S^1,S^2,\ldots,S^k\} \subseteq CP \text{ such that } \Delta \subseteq \bigcup_{j=1}^{k} Q(S^j)$$

Then consider $S^o = \frac{1}{k}\sum_{j=1}^{k} S^j$. By WPO, $F(S^o)$ must be on the Pareto frontier of the closed convex set S^o, so by the Supporting Hyperplane Theorem (see Rockafellar [13], Section 11), there must exist some vector $p \in \Delta$ such that $p\cdot F(S^o) \geq p\cdot x$ for all x in S^o. However, since the $Q(S^j)$ cover Δ, we may assume (renumbering if necessary) that $p \in Q(S^1)$. So for some x^1 in S^1, $p\cdot x^1 > p\cdot F(S^1)$. Then by linearity, we get

$$p\cdot F(S^o) = p\cdot\left(\frac{1}{k}\sum_{j=1}^{k} F(S^j)\right) = \frac{1}{k}\left(\sum_{j=1}^{k} p\cdot F(S^j)\right) < \frac{1}{k}\left(p\cdot x^1 + \sum_{j=2}^{k} p\cdot F(S^j)\right)$$

$$= p\cdot\left(\frac{1}{k}\left(x^1 + \sum_{j=2}^{k} F(S^j)\right)\right).$$

But $\frac{1}{k}\left(x^1 + \sum_{j=2}^{k} F(S^j)\right) \in \frac{1}{k}\sum_{j=1}^{k} S^j = S^o$, so this inequality implies $p \in Q(S^o)$, a contradiction of the way p was constructed. To avoid this contradiction, we must conclude that F is utilitarian.

4. Colinear and concave choice functions

Let $E: CP \to \mathbb{R}^n$ be the choice function which always selects the highest feasible allocation giving all individuals equal utility. That is, E is weakly Pareto-optimal and satisfies:

$$E_1(S) = E_2(S) = \ldots = E_n(S)$$

for every choice problem S. Since all our choice problems are closed comprehensive sets, there will always be a unique point on the weakly Pareto-optimal frontier which satisfies this equity condition, so E is well-defined.

The choice function E is consistent with the maximin principle recommended by Rawls [12], in that:

$$\underset{i}{\text{minimum}}(E_i(S)) = \underset{x \in S}{\text{maximum}}(\underset{i}{\text{minimum}}(x_i)).$$

(If not, then there would exist some x in S such that $x_j > \min_i E_i(S) = E_j(S)$ for every j, which would contradict the Pareto-optimality of E(S).)

By generalizing the equity constraints which define E, we can introduce the class of colinear choice functions. We say that $F: CP \to \mathbb{R}^n$ is colinear iff F is weakly Pareto-optimal and there exist numbers $u_1, u_2, \ldots, u_n$ and $c_1 > 0$, $c_2 > 0, \ldots, c_n > 0$ such that, for every S in CP:

$$\frac{F_1(S) - u_1}{c_1} = \frac{F_2(S) - u_2}{c_2} = \ldots = \frac{F_n(S) - u_n}{c_n} .$$

That is, when $u = (u_1, \ldots, u_n)$ and $c = (c_1, \ldots, c_n)$, a colinear choice function F always selects the best point in S of the form

$$F(S) = u + \alpha c,$$

for some number α. Since we consider only closed and comprehensive sets, there will always be a unique weakly Pareto-optimal point in S on the line $\{u + \alpha c | \alpha \in \mathbb{R}\}$. So the colinear choice function is well-defined once the vectors u and $c > \underset{\sim}{0}$ are specified.

A colinear choice function like E is generally neither linear nor utilitarian. For example, (letting n=2) suppose that S = H((4,4),(0,10)) and T = H((4,4),(10,0)). Then E(S) = (4,4) and E(T) = (4,4). But $\frac{1}{2}S + \frac{1}{2}T$ = H((4,4),(7,2),(2,7),(5,5)). So we get:

$$E(\frac{1}{2}S + \frac{1}{2}T) = (5,5) \neq (4,4) = \frac{1}{2}E(S) + \frac{1}{2}E(T).$$

Although colinear choice functions do not satisfy linearity, they do satisfy a weaker property.

