Département de sciences économiques

2006-03

Consistent Relations

BOSSERT, Walter

Université de Montréal

Département de sciences économiques
Université de Montréal
Faculté des arts et des sciences
C.P. 6128, succursale Centre-Ville
Montréal (Québec) H3C 3J7
Canada
http://www.sceco.umontreal.ca
SCECO-information@UMontreal.CA
Téléphone : (514) 343-6539
Télécopieur : (514) 343-7221

Ce cahier a également été publié par le Centre interuniversitaire de recherche en économie quantitative (CIREQ) sous le numéro 03-2006.

This working paper was also published by the Center for Interuniversity Research in Quantitative Economics (CIREQ), under number 03-2006.

ISSN 0709-9231

Consistent Relations*

Walter Bossert

Département de Sciences Economiques and CIREQ
Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal QC H3C 3J7
Canada
FAX: (+1 514) 343 7221
e-mail: walter.bossert@umontreal.ca

This version: February 2006

Abstract. Consistency, a natural weakening of transitivity introduced in a seminal contribution by Suzumura (1976b), has turned out to be an interesting and promising concept in a variety of areas within economic theory. This paper summarizes its recent applications and provides some new observations in welfarist social choice and in population ethics. In particular, it is shown that the conclusion of the welfarism theorem remains true if transitivity is replaced by consistency and that an impossibility result in variable-population social-choice theory turns into a possibility if transitivity is weakened to consistency. *Journal of Economic Literature* Classification Nos.: D01, D63, D71.

Keywords: Consistency, Ordering Extensions, Rationalizability, Welfarism, Population Ethics.

* Financial support through a grant from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

1 Introduction

Binary relations are at the heart of much of economic theory, both in the context of individual choice and in multi-agent decision problems. A fundamental coherence requirement imposed on a relation is the well-known transitivity axiom. If a relation is interpreted as a goodness relation, transitivity postulates that whenever one alternative is at least as good as a second and the second alternative is, in turn, at least as good as a third, then the first alternative is at least as good as the third. However, from an empirical as well as a conceptual perspective, transitivity is frequently considered too demanding and weaker notions of coherence have been proposed in the literature. Two alternatives that have received a considerable amount of attention in the literature are quasi-transitivity and acyclicity. Quasi-transitivity demands that the asymmetric factor of a relation (the betterness relation) is transitive, whereas acyclicity rules out the presence of betterness cycles. Quasi-transitivity is implied by transitivity and implies acyclicity. The reverse implications are not valid.

Suzumura (1976b) introduced an interesting alternative weakening of transitivity and showed that it can be considered a more intuitive weakening than quasi-transitivity. This notion of coherence is called consistency, and it rules out the presence of cycles with at least one instance of betterness. Thus, the axiom is stronger than acyclicity. Moreover, consistency is weaker than transitivity and is equivalent to it in the presence of reflexivity and completeness. Quasi-transitivity and consistency are independent.

Consistency is exactly what is needed to avoid the phenomenon of a money pump. If consistency is violated by an agent's goodness relation, there exists a cycle with at least one instance of betterness. In this case, the agent under consideration is willing to trade an alternative for another alternative (where 'willing to trade' is interpreted as being at least as well-off after the trade as before), the second alternative for a third and so on until an alternative is reached such that getting back the original alternative is better than retaining possession of the last alternative in the chain. Thus, at the end of such a chain of exchanges, the agent is willing to give up the last alternative and, in addition, to pay a positive amount in order to get back the original alternative.

An important property of consistency is that it is necessary and sufficient for the existence of an ordering extension of a relation. Szpilrajn (1930) showed that, for any asymmetric and transitive relation, there exists an asymmetric, transitive and complete relation that contains the original relation. An analogous result applies if asymmetry is replaced with reflexivity. Suzumura (1976b) has shown that the transitivity assumption can be weakened to consistency without changing the conclusion regarding the existence of an ordering extension. Moreover, consistency is the weakest possible property that guarantees this existence result. Because extension theorems are of considerable importance in many applications of set theory, this is a fundamental result and illustrates the significance of the consistency property.

The purpose of this paper is to review the uses of consistency in a variety of applications and to provide some new observations, with the objective of further underlining the importance of this axiom. The first step is a statement of Suzumura's (1976b) extension theorem in the following section, followed by an application in the theory of rational choice due to Bossert, Sprumont and Suzumura (2005a) in Section 3. The last two sections provide new observations. In Section 4, a variant of the welfarism theorem that assumes consistency instead of transitivity is provided, and Section 5 illustrates how an impossibility result in population ethics can be turned into a possibility by weakening transitivity to consistency.

2 Relations and Extensions

Suppose X is a non-empty set of alternatives and $R \subseteq X \times X$ is a (binary) relation on X which is interpreted as a *goodness* relation, that is, $(x, y) \in R$ means that x is considered at least as good as y by the agent (or society) under consideration. The *asymmetric factor* of R is defined by

$$P(R) = \{(x, y) \mid (x, y) \in R \text{ and } (y, x) \notin R\},$$

and the *symmetric factor* of R is

$$I(R) = \{(x, y) \mid (x, y) \in R \text{ and } (y, x) \in R\}.$$

Given the interpretation of R as a goodness relation, $P(R)$ is the *better-than* relation corresponding to R and $I(R)$ is the *equally-good* relation associated with R.

The *transitive closure* $tc(R)$ of a relation R is defined by

$$\begin{aligned} tc(R) \quad = \quad & \{(x, y) \mid \text{ there exist } M \in \mathbb{N} \text{ and } x^0, \ldots, x^M \in X \text{ such that} \\ & x = x^0, (x^{m-1}, x^m) \in R \text{ for all } m \in \{1, \ldots, M\} \text{ and } x^M = y\}. \end{aligned}$$

A relation R is *reflexive* if, for all $x \in X$,

$$(x, x) \in R$$

and R is *asymmetric* if

$$R = P(R).$$

Furthermore, R is *complete* if, for all $x, y \in X$,

$$x \neq y \;\Rightarrow\; (x, y) \in R \text{ or } (y, x) \in R.$$

and R is *transitive* if, for all $x, y, z \in X$,

$$(x, y) \in R \text{ and } (y, z) \in R \;\Rightarrow\; (x, z) \in R.$$

R is *consistent* if, for all $x, y \in X$,

$$(x, y) \in tc(R) \Rightarrow (y, x) \notin P(R).$$

A *quasi-ordering* is a reflexive and transitive relation and an *ordering* is a complete quasi-ordering.

