

DISCUSSION PAPER NO. 352

IMPOSSIBILITY RESULTS IN THE AXIOMATIC THEORY
OF INTERTEMPORAL CHOICE

by

Edward Packel, Lake Forest College

Revised December 1978

IMPOSSIBILITY RESULTS IN THE AXIOMATIC THEORY
OF INTERTEMPORAL CHOICE*

by

Edward Packel, Lake Forest College

I. Introduction

Problems posed by questions of intertemporal choice are at once intriguing, elusive, and imminent. Especially when we try to cope with consequences of scarce and nonrenewable resources, we see that decisions must be made in the present which will have profound and often irreversible effects upon future generations. Clearly the present generation can never have sufficient wisdom to act omnisciently in such matters, so we must ask more modestly whether society can act wisely, prudently, and fairly under the assumption that resource constraints upon and preferences of future generations are known.

A commonly used approach in intertemporal decisionmaking is the discounting of future benefits by means of an appropriately chosen "discount rate." Thus, a multiplicative factor r , less than but hopefully appropriately close to unity, is somehow chosen; and successive generations beyond the present have their utilities discounted by increasing powers of r . Can such an r be chosen to be equitable in some sense? Can we, in this or any other fashion, posit intertemporal choice procedures which are "fair" to present and future generations?

In a foundational paper, Ferejohn and Page [1] formulate the problem of intertemporal choice from an axiomatic perspective. This is done in Arrowian fashion by considering the problem of aggregating the preference orderings

*The author wishes to thank Ken Shepsle and an anonymous referee for their helpful comments.

R_1 of an infinite sequence of future generations into an intertemporal societal preference R . Using this approach and results akin to Arrow's impossibility theorem, it is shown in [1] that no aggregation procedure can simultaneously satisfy standard conditions of Pareto optimality, independence, nondictatorship, and a natural condition of stationarity introduced by Koopmans in [2].

In this paper we introduce a generalized formulation of intertemporal choice under which we strengthen the impossibility result of [1] by showing that the Pareto condition and stationarity are by themselves inconsistent. We then show that stationary aggregation procedures satisfying the Pareto and independence conditions are possible if the domain of generational preferences is restricted; but that, even under such restricted domain, the societal preference must be dictatorial. We then show how these results reinforce the case made in [1] against indiscriminate use of discount rates in problems of intertemporal aggregation.

II. Axiomatic Formulation of Intertemporal Choice

We assume an infinite sequence of generations i , each of which has a set X_i of "instantaneous" states of the economy. The decision to impose no limit on the number of future generations stems both from the inability to decide just where an arbitrary cutoff might be and from the fact that more general and interesting results can be obtained without such a limitation. The set of present and future states of the world or programs is a subset X of the Cartesian product $X_1 \times X_2 \times \dots$. A natural special case would be to let all X_i be identical and to take X as the full Cartesian product X^{∞} of these identical factors. More generally, the program space X might satisfy a variety of finitistic and presumably complex constraints. In order to lay bare the assumptions needed for our main results and to avoid the seemingly impossible task of formulating such constraints, we restrict ourselves to program spaces X satisfying the following conditions:

$$X1: (x_1, x_2, x_3, \dots) \in X \Rightarrow (x_2, x_3, \dots) \in X.$$

$$X2: \exists \text{ distinct } a \text{ and } b \text{ such that } x_i = a \text{ or } b \quad \forall i \Rightarrow (x_1, x_2, \dots) \in X.$$

The first condition requires that if a certain overall intergenerational program is possible, then moving the program up one generation is also possible. A by-product of this assumption in most reasonable situations is that $i < j \Rightarrow X_i \supseteq X_j$. The possibility that the sets X_i may shrink as i increases is consistent with the fact that we are formulating a theory of nonrenewable resources and possibly irreversible decisions.