A choice function $F: CP \to \mathbb{R}^n$ is <u>concave</u> iff:

$$F(\lambda S + (1-\lambda)T) \geq \lambda F(S) + (1-\lambda)F(T)$$

for every pair of choice problems S and T in CP, and for every number λ such that $0 \leq \lambda \leq 1$ and $\lambda S + (1-\lambda)T \in CP$. (Notice that the formula above is a vector inequality in $\mathbb{R}^n$, meaning that the inequality holds in every component.)

In the last section we saw that a linear choice function is one for which every individual can expect the same utility from a social choice whether it is planned ahead "today" or made on a situational basis "tomorrow". In these terms, a concave choice function is one for which every individual's expected utility from planning ahead ($F_i(\lambda S + (1-\lambda)t)$) is always greater than or equal to his expected utility from situational judgments ($\lambda F_i(S) + (1-\lambda)F_i(T)$). So, when a concave choice function is used, the timing of social choices can make a difference; but timing would never be a cause for dispute, because all individuals would agree that earlier (planned-ahead) choices yield better expected outcomes.

For illustration, let us return to the numerical example given above, with S = H((4,4),(0,10)) and T = H((4,4),(10,0)). Suppose that a fair coin is about to be tossed: if it comes up Heads then the two individuals will be offered the choice problem S, and if it comes up tails, then they will be offered the choice problem T. Applying the choice function E after the coin toss will yield the equitable (and <u>ex post</u> Pareto-optimal) allocation (4,4), no matter how the coin may fall. But if the individuals plan their group choice before the coin is tossed, then the choice function E selects the equitable (and ex ante Pareto-optimal) allocation (5,5), which is implemented by planning to take (0,10) if Heads and (10,0) if Tails. Thus, we have the concavity relation:

$$E(\tfrac{1}{2}S + \tfrac{1}{2}T) = (5,5) \geq (4,4) = \tfrac{1}{2}E(S) + \tfrac{1}{2}E(T).$$

So before the coin is tossed, there is no dispute about whether to plan the group choices immediately or to wait until after the coin is tossed; both want to plan the group choices immediately.

Of course any linear utilitarian choice function must also be concave, since linearity implies concavity trivially. Furthermore, any colinear choice function must be concave. To prove this fact, suppose that:

$$F(S) = u + \alpha c, \quad F(T) = u + \beta c,$$

$$\text{and} \quad F(\lambda S + (1-\lambda)T) = u + \gamma c.$$

Since

$$\lambda F(S) + (1-\lambda)F(T) = u + (\lambda\alpha + (1-\lambda)\beta)c \in \lambda S + (1-\lambda)T,$$

and since $F(\lambda S + (1-\lambda)T)$ must be undominated in $\lambda S + (1-\lambda)T$, we must have:

$$\gamma \geq \lambda\alpha + (1-\lambda)\beta,$$

and so:

$$F(\lambda S + (1-\lambda)T) \geq \lambda F(S) + (1-\lambda)F(T).$$

There are concave choice functions which are neither utilitarian nor colinear. For example, let n=2, and let F' and F'' be linear utilitarian choice functions such that

$$F'(S) \text{ maximizes } \tfrac{1}{3}x_1 + \tfrac{2}{3}x_2 \text{ over } x \in S, \text{ and}$$

$$F''(S) \text{ maximizes } \tfrac{2}{3}x_1 + \tfrac{1}{3}x_2 \text{ over } x \in S.$$

Then let $F'''(S) = \frac{1}{2}F'(S) + \frac{1}{2}F''(S)$ (so that F''' is linear but not weakly

Pareto-optimal), and let $F(S) = (F_1(S), F_2(S))$ be the unique point on the weakly Pareto-optimal frontier of S such that:

$$F_1(S) - F_1'''(S) = F_2(S) - F_2'''(S).$$

It can be checked that the $F: CP \to \mathbb{R}^2$ thus defined is a concave and weakly Pareto-optimal choice function, but it is not utilitarian or colinear. However, this choice function may also seem rather complicated, in that it depends on $F'(S)$ and $F''(S)$, two alternatives which are not actually selected. That is, this choice function violates the principle of independence of irrelevant alternatives.