The notion of consistency is due to Suzumura (1976b) and it is equivalent to the requirement that any cycle must be such that all relations involved in this cycle are instances of equal goodness—betterness cannot occur. Clearly, this requirement implies (but is not implied by) the well-known *acyclicity* axiom which rules out the existence of betterness cycles (cycles where *all* relations involve the asymmetric factor of the relation). Consistency and *quasi-transitivity,* which requires that $P(R)$ is transitive, are independent. Transitivity implies consistency but the reverse implication is not true in general. However, if R is reflexive and complete, consistency and transitivity are equivalent.

A relation R' is an *extension* of a relation R if

$$R \subseteq R' \text{ and } P(R) \subseteq P(R').$$

If an extension R' of R is an ordering, we refer to R' as an *ordering extension* of R. One of the most fundamental results on extensions of binary relations is due to Szpilrajn (1930) who showed that any transitive and asymmetric relation has a transitive, asymmetric and complete extension. The result remains true if asymmetry is replaced with reflexivity, that is, any quasi-ordering has an ordering extension. Arrow (1951, p. 64) stated this generalization of Szpilrajn's theorem without a proof and Hansson (1968) provided a proof on the basis of Szpilrajn's original theorem.

While the property of being a quasi-ordering is sufficient for the existence of an ordering extension of a relation, this is not necessary. As shown by Suzumura (1976b), consistency is *necessary and sufficient* for the existence of an ordering extension. This observation is stated formally in the following theorem; see Suzumura (1976b, pp. 389–390) for a proof.

Theorem 1 *A relation R has an ordering extension if and only if R is consistent.*

Theorem 1 is an important result. It establishes that consistency is the weakest possible property of a relation that still guarantees the existence of an ordering extension. Note that quasi-transitivity (which, as mentioned earlier, is logically independent of consistency) has nothing to do with the possibility of extending a binary relation to an ordering.

3 Rational Choice

Consistency has recently been examined in the context of *rational choice.* Observed (or observable) choices are *rationalizable* if there exists a relation such that, for any feasible set, the set of chosen alternatives coincides with the set of greatest or maximal elements according to this relation.

Following the contributions of Richter (1966; 1971), Hansson (1968), Suzumura (1976a; 1977; 1983, Chapter 2) and others, the approach to rational choice analyzed in this paper is capable of accommodating a wide variety of choice situations because no restrictions (other than non-emptiness) are imposed on the domain of a choice function. Letting $\mathcal{X}$ denote the power set of X excluding the empty set, a *choice function* is a mapping $C\colon \Sigma \to \mathcal{X}$ such that $C(S) \subseteq S$ for all $S \in \Sigma$, where $\Sigma \subseteq \mathcal{X}$ with $\Sigma \neq \emptyset$ is the domain of C.

The *direct revealed preference relation* $R_C \subseteq X \times X$ of a choice function C with an arbitrary domain Σ is defined as

$$R_C = \{(x, y) \mid \text{ there exists } S \in \Sigma \text{ such that } x \in C(S) \text{ and } y \in S\}.$$

The *(indirect) revealed preference relation* of C is the transitive closure $tc(R_C)$ of the direct revealed preference relation R_C.

A choice function C is *greatest-element rationalizable* if there exists a relation R on X such that

$$C(S) = \{x \in S \mid (x, y) \in R \text{ for all } y \in S\}.$$

If such a relation R exists, it is called a *rationalization* of C. The most common alternative to greatest-element rationalizability is *maximal-element* rationalizability which requires the existence of a relation R such that, for all feasible sets S, $C(S)$ is equal to the set of maximal elements in S according to R, that is, no element in S is strictly preferred to any element in $C(S)$. Bossert, Sprumont and Suzumura (2005b) provide a detailed analysis of maximal-element rationalizability. Logical relationships between and characterizations of various notions of rationalizability, both on arbitrary domains and under more specific domain assumptions, can be found in Bossert, Sprumont and Suzumura (2006).

In order to interpret a rationalization as a goodness relation, it is usually required that it satisfy additional properties such as the *richness* axioms reflexivity and completeness, or one of the *coherence* properties acyclicity, quasi-transitivity, consistency and transitivity. The full set of rationalizability notions that can be obtained by combining one or both (or none) of the richness properties with one (or none) of the coherence properties is analyzed in Bossert and Suzumura (2005). They show that, if all these combinations are available, it is sufficient to restrict attention to greatest-element rationalizability: for each notion of maximal-element rationalizability, there exists a notion of greatest-element rationalizability (possibly involving different richness and coherence properties) that is equivalent. Thus, restricting attention to greatest-element rationalizability does not involve any loss of generality.

Bossert, Sprumont and Suzumura (2005a) have characterized all notions of rationalizability when the coherence property required is consistency. As mentioned earlier, consistency and transitivity are equivalent in the presence of reflexivity and completeness. Thus, greatest-element rationalizability by a reflexive, complete and consistent relation is equivalent to greatest-element rationalizability by an ordering and Richter's (1966; 1971) results apply; see the next theorem. Moreover, greatest-element rationalizability by a complete and consistent relation implies greatest-element rationalizability by a reflexive, complete and consistent relation, and greatest-element rationalizability by a consistent relation implies greatest-element rationalizability by a reflexive and consistent relation, and analogous observations apply in the case of maximal-element rationalizability; see Bossert, Sprumont and Suzumura (2005a, Theorem 1). As pointed out in Bossert, Sprumont and Suzumura (2006), as soon as the coherence properties quasi-transitivity or acyclicity are imposed, reflexivity no longer is guaranteed as an additional property of a rationalization. Thus, consistency stands out from these alternative weakenings of transitivity in this regard: as is the case for transitive greatest-element (or maximal-element) rationalizability, any notion of consistent greatest-element (or maximal-element) rationalizability is equivalent to the definition that is obtained if reflexivity is added as a property of a rationalization.

