Theoretical Economics 2 (2007), 187–202 1555-7561/20070187

Can intergenerational equity be operationalized?

WILLIAM R. ZAME
Department of Economics, University of California, Los Angeles

A long Utilitarian tradition has the ideal of equal regard for all individuals, both those now living and those yet to be born. The literature formalizes this ideal as asking for a preference relation on the space of infinite utility streams that is complete, transitive, invariant to finite permutations, and respects the Pareto ordering; an *ethical preference relation*, for short. This paper argues that operationalizing this ideal is problematic. Most simply, every ethical preference relation has the property that almost all (in the sense of outer measure) pairs of utility streams are indifferent. Even if we abandon completeness and respect for the Pareto ordering, every irreflexive preference relation that is invariant to finite permutations has the property that almost all pairs of utility streams are incomparable (not strictly ranked). Moreover, no ethical preference relation is measurable. As a consequence, the existence of an ethical preference relation is independent of the axioms used in almost all of formal economics and all of classical analysis. Finally, even if an ethical preference relation exists, it cannot be "explicitly described." These results have implications for game theory, for macroeconomics, and for economic development.

KEYWORDS. Intergenerational equity, infinite utility streams, long run averages, overtaking criterion, Utilitarianism.

JEL CLASSIFICATION. D60, D70, D90.

1. INTRODUCTION

A long Utilitarian tradition has the ideal of equal regard for all individuals—those now living and those not yet born. As Sidgwick (1907, p. 424) argues: "the time at which a man exists cannot affect the value of his happiness from a universal point of view . . . the interests of posterity must concern a Utilitarian as much as those of his contemporaries." Similarly, Ramsey (1928) asserts that any argument for preferring one generation over another must arise "merely from the weakness of the imagination," and Rawls (1971) makes similar arguments. The literature formalizes this goal by asking for a complete

William R. Zame: zame@econ.ucla.edu

This paper previously circulated under the titles "Can Utilitarianism be Operationalized?" and "Is Intergenerational Equity Possible?" I am grateful to Kaushik Basu and Mike Rothschild for stimulating my interest in the subject, to Harvey Friedman and the late John Myhill for teaching me most of what I know of model theory, to Jeff Ely, Shachar Kariv, Tom Liggett, Tony Martin, Pierre-Olivier Weill, the editor, and two referees for helpful comments, and to the National Science Foundation for financial support under grants 0317752, 0518936, and 0617027. Views expressed in this paper are those of the author and not of any funding agency.

transitive preference relation on the space of infinite utility streams (utilities of current and future individuals or of current and future generations) that is invariant under permuting the utility of any two individuals or generations (displays *intergenerational equity*) and respects the Pareto ordering—an *ethical preference relation*, for short.

This paper argues that operationalizing this ideal is problematic. Most simply (Theorem 1), every ethical preference relation has the property that almost all (in the sense of outer measure) pairs of utility streams are indifferent. Indeed, even if we abandon completeness and respect for the Pareto ordering, every irreflexive preference relation that displays intergenerational equity has the property that almost all pairs of utility streams are incomparable (not strictly ranked). More subtly (Theorem 2), no ethical preference relation is measurable. As a consequence (Theorem 3), the existence of a complete transitive preference relation that displays intergenerational equity and respects the Pareto ordering is independent of the axioms used in almost all of formal economics and all of classical analysis. (These are the Zermelo–Fraenkel Axioms, which formalize naive set theory in a consistent way, avoiding paradoxes such as the set of sets that are not elements of themselves, and the Axiom of Dependent Choice, which asserts the possibility of making a sequence of choices, each choice depending on the previous choices—but not the full Axiom of Choice.[1]) Moreover (Theorem 4), even if such a preference relation exists, it cannot be "explicitly described."

These results can also be interpreted in the context of repeated games (and in macroeconomics and in economic development), with respect to the question of modeling the preferences of an agent (or a planner) over an infinite stream of payoffs. The usual assumptions are that preferences are increasing (high payoffs are preferred to low payoffs) and display impatience (high payoffs are preferred soon rather than late); the discounted sum of utilities for period payoffs is the canonical example of a preference relation with these properties. But how should we model the preferences of an agent who is *perfectly patient*—indifferent to the order in which payoffs arrive, or at least to the order in which any finite sequence of payoffs arrive? Requiring that preferences over infinite utility streams be increasing and display perfect patience is formally equivalent to requiring that they respect the Pareto ordering and display intergenerational equity, so precisely the same conclusions as above obtain.