Condition X2 simply says that some pair of states a and b will be common to all generations and that any program involving a 's and b 's must be admissible. This is not an unreasonable assumption in that there will certainly be some possible states (presumably involving substantial use of only renewable resources) which can be achieved through all generations. Note that the specific program space X^∞ mentioned earlier easily satisfies conditions X1 and X2, which we assume to hold throughout the remainder of the paper.

Let \mathcal{R} denote the set of complete, transitive binary relations on X . With each generation i we associate a preference ordering $R_i \in \mathcal{R}$. Thus, R_i reflects generation i 's views about how states of the world across all generations should be ordered. We avoid the delicate issue of how and to what extent the preferences of one generation might be affected by those of another by working only within a final profile $\pi = (R_1, R_2, \dots)$ which results from each generation's ultimate ordering decision. Let $\mathcal{R}^\infty = \mathcal{R} \times \mathcal{R} \times \dots$ denote the set of all possible profiles of generational preferences and let \mathcal{B} denote the set of complete (not necessarily transitive) binary relations on X .

We can now formulate the problem of intertemporal choice as the attempt to find functions $F : \mathcal{R}^\infty \rightarrow \mathcal{B}$ that satisfy appropriate social choice and equity conditions. Let P_i denote the strict preference associated with R_i ($xP_i y \Leftrightarrow \sim yR_i x$) and let P denote the strict aggregate preference associated with $R = F(\pi)$ for a fixed profile $\pi = (R_1, R_2, \dots)$. We define the conditions needed on the choice function F as follows:

Pareto: $\forall x, y \in X, x P_i y \forall i \Rightarrow x P y$.

Independence: $\forall x, y \in X$, if $x R_i y \Leftrightarrow x R'_i y \forall i$, then $x R y \Leftrightarrow x R' y$.

Stationarity: $\forall (x_1, x_2, x_3, \dots)$ and $(x_1, y_2, y_3, \dots) \in X, \forall \pi \in \mathcal{R}^\infty$,
 $(x_1, x_2, x_3, \dots) R (x_1, y_2, y_3, \dots) \Leftrightarrow (x_2, x_3, \dots) R (y_2, y_3, \dots)$.

The first two conditions on F are standard Arrow-type axioms. Little need be said in "defense" or explanation of the Pareto condition. Independence in our context says that the intertemporal ordering of two programs should depend only on the pairwise generational orderings between them and not on the relative generational orderings of other programs. The standard and cogent arguments in support of independence in intratemporal social choice theory apply just as well to our setting (see, for instance, Plott [3]).

Stationarity, introduced by Koopmans [2], requires invariance of F under forward generational program shifts, provided that the unshifted programs belong to X and agree in their initial components. Thus, whatever ordering exists between two programs with identical first generation states must exist between the corresponding accelerated programs which move each state forward one generation and conversely. From an abstract standpoint, this intergenerational shift "invariance" seems reasonable as a principle of intertemporal equity. Some thought about specifics, however, suggests scenarios where matters

of timing are important in such a way that stationarity of a choice function might not be appropriate. In fact, Diamond [4] shows in a somewhat different setting that preferences which satisfy stationarity and two similar preference axioms fail to provide "equal treatment for all generations." The main point is that stationarity, plausible though not perhaps compelling as it may be, will be seen to be a consequence of employing a discount rate in a fairly general intertemporal choice setting. Thus, the impossibility results which follow have an interesting and meaningful axiomatic base.

III. Impossibility Results

Ferejohn and Page [1] have shown that any choice function $F: \mathcal{R}^\infty \rightarrow \mathcal{R}$ satisfying Pareto, independence, and stationarity must be a dictatorship of the first or present generation. Thus, such an F has the property that $\forall x, y \in X, xP_1y \Rightarrow xPy$. We establish in what follows the stronger result that, even without requiring transitivity in the range of F , Pareto and stationarity are inconsistent. We then show that restricting the domain of F allows for transitive choice functions which are Pareto, independent, and stationary. By modifying the arguments of [1], we then prove that any such functions must result in first generation dictatorships.