Richter (1971) showed that the following axiom is necessary and sufficient for greatest-element rationalizability by a transitive relation and by an ordering. Thus, the existence of a rationalizing relation that is not merely a quasi-ordering but an ordering follows from greatest-element rationalizability by a transitive relation. This observation sets transitive greatest-element rationalizability apart from other notions of greatest-element rationalizability involving weaker coherence requirements.

Transitive-closure coherence. For all $S \in \Sigma$, for all $x \in S$,

$$(x, y) \in tc(R_C) \text{ for all } y \in S \;\Rightarrow\; x \in C(S).$$

We now obtain the following result; see Bossert, Sprumont and Suzumura (2005a) for a proof.

Theorem 2 *C is greatest-element rationalizable by a (reflexive,) complete and consistent relation if and only if C satisfies transitive-closure coherence.*

If completeness is dropped as a requirement imposed on a rationalization, a weaker notion of greatest-element rationalizability is obtained. In contrast to greatest-element rationalizability by a quasi-transitive or an acyclical relation which leads to much more complex necessary and sufficient conditions (see Bossert and Suzumura, 2005), requiring a rationalization to be consistent preserves the intuitive and transparent nature of the characterization stated in Theorem 2. As is the case for the transitive closure, there is a unique minimal consistent relation that has to be respected by any consistent rationalization, namely, the *consistent closure* of R_C. The consistent closure $cc(R)$ of a relation R is defined by

$$cc(R) = R \cup \{(x, y) \mid (x, y) \in tc(R) \text{ and } (y, x) \in R\}.$$

Clearly, $R \subseteq cc(R) \subseteq tc(R)$. Moreover, just as $tc(R)$ is the unique smallest transitive relation containing R, $cc(R)$ is the unique smallest consistent relation containing R. This latter property is crucial in obtaining a clear-cut and intuitive rationalizability result even without imposing completeness (in which case consistent greatest-element rationalizability is not equivalent to transitive greatest-element rationalizability). In contrast, there is no such thing as a quasi-transitive closure or an acyclical closure of a relation, which explains why rationalizability results involving these coherence properties are much more complex.

Greatest-element rationalizability by means of a consistent (and reflexive but not necessarily complete) relation can now be characterized by employing a natural weakening of transitive-closure consistency: all that needs to be done is to replace the transitive closure of the direct revealed preference relation by its consistent closure.

Consistent-closure coherence. For all $S \in \Sigma$, for all $x \in S$,

$$(x, y) \in cc(R_C) \text{ for all } y \in S \;\Rightarrow\; x \in C(S).$$

The following characterization is due to Bossert, Sprumont and Suzumura (2005a).

Theorem 3 *C is greatest-element rationalizable by a (reflexive and) consistent relation if and only if C satisfies consistent-closure coherence.*

4 Welfarism

Following Arrow's (1951; 1963) impossibility theorem, one route of escape from its negative consequences that has been chosen in the subsequent literature is to assume that a social ranking is established on the basis of a richer informational framework. In Arrow's setup, the individual goodness relations form the informational basis of collective choice. This approach rules out, in particular, interpersonal comparisons of well-being. An informationally richer environment is obtained if a social ranking is allowed to depend on *utility* profiles instead of profiles of goodness relations, and these utilities can be assumed to carry more than just ordinally measurable and interpersonally non-comparable information regarding the well-being of the agents. Under an implicit regularity assumption that guarantees the existence of representations of the individual goodness relations, the Arrow framework is included as a special case: it corresponds to the informational assumption of ordinal measurability and interpersonal non-comparability.

The universal set of alternatives X is assumed to contain at least three elements. The set of all utility functions $U \colon X \to \mathbb{R}$ is denoted by $\mathcal{U}$ and its n-fold Cartesian product is $\mathcal{U}^n$. A *utility profile* is an n-tuple $\mathbf{U} = (U_1, \ldots, U_n) \in \mathcal{U}^n$.

A *collective choice functional* is a mapping $F \colon \mathcal{E} \to \mathcal{B}$ where $\mathcal{E} \subseteq \mathcal{U}^n$ is the domain of this functional, assumed to be non-empty, and $\mathcal{B}$ is the set of all binary relations on X. For each utility profile $\mathbf{U} \in \mathcal{E}$, $F(\mathbf{U})$ is the social preference corresponding to $\mathbf{U}$. A *reflexive and consistent collective choice functional* is a collective choice functional F such that $F(\mathbf{U})$ is reflexive and consistent for all $\mathbf{U} \in \mathcal{E}$, and a *social-evaluation functional* is a collective choice functional F such that $F(\mathbf{U})$ is an ordering for all $\mathbf{U} \in \mathcal{E}$. Informational assumptions regarding the measurability and interpersonal comparability of individual utilities can be expressed by requiring the collective choice functional to be constant on sets of utility profiles that contain the same information. For example, if utilities are *cardinally measurable and fully comparable,* any utility profile $\mathbf{U}'$ that is obtained from a profile $\mathbf{U}$ by applying the same increasing affine transformation to all individual utility functions carries the same information as $\mathbf{U}$ itself. Thus, the collective choice functional must assign the same social ranking to both profiles. See Blackorby, Donaldson and Weymark (1984) or Bossert and Weymark (2004) for discussions of information assumptions in social-choice theory.