This paper contributes to a substantial literature. Diamond (1965) shows that a complete transitive preference relation that displays intergenerational equity and respects the Pareto ordering cannot be continuous in the topology induced by the supremum norm. Basu and Mitra (2003) show that such a preference relation—whether continuous or not—cannot be represented by a (real-valued) utility function. On the other hand, Svensson (1980) proves that such preference relations do exist.[2]

A sketch of Svensson's argument will help to place the results of the present paper in context. Write $X = [0,1]^{\mathbb{N}}$ for the space of utility streams and define an incomplete

[1]Perhaps the most important uses of the full Axiom of Choice in economics are through applications of non-standard analysis; see Anderson (1991) and Anderson and Raimondo (2005) for example.

[2]Fleurbaey and Michel (2003), Hara et al. (2006), Basu and Mitra (forthcoming), and Bossert et al. (forthcoming) provide further results, both positive and negative.

preference relation $\succ$ on X in the following way: $y \succ x$ exactly when there is a finite permutation y^* of y (that is a reordering of finitely many of the terms of y) such that $y^* > x$ (i.e., $y^* \geq x$ and $y^* \neq x$). The relation $\succ$ is irreflexive, transitive, displays intergenerational equity and respects the Pareto ordering, but it is incomplete: some—indeed, many—pairs of utility streams are not comparable. However we can use Szpilrajn's (1930) extension lemma to find an extension $\succeq$ of $\succ$ to a complete transitive preference relation on X. This extension $\succeq$ automatically displays intergenerational equity and respects the Pareto ordering; i.e., it is an ethical preference relation.

To relate this argument to the results of this paper, begin by noticing that the set of pairs $(x, y) \in X \times X = [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}}$ for which $y > x$ has measure 0 (with respect to the product of Lebesgue measure on all factors). Because there are only a countable number of finite permutations of the integers $\mathbb{N}$, it follows that the set of pairs $(x, y) \in X \times X = [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}}$ for which $y \succ x$ or $x \succ y$ also has measure 0, so the complementary set of pairs (x, y) that are incomparable has measure 1. Theorem 1 asserts that *all* irreflexive preference relations that display intergenerational equity share this property.

As noted above, the existence of an extension of the incomplete relation $\succ$ to a complete relation $\succeq$ depends on Szpilrajn's extension lemma; the proof of the latter result depends on the Axiom of Choice. Theorem 3 asserts that reliance on the Axiom of Choice is unavoidable.[3] The proof of this result relies on Theorem 2, which asserts that no ethical preference relation is measurable. (At the same time as this paper was originally circulated, Lauwers (2006) circulated a very different proof of a similar independence result, but for a restricted domain. Lauwers' work relies on infinite Ramsey theory rather than on measure theory.) Finally, Theorem 4 provides a formal meaning for the idea that reliance on the Axiom of Choice means that no ethical preference relation can be "explicitly described."

The remainder of the paper is organized in the following way. Section 2 presents the formal structure and makes the connections between intergenerational equity and measure for incomplete preference relations (Theorem 1). Section 3 specializes to complete preference relations and shows that no ethical preference relation is measurable (Theorem 2). Section 4 discusses the meaning of independence and of "explicit constructibility" and establishes the independence result (Theorem 3) and the non-constructibility result (Theorem 4). Finally, Section 5 presents similar results (Theorems 1–4) for a restricted domain.

2. Utility streams, preferences and measure

To fix ideas and notation, consider infinite utility streams $x = (x_1, x_2, \dots)$ indexed by time and viewed as utilities of successive generations, so that x_n is the utility of generation n. Write $x \geq y$ if $x_n \geq y_n$ for all n, $x > y$ if $x \geq y$ but $x \neq y$, and $x \gg y$ if $x_n > y_n$ for all n. Write $\mathbb{N}$ for the set of natural numbers (positive integers). I assume here that the possible range of utilities in each period contains the interval $[0,1]$, so the space of utility streams contains $X = [0,1]^{\mathbb{N}}$. (Section 5 addresses preference relations on more

[3] This confirms a conjecture of Fleurbaey and Michel (2003).

restricted domains.) Note that negative results for preference relations on X certainly entail negative results for preference relations on larger domains.