Given $x = (x_1, x_2, \dots) \in X$, we denote the tail of the program x starting with the k^{th} generation by ${}_kx = (x_k, x_{k+1}, \dots)$. We also write x as $(x_1, x_2, \dots, x_{k-1}, {}_kx)$ when convenient. Stationarity of R can then be restated as follows:

$$\forall (x_1, {}_2x) \text{ and } (x_1, {}_2y) \in X, (x_1, {}_2x) R (x_1, {}_2y) \Leftrightarrow {}_2x R {}_2y.$$

We let $\tilde{\mathcal{R}}$ denote the subset of \mathcal{R} consisting of the stationary transitive relations. Finally, we call $R \in \mathcal{R}$ weakly stationary if

$$\forall (x_1, {}_2x) \text{ and } (x_1, {}_2y) \in X, \begin{cases} (x_1, {}_2x) P (x_1, {}_2y) \Rightarrow {}_2x R {}_2y \\ \text{and} \\ {}_2x P {}_2y \Rightarrow (x_1, {}_2x) R (x_1, {}_2y). \end{cases}$$

It is immediate that stationarity implies weak stationarity. It will be important in what follows to note that there exist members of \mathcal{R} that are not weakly stationary. Specifically, let P^1 and P^3 be strict total orderings of X_1 and X_3 respectively with $b P^1 a$ and $a P^3 b$ (recall assumption X2). Define

$$(x_1, x_2, x_3, \dots) \bar{P}(y_1, y_2, y_3, \dots) \Leftrightarrow x_1 P^1 y_1 \quad \text{or} \quad x_1 = y_1 \quad \text{and} \quad x_3 P^3 y_3 ;$$

and $\bar{R} \in \mathcal{R}$ by $x \bar{R} y \Leftrightarrow \sim y \bar{P} x$. If we apply assumption X2 to

$x = (a, a, b, a, a, \dots)$ and $y = (a, b, a, a, a, \dots)$, it can be checked that neither implication in the definition of weak stationarity holds for \bar{R} .

We do not try to place added significance on the idea of weak stationarity.

Its main role is as a tool in the lemma and theorem to follow.

Lemma 1. Given $F: \mathcal{R}^\infty \rightarrow \mathcal{B}$ satisfying Pareto. If $\pi = (\bar{R}, \bar{R}, \dots)$ is a constant profile in \mathcal{R}^∞ and $R = F(\pi)$ is stationary, then \bar{R} must be weakly stationary.

Proof. Given $(x_1, 2^x) \bar{P}(x_1, 2^y)$, the Pareto condition gives $(x_1, 2^x) P(x_1, 2^y)$. Suppose that $2^x \bar{R} 2^y$ fails to hold. Then $2^y \bar{P} 2^x$ and Pareto yields $2^y P 2^x$. Stationarity of R then gives $(x_1, 2^y) P(x_1, 2^x)$. This contradicts $(x_1, 2^x) P(x_1, 2^y)$ as deduced earlier and establishes the first part of the definition of weak stationarity. The second part of the definition follows by an entirely analogous argument.

Theorem 1. There does not exist a function $F: \mathcal{R}^\infty \rightarrow \mathcal{B}$ satisfying Pareto and stationarity.

Proof. Choose any $\bar{R} \in \mathcal{R}$ that is not weakly stationary and consider the constant profile $\pi = (\bar{R}, \bar{R}, \dots)$. By Lemma 1 it follows that $R = F(\pi)$ cannot be a stationary relation and hence F cannot be stationary.

The impossibility result of Theorem 1 arises by considering constant profiles whose components are not weakly stationary. The next theorem shows that restricting the domain permits the existence of choice functions with some desired properties.

Theorem 2. There exists a function $F: \tilde{R} \times R^\infty \rightarrow \tilde{R}$ satisfying Pareto and independence.

Proof. Given $\pi = (R_1, R_2, \dots)$, define $F(\pi) = R_1$. It is immediate that F is Pareto, independent, and stationary.