A fundamental result in this setting is the *welfarism theorem;* see, for instance, d'Aspremont and Gevers (1977) and Hammond (1979). A social-evaluation functional F is *welfarist* if, for any utility profile $\mathbf{U}$ and for any two alternatives x and y, the social ranking of x and y according to the social ordering assigned to the profile $\mathbf{U}$ by F depends on the two utility vectors $\mathbf{U}(x) = (U_1(x), \ldots, U_n(x))$ and $\mathbf{U}(y) = (U_1(y), \ldots, U_n(y))$ only. Thus, a *single* ordering of utility vectors is sufficient to rank the alternatives for *any* profile. The welfarism theorem states that, provided that the domain of the social-welfare functional consists of all possible utility profiles, welfarism is equivalent to the conjunction of Pareto indifference and independence of irrelevant alternatives.

In this section, it is illustrated that the welfarism theorem has an analogous formulation for reflexive and consistent collective choice functionals: even if every social ranking is merely required to be reflexive and consistent rather than an ordering, the conjunction of the two axioms is (under the unlimited-domain assumption) equivalent to the existence of a single reflexive and consistent relation R defined on utility vectors that is sufficient to obtain the social ranking for any utility profile. This relation $R \subseteq \mathbb{R}^n \times \mathbb{R}^n$ is referred to as a *social-evaluation relation.* The requisite axioms are the following.

Unlimited domain. $\mathcal{E} = \mathcal{U}^n$.

Pareto indifference. For all $x, y \in X$ and for all $\mathbf{U} \in \mathcal{E}$,

$$U_i(x) = U_i(y) \; \forall i \in N \;\Rightarrow\; (x, y) \in I\left(F(\mathbf{U})\right).$$

Independence of irrelevant alternatives. For all $x, y \in X$ and for all $\mathbf{U}, \mathbf{U}' \in \mathcal{E}$ such that $U_i(x) = U_i'(x)$ and $U_i(y) = U_i'(y)$ for all $i \in N$,

$$(x, y) \in F(\mathbf{U}) \Leftrightarrow (x, y) \in F(\mathbf{U}') \;\text{ and }\; (y, x) \in F(\mathbf{U}) \Leftrightarrow (y, x) \in F(\mathbf{U}').$$

The following theorem generalizes the standard welfarism theorem by allowing social relations to be intransitive and incomplete while retaining the consistency requirement.

Theorem 4 *Suppose that a reflexive and consistent collective choice functional F satisfies unlimited domain. F satisfies Pareto indifference and independence of irrelevant alternatives if and only if there exists a reflexive and consistent social-evaluation relation $R \subseteq \mathbb{R}^n \times \mathbb{R}^n$ such that, for all $x, y \in X$ and for all $\mathbf{U} \in \mathcal{U}^n$,*

$$(x, y) \in F(\mathbf{U}) \Leftrightarrow (\mathbf{U}(x), \mathbf{U}(y)) \in R. \tag{1}$$

Proof. The 'if' part of the theorem is straightforward to verify. To prove the converse implication, suppose that F is a reflexive and consistent collective choice functional satisfying unlimited domain, Pareto indifference and independence of irrelevant alternatives. Define the relation $R \subseteq \mathbb{R}^n \times \mathbb{R}^n$ as follows. For all $u, v \in \mathbb{R}^n$, $(u, v) \in R$ if and only if there exist $x, y \in X$ and $\mathbf{U} \in \mathcal{U}^n$ such that $\mathbf{U}(x) = u$, $\mathbf{U}(y) = v$ and $(x, y) \in F(\mathbf{U})$. That R is well-defined follows as in the standard welfarism theorem; see, for instance, Blackorby, Donaldson and Weymark (1984) or Bossert and Weymark (2004). Once R is well-defined, (1) is immediate and, furthermore, R is reflexive because $F(\mathbf{U})$ is reflexive for all $\mathbf{U} \in \mathcal{U}^n$. The proof is completed by showing that R is consistent. Let $u, v \in \mathbb{R}^n$ be such that $(u, v) \in tc(R)$. By definition of the transitive closure of a relation, there exist $M \in \mathbb{N}$ and $u^0, \ldots, u^M \in \mathbb{R}^n$ such that $u = u^0$, $(u^{m-1}, u^m) \in R$ for all $m \in \{1, \ldots, M\}$ and $u^M = v$. By definition of R, there exist $x^0, \ldots, x^M \in X$ and $\mathbf{U}^1, \ldots, \mathbf{U}^M \in \mathcal{U}^n$ such that $\mathbf{U}^{m-1}(x^{m-1}) = u^{m-1}$, $\mathbf{U}^{m-1}(x^m) = u^m$ and $(x^{m-1}, x^m) \in F(\mathbf{U}^{m-1})$ for all $m \in \{1, \ldots, M\}$. By unlimited domain, there exists $\mathbf{V} \in \mathcal{U}^n$ such that $\mathbf{V}(x^m) = u^m$ for all $m \in \{0, \ldots, M\}$. Using (1), it follows that $(x^{m-1}, x^m) \in F(\mathbf{V})$ for all $m \in \{1, \ldots, M\}$. Because $F(\mathbf{V})$ is consistent, it follows that $(x^M, x^0) \notin P(F(\mathbf{V}))$. Thus, by (1), $(v, u) = \big(\mathbf{V}(x^M), \mathbf{V}(x^0)\big) \notin P(R)$ and R is consistent. ■

5 Population Ethics

The traditional social-choice framework with a fixed population is unable to capture important aspects of many public-policy decision problems. For example, decisions on funds devoted to prenatal care, the intergenerational allocation of resources and the design of aid packages to developing countries involve endogenous populations: depending on the selected alternative, some individuals may or may not be brought into existence. To address this issue, a social ranking must be capable of comparing alternatives with different population sizes.