By a *finite permutation* I mean a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ that differs from the identity only on a finite set. Write $\mathbb{F}$ for the group of finite permutations, and $\mathbb{F}_2$ for the subset of permutations that interchange two elements of $\mathbb{N}$ and are the identity elsewhere. For $\sigma \in \mathbb{F}$, define $T_\sigma : X \to X$ by $T_\sigma(x)_n = x_{\sigma(n)}$.

Although we are most interested in complete preference relations, it is convenient to begin by thinking about incomplete preference relations. Say that an irreflexive preference relation $\succ$ on X *displays intergenerational equity* if it is invariant under permutations in $\mathbb{F}$, in the sense that for all $x, y \in X$ and all $\sigma, \tau \in \mathbb{F}$ we have

$$y \succ x \iff T_\tau(y) \succ T_\sigma(x).$$

Because every permutation in $\mathbb{F}$ is the composition of permutations in $\mathbb{F}_2$, a transitive irreflexive relation is invariant under permutations in $\mathbb{F}$ exactly when it is invariant under permutations in $\mathbb{F}_2$, so for transitive relations the definition given is consistent with the discussion in the Introduction.

The Introduction gives one example of an incomplete preference relation that displays intergenerational equity; the literature provides a number of other (perhaps more familiar) examples:

- **Long run averages**

$$x \succ y \iff \liminf \left(\frac{1}{n} \sum_{i=1}^{n} (x_i - y_i) \right) > 0$$

(Aumann and Shapley 1994)

- **Overtaking**

$$x \succ y \iff \liminf \sum_{i=1}^{n} (x_i - y_i) > 0$$

(Rubinstein 1979)

- **Patient limit**

$$x \succ y \iff \exists \delta_0 \in (0,1) \text{ such that } \sum_{i=1}^{\infty} \delta^i x_i > \sum_{i=1}^{\infty} \delta^i y_i \text{ for all } \delta \in (\delta_0, 1).$$

These preference relations share the feature that they are incomplete; indeed, most pairs are incomparable. As shown below, this conclusion is inescapable.

To formalize this statement, we need a natural measure on utility streams. To this end, write λ for Lebesgue measure on (Borel subsets of) $[0,1]$. Let Λ be the infinite product measure on (Borel subsets of) $X = [0,1]^{\mathbb{N}}$, and let $\mathbf{\Lambda} = \Lambda \times \Lambda$ be the product measure on (Borel subsets of) $X \times X = [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}}$. Recall that *inner measure* and *outer measure* are defined for arbitrary subsets $E \subset X \times X$ by

$$\begin{aligned}\boldsymbol{\Lambda}_{\text{in}}(E) &= \sup\{\boldsymbol{\Lambda}(B) : B \text{ Borel, } B \subset E\}\\ \boldsymbol{\Lambda}_{\text{out}}(E) &= \inf\{\boldsymbol{\Lambda}(B) : B \text{ Borel, } B \supset E\}.\end{aligned}$$

Both the sup and inf above are attained; that is, for each $E \subset X \times X$ there are Borel sets $E_{\text{in}} \subset E$ and $E_{\text{out}} \supset E$ such that $\boldsymbol{\Lambda}(E_{\text{in}}) = \boldsymbol{\Lambda}_{\text{in}}(E)$ and $\boldsymbol{\Lambda}(E_{\text{out}}) = \boldsymbol{\Lambda}_{\text{out}}(E)$. The subset $E \subset X \times X$ is *measurable* if its inner and outer measures are equal ($\boldsymbol{\Lambda}_{\text{in}}(E) = \boldsymbol{\Lambda}_{\text{out}}(E)$), in which case we write $\boldsymbol{\Lambda}(E)$ for the common value.

The basic result is the following.

THEOREM 1. *If $\succ$ is an irreflexive preference relation on $X = [0,1]^{\mathbb{N}}$ that displays intergenerational equity then*

$$\boldsymbol{\Lambda}_{\text{out}}\big(\{(x,y) : x \not\succ y \text{ and } y \not\succ x\}\big) = 1.$$

PROOF. Define an inversion

$$\iota : X \times X \to X \times X$$

by $\iota(x,y) = (y,x)$. Note that $\iota \circ \iota$ is the identity, and that ι is measurable and measure-preserving. Write

$$\begin{aligned}R &= \{(x,y) : y \succ x\}\\ L &= \{(x,y) : x \succ y\}\\ I &= \{(x,y) : x \not\succ y \text{ and } y \not\succ x\}.\end{aligned}$$

Evidently R, L, and I are disjoint and their union is $X \times X$. Notice that $\boldsymbol{\Lambda}_{\text{out}}(I) = 1$ exactly when $\boldsymbol{\Lambda}_{\text{in}}(R) = \boldsymbol{\Lambda}_{\text{in}}(L) = 0$. Because ι is measure-preserving and $\iota(R) = L$, it is enough to prove that $\boldsymbol{\Lambda}_{\text{in}}(R) = 0$.