The choice function used in Theorem 2 gives the present generation full dictatorial power to impose its hopefully altruistic view of current and future possible states of the world, independent of the preferences of future generations. The final theorem shows that, even with a domain collapsed to \tilde{R}^∞ , a first generation dictatorship must result if the specified social choice conditions are required. To prove this result we utilize the approach of Hansson [5] to Arrow's impossibility theorem as outlined below.

Given a social choice function F and a set of voters (in our case generations) N , the collection of decisive sets with respect to F is defined by

$$W_F = \{C \subseteq N \mid xP_i y \ \forall i \in C \Rightarrow xPy\}.$$

The "decisive set" formulation of Arrow's theorem says that, if $F: \mathcal{R}^N \rightarrow \mathcal{R}$ satisfies Pareto and independence, then the collection W_F must be an ultrafilter as defined by properties U1 to U4 below:

$$U1: \quad \emptyset \notin W_F \text{ and } N \in W_F$$

$$U2: \quad K, L \in W_F \Rightarrow K \cap L \in W_F$$

$$U3: \quad K \in W_F \text{ and } K \subseteq L \Rightarrow L \in W_F$$

$$U4: \quad \forall K \subseteq N, \text{ either } K \in W_F \text{ or } N \setminus K \in W_F.$$

We first prove a lemma which enables us to extend the conclusion that W_F must be an ultrafilter to the situation where $N = \{1, 2, \dots\}$ and $F: \tilde{\mathcal{R}}^\infty \rightarrow \tilde{\mathcal{R}}$. We then use stationarity to prove that the ultrafilter W_F must have the form $W_F = \{C \subseteq N \mid 1 \in C\}$, from which it follows directly that the first generation is a dictator. It should be noted that if N is assumed to be finite rather than infinite, a dictatorship of some generation must result without stationarity (Arrow's theorem), though not necessarily a first generation dictatorship.

Lemma 2. If $F: \tilde{\mathcal{R}}^\infty \rightarrow \tilde{\mathcal{R}}$ satisfies Pareto and independence, then W_F is an ultrafilter.

Proof. We look at the crucial steps of the proof given in [5] and we show that restriction to profiles whose components are stationary still permits

these steps to be carried out. For the proof of ultrafilter property U2, assume $K, L \in W_F$ and consider any profile $\pi \in \tilde{\mathcal{R}}^\infty$ for which $xP_i y \forall i \in K \cap L$. By choosing $z = (z_1, z_2, \dots) \in X$ such that $\forall t \quad t^z \neq t^x$ and $t^z \neq t^y$ (possibly by assumption X2), we can choose a profile $\pi' \in \tilde{\mathcal{R}}^\infty$ which agrees with π on its restriction to x, y and has

$$xP'_i z \quad \text{and} \quad yP'_i z \quad \forall i \in K \setminus L$$

$$xP'_i z \quad \text{and} \quad zP'_i y \quad \forall i \in K \cap L$$

$$zP'_i x \quad \text{and} \quad zP'_i y \quad \forall i \in L \setminus K.$$

The components of π' can be chosen to be stationary since they are already stationary as regards x and y ; and stationarity involving z (and any other members of X) can be imposed (because $t^z \neq t^x$ and $t^z \neq t^y$) without contradiction. Use of $K, L \in W_F$ and transitivity now gives $xP'_i y$, and independence then gives xPy , so $K \cap L \in W_F$. A second crucial step occurs in proving property U4, but again the ideas used above generate the desired stationary profiles with no difficulty. From this point Hansson's proof goes through unchanged, though we omit the details.

Theorem 3. If $F: \tilde{\mathcal{R}}^\infty \rightarrow \tilde{\mathcal{R}}$ satisfies Pareto and independence, then $W_F = \{C \subseteq N \mid 1 \in C\}$ and hence the present generation must be dictatorial.