The possibility of extending the welfarist approach to a variable-population environment has been examined in a variety of contributions, most notably in applied ethics; see, for instance, Parfit (1976, 1982, 1984). Impossibility results arise frequently in this area, and it is therefore of interest to examine the possibilities of escaping these negative conclusions. The purpose of this section is to illustrate that weakening transitivity to consistency turns these impossibilities into possibilities. Of course, in order to ensure that consistency is indeed weaker than transitivity, we cannot impose reflexivity, completeness and consistency—as mentioned earlier, consistency and transitivity coincide in the presence of the two richness conditions. Therefore, the question arises whether reflexivity and completeness rather than transitivity are, to a large extent, responsible for the impossibilities. This is not the case: although most of the impossibility results in this area have been established for orderings, they remain true if reflexivity and completeness are dropped.

A variable-population version of a social-evaluation relation is defined on the set of utility vectors of *any* dimension, that is, it is a relation $R \subseteq \Omega \times \Omega$, where $\Omega = \cup_{n \in \mathbb{N}} \mathbb{R}^n$. The components of a utility vector $u \in \Omega$ are interpreted as the *lifetime* utilities of those alive in the requisite alternative. For an individual who is alive, a *neutral* life is one which is as good as one without experiences. A life above neutrality is worth living, a life below neutrality is not. Following standard practice in population ethics, a lifetime-utility level of zero is assigned to neutrality.

In Blackorby, Bossert and Donaldson (2006), it is shown that there exists no variable-population social-evaluation ordering satisfying four axioms that are common in the literature. This result can be generalized by noting that it does not make use of reflexivity of completeness—all that is needed is the transitivity of R.

The first axiom is *minimal increasingness.* It requires that, for any fixed population size, if all individuals have the same utility in two utility vectors, then the vector where everyone's utility is higher is better according to R. We use $\mathbf{1}_n$ to denote the vector of $n \in \mathbb{N}$ ones.

Minimal increasingness. For all $n \in \mathbb{N}$ and for all $\beta, \gamma \in \mathbb{R}$,

$$\beta > \gamma \;\Rightarrow\; (\beta \mathbf{1}_n, \gamma \mathbf{1}_n) \in P(R).$$

Minimal increasingness is a weak unanimity property: it only applies if everyone has the same utility in both alternatives to be compared.

Another fixed-population axiom is *weak inequality aversion.* This axiom demands that, for any fixed population size, perfect equality is at least as good as any distribution of the same total utility.

Weak inequality aversion. For all $n \in \mathbb{N}$ and for all $u \in \mathbb{R}^n$,

$$\left(\left(\frac{1}{n}\sum_{i=1}^{n} u_i\right)\mathbf{1}_n, u\right) \in R.$$

Sikora (1978) suggests a variable-population version of the Pareto principle. The axiom usually is defined as the conjunction of the strong-Pareto principle and the requirement that the addition of an individual above neutrality to a utility-unaffected population is a social improvement. Because strong Pareto will be introduced as a separate axiom later on and is not needed for the impossibility result, we use the second part of the property only.

Pareto plus. For all $n \in \mathbb{N}$, for all $u \in \mathbb{R}^n$ and for all $a \in \mathbb{R}_{++}$,

$$((u, a), u) \in P(R).$$

In the axiom statement, the population common to u and (u, a) is unaffected and, thus, in order to defend the axiom on individual-goodness grounds, it must be argued that a level of well-being above neutrality is better than non-existence. Thus, the axiom extends the Pareto condition to situations where a person is not alive in all alternatives to be compared. While it is possible to compare alternatives with different populations from a social point of view (which is the issue addressed in population ethics), it is not clear that such a comparison can be made from the viewpoint of an individual if the person is not alive in one of the alternatives. It is therefore difficult to interpret this axiom as a Pareto condition because it appears to be based on the idea that people who do not exist have interests that should be respected. There is, therefore, an important asymmetry in the assessment of alternatives with different populations. It is perfectly reasonable to say that an individual considers life worth living if the person is alive with a positive level of lifetime well-being, but that does not justify the claim that a person who does not exist gains from being brought into existence with a lifetime utility above neutrality.

As is the case for Pareto plus, the final axiom used in our impossibility result applies to comparisons across population sizes. A variable-population social-evaluation relation leads to the *repugnant conclusion* if population size can always be substituted for well-being, no matter how close to neutrality the utilities of a large population are. That is, mass poverty may be considered superior to some alternatives in which fewer people lead very good lives. This property has been used by Parfit (1976, 1982, 1984) to argue against *classical utilitarianism,* the variable-population social-evaluation ordering that ranks utility vectors on the basis of their total utilities. If Parfit's view is accepted, R should be required to avoid the repugnant conclusion.

Avoidance of the repugnant conclusion. There exist $n \in \mathbb{N}$, $\xi \in \mathbb{R}_{++}$ and $\varepsilon \in (0, \xi)$ such that, for all $m > n$,

$$(\varepsilon \mathbf{1}_m, \xi \mathbf{1}_n) \notin P(R).$$

Blackorby, Bossert and Donaldson (2006, Theorem 2) show that there exists no variable-population social-evaluation ordering satisfying the above four axioms; see Blackorby, Bossert and Donaldson (2005), Blackorby, Bossert, Donaldson and Fleurbaey (1998), Blackorby and Donaldson (1991), Carlson (1998), McMahan (1981) and Parfit (1976, 1982, 1984) for similar observations. The following theorem shows that reflexivity and completeness are not required—transitivity of R is sufficient to generate the impossibility.