There is a Borel set $R_{\text{in}} \subset R$ such that $\boldsymbol{\Lambda}(R_{\text{in}}) = \boldsymbol{\Lambda}_{\text{in}}(R)$. For each $\sigma, \tau \in \mathbb{F}$, write

$$R_{\text{in}}^{\sigma\tau} = \{(T_\sigma(x), T_\tau(y)) : (x,y) \in R_{\text{in}}\}$$

and note that $R_{\text{in}}^{\sigma\tau}$ is a Borel set. Define

$$A = \bigcup_{\sigma,\tau \in \mathbb{F}} R_{\text{in}}^{\sigma\tau}.$$

Because A is the countable union of Borel sets, it also is a Borel set. Intergenerational equity of $\succ$ guarantees that $R_{\text{in}}^{\sigma\tau} \subset R$ for each σ, τ, so $A \subset R$. The construction guarantees that A is invariant, in the sense that

$$(x,y) \in A,\ \sigma, \tau \in \mathbb{F} \Rightarrow (T_\sigma(x), T_\tau(y)) \in A.$$

Because $R_{\text{in}} \subset A \subset R$ and $\boldsymbol{\Lambda}(R_{\text{in}}) = \boldsymbol{\Lambda}_{\text{in}}(R)$, it follows that $\boldsymbol{\Lambda}(R_{\text{in}}) = \boldsymbol{\Lambda}(A) = \boldsymbol{\Lambda}_{\text{in}}(R)$.

For each $x \in X$, write $A_x = \{y \in X : (x,y) \in A\}$ for the vertical section. Note that each A_x is a Borel set and that the map $x \mapsto \Lambda(A_x)$ is a Borel map. Fubini's theorem implies that

$$\mathbf{\Lambda}(A) = \int_X \Lambda(A_x)\, d\Lambda(x).$$

By construction each vertical section is invariant under each T_σ; that is,

$$y \in A_x, \sigma \in \mathbb{F} \Rightarrow T_\sigma(y) \in A_x.$$

The Hewitt–Savage 0–1 law (Billingsley 1995, p. 496) asserts that every measurable subset of X that is invariant under each T_σ has measure either 0 or 1, so each A_x has measure either 0 or 1.

Now write

$$X_0 = \{x \in X : \Lambda(A_x) = 0\}, \qquad X_1 = \{x \in X : \Lambda(A_x) = 1\}$$

The sets X_0, X_1 are disjoint Borel sets and their union is X (because each A_x has measure 0 or 1), so

$$\mathbf{\Lambda}(A) = \int_{X_0} \Lambda(A_x)\, d\Lambda(x) + \int_{X_1} \Lambda(A_x)\, d\Lambda(x).$$

The first integral is 0 (because the integrand is identically 0) and the second integral is $\Lambda(X_1)$ (because the integrand is identically 1), so $\Lambda(X_1) = \mathbf{\Lambda}(A)$.

On the other hand, suppose $x \in X$ and $\sigma \in \mathbb{F}$. If $y \in A_x$ then by definition $(x,y) \in A$ and our construction guarantees that $(T_\sigma(x), y) \in A$. Hence $A_x \subset A_{T_\sigma(x)}$. Similarly, $A_{T_\sigma(x)} \subset A_{T_{\sigma^{-1}}(T_\sigma(x))} = A_x$. That is, $A_x = A_{T_\sigma(x)}$. In particular, if $x \in X_1$ then $\Lambda(A_{T_\sigma(x)}) = \Lambda(A_x) = 1$ so $T_\sigma(x) \in X_1$. That is, X_1 is invariant under T_σ for each $\sigma \in \mathbb{F}$. Another application of the Hewitt–Savage 0–1 law implies that X_1 has measure either 0 or 1.