Proof. We build upon ideas developed in [1] and we freely use that fact that W_F is an ultrafilter by Lemma 2. Using assumption X2, we first pick distinct $a, b \in X_1$ such that all programs with components only

involving a and b belong to X . For each generation i , let P^i be a strict ordering of X_i with $aP^i b$. We define the profile $\pi = (R_1, R_2, \dots) \in \tilde{\mathcal{R}}^\infty$ as follows. For any $x, y \in X$, let j be the earliest generation for which $x_j \neq y_j$. The strict part P_i of R_i is then defined by

$$xP_i y \Leftrightarrow x_{j+i-1} P^{j+i-1} y_{j+i-1}.$$

When $j = 1$, each P_i is defined by the action of P^i on the j^{th} component; and increases in j push the action of P_i out in the sequence by a corresponding amount. It is readily checked that π , by its definition, is in $\tilde{\mathcal{R}}^\infty$. Now consider the specific programs

$$u = (a, b, a, b, a, b, \dots)$$

$$v = (b, a, b, a, b, a, \dots)$$

$$w = (a, a, b, a, b, a, b, \dots)$$

$$z = (b, b, a, b, a, b, a, \dots).$$

Let $C = \{i \mid uP_i v\} = \{1, 3, 5, 7, \dots\}$ (recall that $aP^i b$ for all i). Also, let $D = \{i \mid vP_i u\} = \{2, 4, 6, 8, \dots\}$. By ultrafilter property U4, either $C \in W_F$ or $D \in W_F$. Supposing $D \in W_F$, we would have vPu and thus, by stationarity of F , uPw and zPv . By transitivity, this gives zPw . However, since $\{i \mid wP_i z\} \supseteq D$, we obtain wPz also. Thus $D \in W_F$ leads to a contradiction and we conclude that $C \in W_F$ and uPv . Applying stationarity and uPv , we obtain wPu and vPz . Transitivity then gives wPz . Letting $E = \{i \mid wP_i z\} = \{1, 2, 4, 6, 8, \dots\}$, we must have $E \in W_F$ (otherwise U4 would imply $N \setminus E \in W_F$, giving zPw). From U2 it follows that $C \cap E = \{1\} \in W_F$. Finally, from U3, we obtain the full conclusion of the theorem.

IV. Stationarity and the Discount Rate

A transitive preference ordering R on a set Y of alternatives is said to be represented by a utility function $u:Y \rightarrow \mathbb{R}$ if $\forall x,y \in Y$, $xRy \Leftrightarrow u(x) \geq u(y)$. Given a set of programs $X \subseteq X_1 \times X_2 \times \dots$, let $Y = \bigcup_{i=1}^{\infty} X_i$. We say that $R \in \mathcal{R}$ has a discount rate representation if there exists $u:Y \rightarrow \mathbb{R}$ and $r \in (0,1)$ such that $xRy \Leftrightarrow U^r(x) \geq U^r(y)$, where the intergenerational "utility" U^r is defined on $x = (x_1, x_2, \dots) \in X$ by $U^r(x) = \sum_{i=1}^{\infty} r^{i-1} u(x_i)$. The role of the increasing powers r^{i-1} is simultaneously to discount the preferences of generation i and to ensure a convergent series. We assume implicitly that $U^r(x)$ does converge for all $x \in X$. This convergence could be guaranteed by requiring u to be bounded over Y , but we do not discuss this matter further.

An intertemporal choice function $F: \mathcal{R}^{\infty} \rightarrow \mathcal{R}$ is defined to be a discounting choice function if to each profile $\pi \in \mathcal{R}^{\infty}$ there corresponds a function $u:Y \rightarrow \mathbb{R}$ and a rate $r \in (0,1)$ such that $x F(\pi) y \Leftrightarrow U^r(x) \geq U^r(y)$, with U^r defined as above. The profile dependent function u can be thought of as some sort of societal utility function on the set of all possible instantaneous states of the economy. Ideally, u might reflect each generation's preferences over the more broadly viewed programs (this might somehow be done by using stationarity), but our results do not depend on how this is done or whether it is indeed possible.