In general, a choice function $F: CP \to \mathbb{R}^n$ is <u>independent of irrelevant alternatives</u> (<u>IIA</u>) iff, for any two choice problems S and T in CP,

$$\text{if } S \subseteq T \text{ and } F(T) \in S \text{ then } F(S) = F(T).$$

To interpret this condition, suppose that society faces choice problem T and selects F(T). Now suppose that it is discovered that some points in T really are not feasible, and so the actual feasible set is $S \subseteq T$. Independence of irrelevant alternatives requires that, if the old choice is still feasible ($F(T) \in S$) then it should still be chosen. Any choice function determined by maximization of a social welfare function, or by a social preference ordering over utility allocations, will satisfy independence of irrelevant alternatives.

Recall that CP^o is the set of all choice problems which can be generated as comprehensive-convex hulls of finite sets of points in $\mathbb{R}^n$. If we impose the regularity condition that $CP \supseteq CP^o$, then we can state our main result for this section, as follows.

<u>Theorem 2</u>. Suppose that $CP \supseteq CP^o$, and suppose that $F: CP \to \mathbb{R}^n$ is weakly Pareto-optimal, concave, and independent of irrelevant alternatives. Then F is either utilitarian or colinear.

Proof of Theorem 2.

The proof will be broken up into a series of lemmas and definitions.

Definition. For any allocations x and y in $\mathbb{R}^n$, let $f(x) = F(H(x))$ and let $f(x,y) = F(H(x,y))$.

Definition. Let M be the set of all allocations in $\mathbb{R}^n$ which F could possibly select; that is:

$$M = \{x \mid x = F(S) \text{ for some } S \in CP\}.$$

Lemma 1. $M = \{x \mid x \in \mathbb{R}^n \text{ and } x = f(x)\}$.

Proof. If $x = f(x) = F(H(x))$ then $x \in M$, for $S = H(x)$. On the other hand, if $x = F(S)$ for any S, then $H(x) \subseteq S$ by WPO, and so $x = F(H(x))$ by IIA.

Lemma 2. M is convex.

Proof. Suppose $x = f(x)$, $y = f(y)$, $0 \leq \lambda \leq 1$, and $z = \lambda x + (1 - \lambda)y$. Then $H(z) = \lambda H(x) + (1 - \lambda)H(y)$, so $f(z) = F(H(z)) \geq \lambda F(H(x)) + (1 - \lambda)F(H(y))$

$$= \lambda x + (1 - \lambda)y = z.$$

So $f(z) \geq z$ by concavity. But $f(z) \leq z$ by WPO. So $f(z) = z$, and $z \in M$.

Definition. Let CP* be the set of all choice problems generated by utility allocations in $\mathbb{R}^n$. That is:

$$CP^* = \{H(x,y,\ldots,z) \mid x \in M,\ y \in M, \ldots, z \in M\}.$$

Lemma 3. If $S \in CP^*$ and $T \in CP^*$ and $0 < \lambda < 1$, then $F(\lambda S + (1 - \lambda)T) = \lambda F(S) + (1-\lambda)F(T)$

Proof. Let $x = F(S)$, $y = F(T)$. Choose $w \in S \cap M$ and $z \in S \cap M$ so that $\lambda w + (1-\lambda)z \geq F(\lambda S + (1-\lambda)T)$. (This can be done because, by convexity of M, $S \in CP^*$ must be generated from $S \cap M$ by disposal of utility only. Furthermore, by feasibility, there must be some $w^* \in S$ and $z^* \in T$ such that $\lambda w^* + (1-\lambda)z^* = F(\lambda S + (1-\lambda)T)$. Then choose $w \in S \cap M$ and $z \in T \cap M$ so that $w \geq w^*$ and $z \geq z^*$.) Then by IIA and convexity of M, we must have $\lambda w + (1-\lambda)z = F(\lambda S + (1-\lambda)T)$.