Because $\Lambda(X_1) = \mathbf{\Lambda}(A) = \mathbf{\Lambda}_{\text{in}}(R)$, it follows that $\mathbf{\Lambda}_{\text{in}}(R)$ is either 0 or 1. However, $\mathbf{\Lambda}(X \times X) \geq \mathbf{\Lambda}_{\text{in}}(R) + \mathbf{\Lambda}_{\text{in}}(L)$ and $\mathbf{\Lambda}_{\text{in}}(R) = \mathbf{\Lambda}_{\text{in}}(R)$. Thus, if $\mathbf{\Lambda}_{\text{in}}(R) = 1$ then $\mathbf{\Lambda}(X \times X) \geq 2$. Because $\mathbf{\Lambda}$ is a probability measure, this is impossible, so we conclude that $\mathbf{\Lambda}_{\text{in}}(R) = \mathbf{\Lambda}_{\text{in}}(L) = 0$ and hence that $\mathbf{\Lambda}_{\text{out}}(I) = 1$, as asserted. □

3. Non-measurability

We now turn our attention to complete, transitive preference relations. Say that a complete transitive preference relation $\succeq$ *displays intergenerational equity* exactly when its irreflexive part $\succ$ does so. In view of our earlier discussion, this means that for all $x, y \in X$ and all $\sigma, \tau \in \mathbb{F}$ (or just all $\sigma, \tau \in \mathbb{F}_2$) we have

$$y \succ x \iff T_\tau(y) \succ T_\sigma(x).$$

It is easily checked that $\succeq$ displays intergenerational equity if and only if it has the property

$$x \in X, \sigma \in \mathbb{F}_2 \Rightarrow x \sim T_\sigma(x).$$

A preference relation *respects the weak Pareto ordering* if $y \gg x$ implies $y \succ x$. A complete transitive (weak) preference relation on X is *ethical* if it displays intergenerational equity and respects the weak Pareto ordering.

Say that the preference relation $\succeq$ is *measurable* if its graph

$$\mathscr{G} = \{(x,y) \in X \times X : y \succeq x\}$$

is a measurable subset of $X \times X$.

THEOREM 2. *No ethical preference relation on $X = [0,1]^{\mathbb{N}}$ is measurable.*

PROOF. Suppose, to the contrary, $\succeq$ is an ethical preference relation on $[0,1]^{\mathbb{N}}$ whose graph $\mathscr{G}$ is measurable. As in the proof of Theorem 1, let $\iota : X \times X \to X \times X$ be the inversion $\iota(x,y) = (y,x)$ and define

$$\begin{aligned} R &= \{(x,y) : y \succ x\} \\ L &= \{(x,y) : x \succ y\} \\ I &= \{(x,y) : x \not\succ y \text{ and } y \not\succ x\} = \{(x,y) : x \sim y\}. \end{aligned}$$

Note that $I = \mathscr{G} \cap \iota(\mathscr{G})$, $R = \mathscr{G} \setminus I$, and $L = \iota(R)$, so that R, L, and I are all measurable. It follows from Theorem 1 that $\mathbf{\Lambda}(R) = \mathbf{\Lambda}(L) = 0$ and $\mathbf{\Lambda}(I) = 1$.

Proceeding as in the proof of Theorem 1, construct a Borel set $J \subset I$ such that $\mathbf{\Lambda}(J) = 1$ and

$$(x,y) \in J, \sigma, \tau \in \mathbb{F} \Rightarrow (T_\sigma(x), T_\tau(y)) \in J.$$

Write J_x for the vertical section. Fubini's theorem guarantees that

$$\mathbf{\Lambda}(J) = \int_X \Lambda(J_x)\, dx.$$

Because $\mathbf{\Lambda}(J) = 1$, it follows that $\Lambda(J_x) = 1$ for almost every $x \in X$; fix any $x^* \in X$ for which $\Lambda(J_{x^*}) = 1$.