Theorem 4 If $F: \mathcal{R}^{\infty} \rightarrow \mathcal{R}$ is a discounting choice function, then $R = F(\pi)$ is stationary for all $\pi \in \mathcal{R}^{\infty}$.

Proof Given $\pi \in \mathcal{R}^{\infty}$, we are given that $xRy \Leftrightarrow U^r(x) \geq U^r(y)$.

Thus, $\forall (x_1, 2x)$ and $(x_1, 2y) \in X$,

$$\begin{aligned}(x_1, 2^x) R (x_1, 2^y) &\Leftrightarrow U^F(x_1, 2^x) \geq U^F(x_1, 2^y) \\ &\Leftrightarrow U^F(2^x) \geq U^F(2^y) \\ &\Leftrightarrow 2^x R 2^y.\end{aligned}$$

It is readily checked that general utility representations and discounting utility representations in particular always satisfy Pareto and Independence. Thus Theorem 4 combined with Theorem 1 says that a discounting choice function with full domain \mathcal{R}^{∞} cannot exist. If the domain is restricted to $\tilde{\mathcal{R}}^{\infty}$, then Theorem 3 applies. We conclude that a search for a fixed discount rate, to be applied somehow to all conceivable intertemporal choice decisions, can at best result in a present generation dictatorship. We also mention that Koopmans [2] connects stationarity and discounting in a much more profound fashion, showing that stationarity together with four other assumptions on preferences serves to completely characterize a discounting rule.

V. Conclusions

We have shown formally that for infinite generation, transitive, intertemporal choice procedures the familiar conditions of Pareto optimality and independence together with the equity condition of stationarity are incompatible unless one is willing to accept a dictatorship. As in the case of Arrow's impossibility theorem, the result is primarily conceptual rather than practical. Real world decisions are seldom made by a pre-specified rule that is broad enough to deal with a full domain of individual or generational preference profiles. Nevertheless, impossibility results give us important clues about inherent and surprising limitations in institutions we often accept unquestioningly.

The appearance of stationarity, introduced in [2] and proposed as a social choice condition in [1], leads to broader and more practical questions involving discount rates. If we accept the full intertemporal choice model as presented, then we are led to question the validity of any debate about "fair" discount rates. Indeed, no discount rate regardless of how close it may be to unity can be fair under the assumptions we have used. Such discounting implies stationarity and is tantamount to giving dictatorial power to the first generation. Since discount rates are in practice applied in much more specialized and varied ways, this criticism of discounting must be tempered somewhat. The results do suggest, however, that discounting should be employed discriminately and with caution.

The dictatorship result of Arrow's theorem depends on the assumption that the social choice function is transitive. If transitivity is weakened to quasitransitivity or acyclicity, then one obtains, respectively, an oligarchy

or a collegial polity (see [5] and [6]) . It is not yet clear to what extent our results carry over to these situations, though it seems reasonable to conjecture that, in the quasitransitive case, the oligarchy must consist of consecutive initial segments of generations. In the acyclic case one would likewise expect the present generation to belong to the collegium (the intersection of all decisive sets) and thus to have, at least in the monotonic case, a power of veto.

REFERENCES

- [1] FEREJOHN, J.A. and PAGE, T. "On the Foundations of Intertemporal Choice," American Journal of Agricultural Economics 60, May 1978: 269-275.
- [2] KOOPMANS, T. "Two Papers on the Representation of Preference Orderings with Independent Components of Consumption and Representation of Preference Orderings Over Time." Cowles Foundation Paper No. 366, 1972.
- [3] PLOTT, C.R. "Axiomatic Social Choice Theory: An Overview and Interpretation." American Journal of Political Science 20, August 1976:511-596.
- [4] DIAMOND, P. "The Evaluation of Infinite Utility Streams," Econometrica 33,1965:170-77.
- [5] HANSSON, B. "The Existence of Group Preference Functions," Public Choice 38, Winter 1976:89-98.
- [6] BROWN, D.J. "Acyclic Choice." Cowles Foundation Paper No. 360, 1973.