By IIA, we have $x = f(x,w)$, since $H(x,w) \subseteq S$. Also:

$$\lambda w + (1-\lambda)z = f(\lambda x + (1-\lambda)z,\ \lambda w + (1-\lambda)z)$$

by IIA from $\lambda S + (1-\lambda)T$. But $z \in M$ implies that $z = f(z)$. So by concavity

$$\lambda w + (1-\lambda)z \geq \lambda f(x,w) + (1-\lambda)f(z) = \lambda x + (1-\lambda)z.$$

Since $\lambda > 0$, we get $w \geq x$. So $x \in H(w)$ and $x = f(w)$ by IIA. But $w \in M$, so $x = w$.

A similar argument shows that $y = z$. Thus $\lambda x + (1-\lambda)y = F(\lambda S + (1-\lambda)T)$.

<u>Lemma 4</u>. There exists some vector $p \in \mathbb{R}^n$ such that $p_i \geq 0$ for all i, $\sum_{i=1}^{n} p_i = 1$, and:

$$\text{for all } S \in CP^*, \text{ and all } x \in S,\ p \cdot x \leq p \cdot F(S).$$

<u>Proof</u>. F is linear on CP* by Lemma 3. CP* is a convex collection of sets, because M is convex. So by Theorem 1, F is utilitarian on CP*.

<u>Definition</u>. Henceforth let p be the vector constructed in Lemma 4.

<u>Lemma 5</u>. For any choice problem $S \in CP$, for any $x \in S \cap M$, $p \cdot F(S) \geq p \cdot x$.

<u>Proof</u>. Let $y = F(S)$. Then $y \in M$ and $x \in M$, so $H(x,y) \in CP^*$. By IIA, $F(H(x,y)) = y$. So $p \cdot y \geq p \cdot x$ by Lemma 4.

<u>Lemma 6</u>. Suppose $p \cdot x > p \cdot f(y)$ and $x \in M$. Then $x = f(x,y)$. (Notice that, since $p \geq \underset{\sim}{0}$ and $y \geq f(y)$, the hypothesis of this lemma will hold if $p \cdot x > p \cdot y$ and $x \in M$.)

Proof. Let $z = f(x,y)$. For some λ such that $0 \leq \lambda \leq 1$, we have $z \leq \lambda y + (1-\lambda)x$. Our goal is to show that $\lambda = 0$.

If $\lambda = 1$ then we would have $z \leq y$, implying $z = f(y)$ by IIA. But Lemma 5 requires $p \cdot z \geq p \cdot x$, which would contradict the hypothesis if $z = f(y)$.

So we know $\lambda < 1$. By IIA we know $z = f(x, \lambda y + (1-\lambda)x)$. But by concavity:

$$f(x, \lambda y + (1-\lambda)x) \geq \lambda f(x,y) + (1-\lambda)f(x) = \lambda z + (1-\lambda)x.$$

So $z \geq \lambda z + (1-\lambda)x$, and thus $z \geq x$, since $1-\lambda > 0$. So $\lambda y + (1-\lambda)x \geq z \geq x$. This implies that either $\lambda = 0$ or $y \geq x$. But $y \geq x$ is not possible, by Lemma 5 and the hypothesis of this lemma. So $\lambda = 0$ and $z \leq x$. Therefore $z = x$.

Lemma 7. Suppose $x \in M$, $y \in M$, $\lambda \in \mathbb{R}$, and $z = \lambda x + (1-\lambda)y$. Then $p \cdot f(z) = p \cdot z$.

Proof. If $0 \leq \lambda \leq 1$, then this lemma follows trivially from Lemma 2.

Consider now the case $\lambda > 1$. Suppose that, contrary to the lemma, for some ϵ, $p \cdot z - p \cdot f(z) > \epsilon > 0$. Let $w = f(z) + \epsilon \underset{\sim}{1}$, where $\underset{\sim}{1} = (1,\ldots,1) \in \mathbb{R}^n$.

If $f(z,w)$ were in $H(z)$ then we would have $f(z,w) = f(z)$ by IIA. But this would violate WPO, since $w > f(z)$. So $z \not\geq f(z,w)$.