Choose and fix a sequence $\{b_n\}$ of real numbers such that

$$0 < b_n < 1 \text{ and } \prod_{n=1}^{\infty} b_n > \tfrac{1}{2}.$$

Set

$$D = \prod_{n=1}^{\infty} [0, b_n]$$

and define a map $f : D \to X$ by

$$f(x)_n = x_n + \frac{1 - b_n}{2}.$$

It is easily checked that f is one-to-one, measurable, and measure-preserving; hence $\Lambda(f(D\cap J_{x^*}))=\Lambda(D\cap J_{x^*})$. Our construction guarantees that $\Lambda(J_{x^*})=1$ and $\Lambda(D)>\frac{1}{2}$, so

$$\Lambda(D)=\Lambda(D\cap J_{x^*})=\Lambda(f(D\cap J_{x^*}))>\tfrac{1}{2}.$$

However, if $y\in(D\cap J_{x^*})$ then $f(y)\gg y$, so $f(y)\succ y$, whence $f(y)\succ x^*$. Thus $(D\cap J_{x^*})$ and $f(D\cap J_{x^*})$ are disjoint. Hence

$$\Lambda(X)\geq\Lambda\big((D\cap J_{x^*})\cup f(D\cap J_{x^*})\big)=\Lambda\big(D\cap J_{x^*}\big)+\Lambda\big(f(D\cap J_{x^*})\big)>\tfrac{1}{2}+\tfrac{1}{2}.$$

Because Λ is a probability measure, we have reached a contradiction, so the proof is complete. □

4. Independence and definability

To understand what independence means, it is useful to recall some mathematical logic.[4] A *language* L consists of the usual logical symbols together with a set of constant symbols, a set of relation symbols, and a set of function symbols. Given a set of sentences $\mathscr{A}$ in the language L, we can use the usual logical rules of inference to derive/prove other sentences. A set of sentences $\mathscr{A}$ is *consistent* if no contradiction can be derived from it; that is, there is no sentence Q with the property that both Q and the negation of Q can be derived from $\mathscr{A}$.

A *model* for the language L consists of an underlying set $\mathbb{M}$ and an *interpretation* of all the symbols of L. The Completeness Theorem asserts that a set of sentences $\mathscr{A}$ in the language L is consistent if and only if there is a model $\mathbb{M}$ of L in which all the sentences of $\mathscr{A}$ are true. Assuming $\mathscr{A}$ is consistent, a sentence P is provable from $\mathscr{A}$ if and only if P is true in every model, and the negation of P is provable from $\mathscr{A}$ if and only if P is false in every model. The sentence P is *independent of* $\mathscr{A}$ (or *undecidable on the basis of* $\mathscr{A}$) if neither P nor its negation can be derived from $\mathscr{A}$. Thus, P is independent of $\mathscr{A}$ if and only if there is a model in which the sentences of $\mathscr{A}$ are true and P is true and a model in which the sentences of $\mathscr{A}$ are true but P is false.

To make all this concrete, consider the familiar language of Euclidean geometry and the sentences that constitute the familiar axioms and postulates of Euclidean geometry. The ordinary plane, with the usual interpretations, is a model of this language, and the familiar axioms and postulates are true in this model, so Euclidean geometry is consistent. A different model (spherical geometry) is obtained by taking the sphere (rather than the plane) as the underlying set, and interpreting a "line" as a great circle on the surface of the sphere. In spherical geometry, all the usual axioms and postulates of Euclidean geometry are true *except* the Parallel Postulate, which asserts that through every point not on a given line there is exactly one line parallel to the given line. (Through

[4] Chang and Keisler (1992), Jech (1978), and Weiss and D'Mello (1997) provide convenient references; the last is especially accessible—and freely downloadable.

any point not on a given great circle, there are *no* great circles parallel to the given great circle.) Hence the Parallel Postulate is *independent* of the other axioms and postulates.[5]

An example more relevant to the present work concerns the existence of non-measurable sets. A familiar construction (due to Vitali) shows that the Zermelo–Fraenkel Axioms together with the Axiom of Choice imply the existence of sets of real numbers that are not Lebesgue measurable. In particular, it follows that in every model of set theory in which the Zermelo–Fraenkel Axioms and the Axiom of Choice are true, there exist non-measurable sets.

However, the full power of the Axiom of Choice is almost never used in formal economics or in classical analysis for that matter. What is used is a much weaker axiom, the Axiom of Dependent Choice. The formal statements of these axioms follow.

Axiom of Choice Let $\mathscr{X}$ be a non-empty family of non-empty sets. Then there is a function

$$f : \mathscr{X} \to \bigcup_{X \in \mathscr{X}} X$$

such that $f(X) \in X$ for each $X \in \mathscr{X}$.