Theorem 5 *There exists no transitive variable-population social-evaluation relation satisfying minimal increasingness, weak inequality aversion, Pareto plus and avoidance of the repugnant conclusion.*

Proof. Suppose R satisfies minimal increasingness, weak inequality aversion and Pareto plus. The proof is completed by showing that R leads to the repugnant conclusion. For any population size $n \in \mathbb{N}$, let $\xi \in \mathbb{R}_{++}$, $\varepsilon \in (0, \xi)$ and $\delta \in \mathbb{R}_{++}$ be such that $0 < \delta < \varepsilon < \xi$. Choose any integer r such that

$$r > n \frac{(\xi - \varepsilon)}{(\varepsilon - \delta)}. \tag{2}$$

Because the numerator and denominator on the right side of the inequality are both positive, r is positive. By Pareto plus,

$$((\xi \mathbf{1}_n, \delta \mathbf{1}_r), \xi \mathbf{1}_n) \in P(R). \tag{3}$$

Average utility in $(\xi \mathbf{1}_n, \delta \mathbf{1}_r)$ is $(n\xi + r\delta)/(n + r)$ so, by minimal inequality aversion,

$$\left(\left(\frac{n\xi + r\delta}{n + r} \right) \mathbf{1}_{n+r}, (\xi \mathbf{1}_n, \delta \mathbf{1}_r) \right) \in R. \tag{4}$$

By (2),

$$\varepsilon > \frac{n\xi + r\delta}{n+r}$$

and, by minimal increasingness,

$$\left(\varepsilon \mathbf{1}_{n+r}, \left(\frac{n\xi + r\delta}{n+r}\right) \mathbf{1}_{n+r}\right) \in P(R). \tag{5}$$

Combining (3), (4), (5) and using transitivity, it follows that $(\varepsilon \mathbf{1}_{n+r}, \xi \mathbf{1}_n) \in P(R)$ and avoidance of the repugnant conclusion is violated. ■

If transitivity is weakened to consistency, the axioms in the theorem statement are compatible. Moreover, three of them can be strengthened and other properties that are commonly imposed in population ethics can be added without obtaining an impossibility.

Expressed in the current setting, the strong-Pareto principle is another fixed-population axiom. If everyone alive in two alternatives with utility vectors u and v has a utility in u that is at least as high as the utility of this person in v with at least one strict inequality, u is better than v. Clearly, this axiom is a strengthening of minimal increasingness.

Strong Pareto. For all $n \in \mathbb{N}$ and for all $u, v \in \mathbb{R}^n$,

$$u_i \geq v_i \text{ for all } i \in \{1, \ldots, n\} \text{ and } u \neq y \;\Rightarrow\; (u, v) \in P(R).$$

Continuity is a condition that prevents the social-evaluation relation R from exhibiting 'large' changes in response to 'small' changes in a utility vector. Again, the axiom imposes restrictions on fixed-population comparisons only.

Continuity. For all $n \in \mathbb{N}$ and for all $u \in \mathbb{R}^n$, the sets $\{v \in \mathbb{R}^n \mid (v, u) \in R\}$ and $\{v \in \mathbb{R}^n \mid (u, v) \in R\}$ are closed in $\mathbb{R}^n$.

Weak inequality aversion can be strengthened by requiring the restriction of R to $\mathbb{R}^n \times \mathbb{R}^n$ to be *strictly S-concave* for any population size $n \in \mathbb{N}$; see, for instance, Marshall and Olkin (1979). Strict S-concavity is equivalent to the conjunction of the *strict transfer principle* familiar from the theory of inequality measurement and *anonymity.* The strict transfer principle requires that a progressive transfer increases goodness, provided the relative rank of the individuals involved in the transfer is unchanged; see Dalton (1920) and Pigou (1912). A social-evaluation relation is anonymous if the individuals in a fixed population are treated impartially, without paying attention to their identities; see Sen

(1973) for a detailed discussion. A *bistochastic $n \times n$ matrix* is a matrix whose entries are between zero and one and all row sums and column sums are equal to one.

Strict S-concavity. For all $n \in \mathbb{N}$, for all $u \in \mathbb{R}^n$ and for all bistochastic $n \times n$ matrices B,

(i) $(Bu, u) \in R$;

(ii) Bu is not a permutation of $u \Rightarrow (Bu, u) \in P(R)$.

Independence of the utilities of unconcerned individuals is a fixed-population separability property introduced by d'Aspremont and Gevers (1977). It requires that only the utilities of those who can possibly be affected by a choice between two fixed-population alternatives should determine their ranking.

Independence of the utilities of unconcerned individuals. For all $n, m \in \mathbb{N}$, for all $u, v \in \mathbb{R}^n$ and for all $w, s \in \mathbb{R}^m$,

$$((u, w), (v, w)) \in R \Leftrightarrow ((u, s), (v, s)) \in R.$$

We now turn to further variable-population axioms. *Existence of critical levels* requires that non-trivial trade-offs between population size and well-being are possible. For any utility vector $u \in \Omega$, it is assumed that there exists a utility level $c \in \mathbb{R}$ (which may depend on u) such that the ceteris-paribus addition of an individual with utility level c to an existing population with utilities u is a matter of indifference according to R.

Existence of critical levels. For all $u \in \Omega$, there exists $c \in \mathbb{R}$ such that

$$((u, c), u) \in I(R).$$

The *negative expansion principle* is dual to Pareto plus. It requires any utility distribution to be ranked as better than one with the ceteris-paribus addition of an individual whose life is not worth living—that is, with a lifetime utility below neutrality.

Negative expansion principle. For all $n \in \mathbb{N}$, for all $u \in \mathbb{R}^n$ and for all $a \in \mathbb{R}_{--}$,

$$(u, (u, a)) \in P(R).$$

Finally, a strengthening of avoidance of the repugnant conclusion is defined. It is obtained by replacing the existential quantifiers in the original axiom with universal quantifiers and replacing the negation of betterness in the conclusion with the negation of the at-least-as-good-as relation. This is a strong property and one might not want to endorse it; the reason why it is used to replace the weaker condition is that it makes the possibility result logically stronger.

Strong avoidance of the repugnant conclusion. For all $n \in \mathbb{N}$, for all $\xi \in \mathbb{R}_{++}$, for all $\varepsilon \in (0, \xi)$ and for all $m > n$,

$$(\varepsilon \mathbf{1}_m, \xi \mathbf{1}_n) \notin R.$$

We do not impose *avoidance of the sadistic conclusion* or any of its variants (see Arrhenius, 2000) because it is implied by some of the properties already defined.