By construction, $p \cdot z > p \cdot w = p \cdot f(x) + \epsilon$. (Notice $p \cdot \underset{\sim}{1} = 1$.) Then by Lemma 6, $f(x, \frac{1}{\lambda}w + \frac{\lambda-1}{\lambda}y) = x$, since $x = \frac{1}{\lambda}z + \frac{\lambda-1}{\lambda}y$. (Remember $\lambda > 1$ is assumed here.) But by concavity, we must have:

$$x = f(x, \frac{1}{\lambda}w + \frac{\lambda-1}{\lambda}y) \geq \frac{1}{\lambda}f(z,w) + \frac{\lambda-1}{\lambda}f(y) = \frac{1}{\lambda}f(z,w) + \frac{\lambda-1}{\lambda}y.$$

Multiplying through by λ, we get $\lambda x + (1-\lambda)y \geq f(z,w)$, a contradiction of the conclusion in the last paragraph. Thus our assumption of $p \cdot z > p \cdot f(z)$ must be impossible, and so $p \cdot z = p \cdot f(z)$ if $\lambda > 1$.

For the case of $\lambda < 0$, the proof is similar to the $\lambda > 1$ case, reversing the roles of x and y, and of λ and $(1-\lambda)$.

Lemma 8. Suppose that F is not utilitarian. Then there exist allocations u and v in $\mathbb{R}^n$ such that $u = f(v)$ and $p\cdot v > p\cdot u$. Also, there exists a vector $c \in \mathbb{R}^n$ such that $c > \underset{\sim}{0}$, $p\cdot c = 1$, and $u - c \in M$.

Proof. If F is not utilitarian, then we can find some $S \in CP$ such that $\sup_{w \in S} p\cdot w > p\cdot F(S)$. So we can choose some $v \in S$ such that $p\cdot v > p\cdot F(S)$. Now $f(v) \leq v$ so $f(v) \in S \cap M$. By Lemma 5, $p\cdot F(S) \geq p\cdot f(v)$. So $p\cdot v > p\cdot u$, when we let $u = f(v)$.

To construct c, first let $d = u - f(u - \underset{\sim}{1})$. We know $u - d = f(u - \underset{\sim}{1}) \in M$. Also, $u - \underset{\sim}{1} \geq f(u - \underset{\sim}{1})$, so $d \geq \underset{\sim}{1} > \underset{\sim}{0}$ and $p\cdot d > p\cdot \underset{\sim}{1} = 1$. Let $c = (\frac{1}{p\cdot d})d$. So $p\cdot c = 1$, $c > \underset{\sim}{0}$, and $u - c = (\frac{1}{p\cdot d})(u - c) + (\frac{p\cdot d - 1}{p\cdot d})u$ is in M (by convexity of M).

Assumption. In Lemmas 9 through 12 we will assume that F is not utilitarian. We will prove that F must be colinear.

Definition. Henceforth, under the assumption that F is not utilitarian, let u,v, and c be as in Lemma 8.

Lemma 9. Suppose $x \in M$ and $p\cdot v > p\cdot x > p\cdot u$. Then $x \geq u$.

Proof. Let $\alpha = \frac{p\cdot v - p\cdot u + 1}{p\cdot v - p\cdot x}$, and let $y = v + \alpha(x-v)$. Notice that $\alpha > 1$ and $p\cdot y = p\cdot u - 1 < p\cdot u$. Since $u \in M$, by Lemma 6, $f(u,y) = u$.

Let $w = \frac{1}{\alpha}u + \frac{\alpha-1}{\alpha}v$. Then $w \leq v$, since $u \lesssim v$ and $\alpha > 1$. By Lemma 6, $f(v,x) = x$, so by IIA we get $f(w,x) = x$. But $H(w,x) = \frac{1}{\alpha}H(u,y) + \frac{\alpha-1}{\alpha}H(v)$. So by concavity of F:

$$x = f(w,x) \geq \frac{1}{\alpha}f(u,y) + \frac{\alpha-1}{\alpha}f(v) = \frac{1}{\alpha}u + \frac{\alpha-1}{\alpha}u = u.$$

Lemma 10. Suppose $x \in M$ and $p\cdot x = p\cdot u$. Then $x \geq u$.

Proof. Suppose not. Then, for some i, $x_i < u_i$. Let $y = f(u+c)$. Since $u \in M$ and $u - c \in M$, Lemma 7 implies $p\cdot y = p\cdot(u+c) = p\cdot u + 1$. But $y \in M$, so for any small positive ϵ $(0 < \epsilon < 1)$ we have $(1-\epsilon)x + \epsilon y \in M$. If we choose ϵ small enough, we get

$$(1-\epsilon)x_i + \epsilon y_i < u_i$$

so $(1-\epsilon)x + \epsilon y \not\geq u$. But $p\cdot v > p\cdot((1-\epsilon)x + \epsilon y) = p\cdot u + \epsilon > p\cdot u$ as long as ϵ is small and positive. So $(1-\epsilon)x + \epsilon y$ will violate Lemma 9 if $x \not\geq u$. So we must have $x \geq u$.

Lemma 11. For any vector d in $\mathbb{R}^n$, if $u - d \in M$ then, for every λ in $\mathbb{R}$, $u + \lambda d \in M$.

Proof. Consider first the case of $\lambda > 0$. Let $y = f(u + \lambda d)$. Then $p\cdot y = p\cdot(u+\lambda d)$ by Lemma 7. So $\frac{1}{1+\lambda}y + \frac{\lambda}{1+\lambda}(u-d) \in M$ (by convexity of M) and $p\cdot(\frac{1}{1+\lambda}y + \frac{\lambda}{1+\lambda}(u-d)) = p\cdot u$. By Lemma 10, we conclude $\frac{1}{1+\lambda}y + \frac{\lambda}{1+\lambda}(u-d) \geq u$. So $y \geq (1+\lambda)u - \lambda(u-d) = u + \lambda d$. But $y = f(u+\lambda d)$ implies $y \leq u + \lambda d$. So $u + \lambda d = y \in M$.

For the case of $\lambda < 0$, observe that the preceding case showed $u - (-d) = u + d \in M$. So, using our results for positive multipliers, we get $u + \lambda d = u + (-\lambda)(-d) \in M$ since $-\lambda > 0$.

Lemma 12. Suppose $x \in M$. Then $x = u + \beta c$, where $\beta = p\cdot x - p\cdot u$.

Proof. By Lemma 11 and the definition of c, $u - \beta c \in M$. So $\frac{1}{2}x + \frac{1}{2}(u - \beta c) \in M$. But $p\cdot(\frac{1}{2}x + \frac{1}{2}(u-\beta c)) = \frac{1}{2}p\cdot x + \frac{1}{2}p\cdot u - \frac{1}{2}\beta\, p\cdot c = \frac{1}{2}(p\cdot x + p\cdot u - \beta) = p\cdot u$. So $\frac{1}{2}x + \frac{1}{2}(u - \beta c) \geq u$ by Lemma 10. Solving for x, we get $x \geq u + \beta c$.

Now let $y = u + (u-x)$. By Lemma 11 we know $y \in M$ and also $u + \beta c \in M$. So $\frac{1}{2}y + \frac{1}{2}(u+\beta c) \in M$. Observe that $p\cdot(\frac{1}{2}y + \frac{1}{2}(u+\beta c)) = \frac{1}{2}p\cdot u - \frac{1}{2}\beta + \frac{1}{2}p\cdot u + \frac{1}{2}\beta = p\cdot u$. So $\frac{1}{2}y + \frac{1}{2}(u+\beta c) \geq u$ by Lemma 10. Thus $y \geq u - \beta c$, and so $x = 2u - y \leq u + \beta c$. Together with $x \geq u + \beta c$ from the preceding paragraph, this proves the lemma.

Since Lemma 12, stating that F is colinear, was proven under the assumption that F is not utilitarian, we have proven the Theorem.

5. Scale invariant choice functions and dictatorships.

Both the colinear and the utilitarian choice functions involve interpersonal comparisons of utility. The equity constraints of the colinear choice functions compare utility levels of different individuals, to determine which allocations are permissible. The utilitarian functional $p \cdot x$ makes interpersonal comparisons implicitly, to determine the tradeoff between the utilities of different individuals. Both of these comparisons depend on the scales in which the individuals' utilities are measured. (For example, suppose two individuals have utility which is linear in money, and they plan to use a utilitarian choice function maximizing the sum of the individuals' utility numbers. The choices selected by this rule would change dramatically if one individual switched his utility scale from dollars to lira while the other individual stayed with dollar units.) One might now ask, if we drop the concavity requirement from Theorem 2, can we then find choice functions which are not so dependent on the way utility is measured? In this section we will show that, even without concavity, we still cannot get scale-invariant choice functions satisfying WPO and IIA, unless we give one individual's preferences exclusive priority over all the others'.