Theorem 6 *There exists a reflexive and consistent variable-population social-evaluation relation satisfying strong Pareto, continuity, strict S-concavity, independence of the utilities of unconcerned individuals, existence of critical levels, Pareto plus, the negative expansion principle and strong avoidance of the repugnant conclusion.*

Proof. An example is sufficient to establish the theorem. Let $g: \mathbb{R} \to \mathbb{R}$ be a continuous, increasing and strictly concave function such that $g(0) = 0$ and define the relation R^* by letting, for all $n, m \in \mathbb{N}$, for all $u \in \mathbb{R}^n$ and for all $v \in \mathbb{R}^m$,

$$\begin{aligned}(u, v) \in R^* \quad \Leftrightarrow \quad & n = m \text{ and } \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i) \\ \text{or} \quad & m = n + 1 \text{ and } \exists \alpha \in \mathbb{R}_- \text{ such that } v = (u, \alpha) \\ \text{or} \quad & n = m + 1 \text{ and } \exists \beta \in \mathbb{R}_+ \text{ such that } u = (v, \beta).\end{aligned}$$

Strong Pareto is satisfied because g is increasing, continuity is satisfied because g is continuous, strict S-concavity follows from the strict concavity of g and independence of the utilities of unconcerned individuals is satisfied because of the additively separable structure of the criterion for fixed-population comparisons. Existence of critical levels is satisfied because zero is a critical level for all $u \in \Omega$, and Pareto plus and the negative expansion principle follow immediately from the definition of R^*. Strong avoidance of the repugnant conclusion is satisfied because $(\varepsilon \mathbf{1}_m, \xi \mathbf{1}_n) \notin R^*$ for all $n \in \mathbb{N}$, for all $\xi \in \mathbb{R}_{++}$, for all $\varepsilon \in (0, \xi)$ and for all $m > n$. That R^* is reflexive is immediate.

It remains to show that R^* is consistent. The first step is to prove that, for all $n, m \in \mathbb{N}$, for all $u \in \mathbb{R}^n$ and for all $v \in \mathbb{R}^m$,

$$(u, v) \in R^* \;\Rightarrow\; \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i) \tag{6}$$

and

$$(u,v) \in P(R^*) \;\Rightarrow\; \sum_{i=1}^{n} g(u_i) > \sum_{i=1}^{m} g(v_i). \tag{7}$$

To prove (6), suppose that $n, m \in \mathbb{N}$, $u \in \mathbb{R}^n$, $v \in \mathbb{R}^m$ and $(u,v) \in R^*$. According to the definition of R^*, there are three possible cases.

Case 1. $n = m$ and $\sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i)$. The conclusion is immediate in this case.

Case 2. $m = n+1$ and $\exists \alpha \in \mathbb{R}_-$ such that $v = (u, \alpha)$. Thus,

$$\sum_{i=1}^{m} g(v_i) = \sum_{i=1}^{n} g(u_i) + g(\alpha) \leq \sum_{i=1}^{n} g(u_i)$$

where the inequality follows because $\alpha \leq 0$ and, by the increasingness of g and the property $g(0) = 0$, $g(\alpha) \leq 0$.

Case 3. $n = m+1$ and $\exists \beta \in \mathbb{R}_+$ such that $u = (v, \beta)$. This implies

$$\sum_{i=1}^{n} g(u_i) = \sum_{i=1}^{m} g(v_i) + g(\beta) \geq \sum_{i=1}^{m} g(v_i)$$

where the inequality follows because $\beta \geq 0$ and thus $g(\beta) \geq 0$.

To prove (7), suppose $n, m \in \mathbb{N}$, $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$ are such that $(u,v) \in P(R^*)$. Again, there are three cases.

Case 1. $n = m$ and $\sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i)$. If $\sum_{i=1}^{m} g(v_i) \geq \sum_{i=1}^{n} g(u_i)$, we obtain $(v,u) \in R^*$ and thus a contradiction to our hypothesis $(u,v) \in P(R^*)$. Therefore, $\sum_{i=1}^{n} g(u_i) > \sum_{i=1}^{m} g(v_i)$.

Case 2. $m = n+1$ and $\exists \alpha \in \mathbb{R}_-$ such that $v = (u, \alpha)$. Thus,

$$\sum_{i=1}^{m} g(v_i) = \sum_{i=1}^{n} g(u_i) + g(\alpha) \leq \sum_{i=1}^{n} g(u_i) \tag{8}$$

as established in the proof of (6). If $\alpha = 0$, it follows that $v = (u, 0)$ which leads to $(v,u) \in R^*$, contradicting our hypothesis $(u,v) \in P(R^*)$. Thus, $\alpha < 0$ and $g(\alpha) < 0$ because $g(0) = 0$ and g is increasing. Therefore, the inequality in (8) is strict.

Case 3. $n = m+1$ and $\exists \beta \in \mathbb{R}_+$ such that $u = (v, \beta)$. This implies

$$\sum_{i=1}^{n} g(u_i) = \sum_{i=1}^{m} g(v_i) + g(\beta) \geq \sum_{i=1}^{m} g(v_i) \tag{9}$$

as established in the proof of (6). If $\beta = 0$, it follows that $u = (v, 0)$ which leads to $(v,u) \in R^*$, again contradicting the hypothesis $(u,v) \in P(R^*)$. Thus, $\beta > 0$ and the inequality in (9) is strict.

To complete the proof, suppose $n, m \in \mathbb{N}$, $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$ are such that $(u, v) \in tc(R^*)$. By repeated application of (6), it follows that $\sum_{i=1}^n g(u_i) \geq \sum_{i=1}^m g(v_i)$. If $(v, u) \in P(R^*)$, (7) implies $\sum_{i=1}^m g(v_i) > \sum_{i=1}^n g(u_i)$, a contradiction. Thus, $(v, u) \notin P(R^*)$ and R^* is consistent. ■

Another impossibility result in population ethics is due to Broome (2004, Chapter 10). Broome suggests that existence is in itself neutral and, thus, the ceteris-paribus addition of an individual to a utility-unaffected population should lead to an equally-good alternative, at least as long as the utility of the added person (if brought into being) is within a non-degenerate interval. This intuition, which Broome calls the *principle of equal existence,* is incompatible with strong Pareto, provided that the social-evaluation relation R is transitive. The impossibility persists if transitivity is weakened to consistency. The following axiom is a weak form of the principle of equal existence.