A <u>transformation of the utility scales</u> is a mapping $L:\mathbb{R}^n \to \mathbb{R}^n$ of the form

$$L(x) = (\alpha_1 x_1 + \beta_1, \alpha_2 x_2 + \beta_2, \ldots, \alpha_n x_n + \beta_n)$$

for some numbers $\alpha_1 > 0, \alpha_2 > 0, \ldots, \alpha_n > 0$, and $\beta_1, \beta_2, \ldots, \beta_n$. Decision theory defines the individuals' von Neumann-Morgenstern utility scales only up to increasing linear transformations. So two choice problems which differ by such a transformation of the utility scales could be interpreted as representing the same underlying social choice situation, if our utility scales have decision-theoretic significance only. If we want to avoid making interpersonal comparisons of utility which cannot be based on decision-theoretic considerations, then our choice function should be invariant under transformations of the utility scales.

A choice function $F:CP \to \mathbb{R}^n$ is thus said to be scale invariant iff:

$$F(\{L(x)|x \in S\}) = L(F(S))$$

for every choice problem S in CP and every transformation L of the utility scales such that $\{L(x)|x \in S\} \in CP$. That is, if we transform the utility scales so that the choice problem S becomes $\{L(x)|x \in S\}$, then the new solution $F(\{L(x)|x \in S\})$ should be the old solution F(S) translated into the new utility scales, if F is scale invariant.

The Nash bargaining solution [10] satisfies the properties of weak Pareto-optimality, independence of irrelevant alternatives, and scale invariance. But Nash's solution is for bargaining problems, which differ from our choice problems in that a bargaining problem has one feasible allocation singled out as the reference point, representing the status quo which would prevail if no other feasible allocation were agreed upon. Without such a reference point, however, the Nash bargaining solution cannot be defined, and the only choice functions which can satisfy these three properties are dictatorships.

We say that an individual i is a dictator for a choice function $F:CP \to \mathbb{R}^n$ iff

$$F_i(S) = \underset{x \in S}{\text{maximum}}(x_i)$$

for every choice problem S in CP. So if i is a dictator for F, then F always selects a feasible allocation which maximizes i's utility. That is, a choice function with a dictator resolves social conflict by giving the dictator everything he wants.

Theorem 3. Suppose $CP \supseteq CP^o$, and let $F:CP \to \mathbb{R}^n$ be a choice function which is scale invariant and independent of irrelevant alternatives. Then there must be some individual i who is a dictator for F.

Proof. We show first that $F(H(w)) = w$, for every vector $w \in \mathbb{R}^n$. If not, then there exists some individual k and some number δ such that

$$w_k - F_k(H(w)) = \delta > 0.$$

(We know that $F(H(w)) \leq w$, since the choice must be feasible.) Let $v \in \mathbb{R}^n$ satisfy:

$$v_j = \begin{cases} w_j & \text{if } j \neq i, \\ w_k - \delta & \text{if } j = k. \end{cases}$$

By scale invariance (translating k's scale by δ) we must have $v_k - F_k(H(v)) = w_k - F_k(H(w)) = \delta$, but independence of irrelevant alternatives implies $F(H(v)) = F(H(w))$. So we get $\delta = w_k - v_k = 0$, a contradiction. Thus $F(H(w)) = w$ must hold.