Principle of equal existence. There exist $u \in \Omega$ and distinct $\alpha, \beta \in \mathbb{R}$ such that

$$((u, \alpha), u) \in I(R) \quad \text{and} \quad ((u, \beta), u) \in I(R). \tag{10}$$

We obtain the following impossibility result.

Theorem 7 *There exists no consistent variable-population social-evaluation relation satisfying strong Pareto and the principle of equal existence.*

Proof. Suppose R satisfies strong Pareto and the principle of equal existence. The proof is completed by showing that R cannot be consistent. By the principle of equal existence, there exist $u \in \Omega$ and distinct utility levels α and β such that (10) is satisfied. Without loss of generality, suppose $\alpha > \beta$. By strong Pareto, $((u, \alpha), (u, \beta)) \in P(R)$ which, together with (10), leads to a violation of consistency. ■

References

Arrhenius, G. (2000), "An impossibility theorem for welfarist axiologies," *Economics and Philosophy* **16**, 247–266.

Arrow, K. (1951; second ed. 1963), *Social Choice and Individual Values,* Wiley, New York.

Blackorby, C., W. Bossert and D. Donaldson (2005), *Population Issues in Social Choice Theory, Welfare Economics, and Ethics,* Cambridge University Press, Cambridge.

Blackorby, C., W. Bossert and D. Donaldson (2006), "Population ethics and the value of life," in: M. McGillivray, ed., *Inequality, Poverty and Well-being,* Macmillan, London, forthcoming.

Blackorby, C., W. Bossert, D. Donaldson and M. Fleurbaey (1998), "Critical levels and the (reverse) repugnant conclusion," *Journal of Economics* **67**, 1–15.

Blackorby, C. and D. Donaldson (1991), "Normative population theory: a comment," *Social Choice and Welfare* **8**, 261–267.

Blackorby, C., D. Donaldson and J. Weymark (1984), "Social choice with interpersonal utility comparisons: a diagrammatic introduction," *International Economic Review* **25**, 327–356.

Bossert, W., Y. Sprumont and K. Suzumura (2005a), "Consistent rationalizability," *Economica* **72**, 185–200.

Bossert, W., Y. Sprumont and K. Suzumura (2005b), "Maximal-element rationalizability," *Theory and Decision* **58**, 325–350.

Bossert, W., Y. Sprumont and K. Suzumura (2006), "Rationalizability of choice functions on general domains without full transitivity," *Social Choice and Welfare,* forthcoming.

Bossert, W. and K. Suzumura (2005), "Rational choice on general domains," *Working Paper,* Université de Montréal.

Bossert, W. and J. Weymark (2004), "Utility in social choice," in: S. Barberà, P. Hammond and C. Seidl, eds., *Handbook of Utility Theory, vol. 2: Extensions,* Kluwer, Dordrecht, 1099–1177.

Broome, J. (2004), *Weighing Lives,* Oxford University Press, Oxford.

Carlson, E. (1998), "Mere addition and the two trilemmas of population ethics," *Economics and Philosophy* **14**, 283–306.

Dalton, H. (1920), "The measurement of the inequality of incomes," *Economic Journal* **30**, 348–361.

d'Aspremont, C. and L. Gevers (1977), "Equity and the informational basis of collective choice," *Review of Economic Studies* **44**, 199–209.

Hammond, P. (1979), "Equity in two person situations: some consequences," *Econometrica* **47**, 1127–1135.

HANSSON, B. (1968), "Choice structures and preference relations," *Synthese* **18**, 443–458.

MARSHALL, A. AND I. OLKIN (1979), *Inequalities: Theory of Majorization and its Applications,* Academic Press, New York.

MCMAHAN, J. (1981), "Problems of population theory," *Ethics* **92**, 96–127.

PARFIT, D. (1976), "On doing the best for our children," in: M. Bayles, ed., *Ethics and Population,* Schenkman, Cambridge, 100–102.

PARFIT, D. (1982), "Future generations, further problems," *Philosophy and Public Affairs* **11**, 113–172.

PARFIT, D. (1984), *Reasons and Persons,* Oxford University Press, Oxford.

PIGOU, A. (1912), *Wealth and Welfare,* Macmillan, London.

RICHTER, M. (1966), "Revealed preference theory," *Econometrica* **41**, 1075–1091.

RICHTER, M. (1971), "Rational Choice," in: J. Chipman, L. Hurwicz, M. Richter and H. Sonnenschein, eds., *Preferences, Utility, and Demand,* Harcourt Brace Jovanovich, New York, 29–58.

SEN, A. (1973), *On Economic Inequality,* Oxford University Press, Oxford.

SIKORA, R. (1978), "Is it wrong to prevent the existence of future generations?," in: R. Sikora and B. Barry, eds., *Obligations to Future Generations,* Temple University Press, Philadelphia, 112–166.

SUZUMURA, K. (1976a), "Rational choice and revealed preference," *Review of Economic Studies* **43**, 149–158.

SUZUMURA, K. (1976b), "Remarks on the theory of collective choice," *Economica* **43**, 381–390.

SUZUMURA, K. (1977), "Houthakker's axiom in the theory of rational choice," *Journal of Economic Theory* **14**, 284–290.

SUZUMURA, K. (1983), *Rational Choice, Collective Decisions and Social Welfare,* Cambridge University Press, New York.

SZPILRAJN, E. (1930), "Sur l'extension de l'ordre partiel," *Fundamenta Mathematicae* **16**, 386–389.