Now consider $D = H(d^1, d^2, \ldots, d^n)$, where each $d^j = (d_1^j, \ldots, d_n^j)$ is the unit vector in $\mathbb{R}^n$ such that

$$d_k^j = \begin{cases} 0 & \text{if} \quad k \neq j, \\ 1 & \text{if} \quad k = j. \end{cases}$$

Let $u = F(D)$. We will show that $u = d^i$ for some individual i. From the preceding paragraph, we can conclude that u is strongly Pareto-optimal in D, so that $\sum_{j=1}^{n} u_j = 1$ and all $u_j \geq 0$. (If not, then there would be some vector w in D such that $w \geq u$, $w \in D$, and $w \neq u$. Then, by IIA, we would get $F(H(w)) = F(D) = u \neq w$, which contradicts the preceding paragraph.) Thus we can select some individual i so that $u_i > 0$. Consider the set $C = H(c^1, c^2, \ldots, c^n)$, where each $c^j = (c_1^j, \ldots, c_n^j)$ is a vector satisfying:

$$c_k^j = \begin{cases} u_k & \text{if } j \neq i \text{ and } j \neq k \neq i \\ 1 & \text{if } j \neq i \text{ and } k = j \\ 0 & \text{if } j \neq i \text{ and } k = i \end{cases}$$

and $c^i = u$.

By independence of irrelevant alternatives, $F(C) = u = c^i$, since $C \subseteq D$. But then scale invariance implies that $F(D) = d^i$. (Use $L_i(x) = \frac{1}{u_i} x_i$ and $L_k(x) = \frac{x_k - u_k}{1 - u_k}$ for $k \neq i$, and check that $L(c^j) = d^j$ for all j.)

We now show that this individual i is a dictator for F. If not, find some choice problem S and some vector y such that $y \in S$ and $y_i > F_i(S)$. Let $z = F(S)$. Since S is comprehensive, we may assume that $y_k < z_k$ for all $k \neq i$. Independence of irrelevant alternatives implies that $z = F(H(y,z))$. Consider the vector $b = (\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}) \in \mathbb{R}^n$. If we let

$$\tilde{L}_i(x) = \frac{(n-1)x_i + y_i - nz_i}{n(y_i - z_i)}$$

$$\text{and} \quad \tilde{L}_k(x) = \frac{x_k - y_k}{n(z_k - y_k)} \quad \text{for all } k \neq i$$

then scale invariance implies $b = F(H(d^k,b))$, because $L(y) = d^i$ and $L(z) = b$. But $b \in D$, so by independence of irrelevant alternatives we must have $d^i = F(H(d^i,b))$. This contradiction proves that no vector y in S can be found to satisfy $y_i > F_i(S)$. Thus individual i must be a dictator for F.

REFERENCES

[1] Arrow, K.J.: Social Choice and Individual Values, 2nd ed., New York: Wiley, 1963.

[2] Bentham, J.: An Introduction to the Principles of Morals and Legislation, London: Athlone Press, 1970.

[3] Deschamps, R., and L. Gevers: "Separability, Risk-bearing, and Social Welfare Judgments", European Economic Review, 10 (1977), 77-94.

[4] Deschamps, R. and L. Gevers: "Leximin and Utilitarian Rules: A Joint Characterization", Journal of Economic Theory, forthcoming.

[5] Harsanyi, J.C.: "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", Journal of Political Economy, 63(1955), 309-321.

[6] Harsanyi, J.C.: "Nonlinear Social Welfare Functions", Theory and Decision, 6(1975), 311-332.

[7] Kalai, E.: "Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons", Econometrica, 45(1977), 1623-1630.

[8] Kalai, E. and D. Schmeidler: "Aggregation Procedure for Cardinal Preference: A Formulation and Proof of Samuelson's Impossibility Conjecture", Econometrica, 45(1977), 1431-1438.

[9] Myerson, R.B.: "Two-Person Bargaining Problems and Comparable Utility", Econometrica, 45(1977), 1631-1637.

[10] Nash, J.: "The Bargaining Problem ", Econometrica, 18(1950), 155-162.

[11] Rawls, J.: A Theory of Justice, Cambridge: Harvard University Press, 1971.

[12] Rawls, J.: "Some Reasons for the Maximin Criterion", American Economic Review, 64(1974), 141-146.

[13] Rockafellar, R.T.: Convex Analysis, Princeton: Princeton U. Press, 1970.

[14] Sen, A.K.: Collective Choice and Social Welfare, San Francisco: Holden-Day, 1970.

[15] Shapley, L.S.: "Utility Comparison and the Theory of Games" in La Decision, Paris: Edition du Centre de la Recherche Scientifique, France, 1969, pp. 251-263.