Axiom of Dependent Choice Let X be a non-empty set and let $\mathscr{R}$ be an entire relation on X (that is, a binary relation such that for each $x \in X$ there is a $y \in X$ with $x \mathscr{R} y$). Then there is a function $f : \mathbb{N} \to X$ such that $f(n) \mathscr{R} f(n+1)$ for each $n \in \mathbb{N}$.

To understand the difference between these two axioms, note that if we are given a set X and an entire relation $\mathscr{R}$ on X, the Zermelo–Fraenkel Axioms by themselves guarantee that for every natural number N there is a function $f : \{1, \dots, N\} \to X$ for which $f(n) \mathscr{R} f(n+1)$ for $n = 1, \dots, N-1$. That is, the Zermelo–Fraenkel Axioms—without any additional choice axiom whatsoever—already guarantee the existence of *arbitrarily long finite sequences* for which $f(n) \mathscr{R} f(n+1)$; the Axiom of Dependent Choice merely guarantees the existence of an *infinite sequence* for which $f(n) \mathscr{R} f(n+1)$. By contrast, the Axiom of Choice guarantees the existence of an *arbitrarily large family* of simultaneous choices.

The Axiom of Dependent Choice is exactly what is needed to carry out many of the familiar constructions of classical analysis—for example, the proof that a bounded sequence of real numbers contains a convergent subsequence.

As noted above, the Zermelo–Fraenkel Axioms together with the Axiom of Choice imply the existence of non-measurable sets. On the other hand, Solovay (1970) constructs a model $\mathbb{M}_1$ of set theory in which the Zermelo–Fraenkel Axioms and the Axiom of Dependent Choice are true and in which all sets of real numbers are Lebesgue measurable.[6] Hence the existence of a non-measurable set of real numbers is independent of the Zermelo–Fraenkel Axioms together with the Axiom of Dependent Choice.

---

[5]The reader may recall that, until the early part of the 19th century, many mathematicians believed that the Parallel Postulate could be derived from other axioms and postulates. Spherical geometry, which was invented by Bolyai, and hyperbolic geometry, which was invented by Lobachevsky, show that this belief was mistaken.

[6]Solovay's construction, like many constructions in model theory, is relative, in the sense that it assumes

The existence of non-measurable sets is relevant in the present context because Theorem 2 guarantees that no ethical preference relation on $X = [0,1]^{\mathbb{N}}$ is measurable. Put differently, the existence of an ethical preference relation on $X = [0,1]^{\mathbb{N}}$ entails the existence of a non-measurable subset of $X \times X$, and hence, as we shall show, the existence of a non-measurable subset of $\mathbb{R}$. It follows that in Solovay's model $\mathbb{M}_1$ there does not exist an ethical preference relation on X, and hence that the existence of ethical preference relations on X is independent. The formal statement and proof are as follows.

THEOREM 3. *The existence of an ethical preference relation on* $X = [0,1]^{\mathbb{N}}$ *is independent of the Zermelo–Fraenkel Axioms and the Axiom of Dependent Choice.*

PROOF. We need to exhibit two models: (i) a model in which the Zermelo–Fraenkel Axioms and the Axiom of Dependent Choice are true and in which there exists an ethical preference relation on $X = [0,1]^{\mathbb{N}}$; (ii) a model in which the Zermelo–Fraenkel Axioms and the Axiom of Dependent Choice are true and in which there does not exist an ethical preference relation on $X = [0,1]^{\mathbb{N}}$. In view of Svensson's theorem, for the first model we can take any model in which the Zermelo–Fraenkel Axioms and the Axiom of Choice are true. For the second model, we take Solovay's (1970) model $\mathbb{M}_1$.

To show that in $\mathbb{M}_1$ there is no ethical preference relation on $[0,1]^{\mathbb{N}}$, suppose to the contrary that such a relation $\succeq$ exists. In view of Theorem 2, the graph $\mathscr{G}$ of $\succeq$ is a non-measurable subset of $[0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}}$. To obtain a contradiction, it is enough to show that this implies the existence of a non-measurable subset of $\mathbb{R}$. Although this is an immediate consequence of a theorem of Maharam (1942) characterizing complete non-atomic probability spaces, an elementary argument (which will also be useful later) is available for the case at hand.

For each $s \in [0,1]$, write $s = .s_1 s_2 \ldots$ for its binary expansion. Define a map $\Phi : [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}} \to [0,1]$ by

$$\Phi\big((s^i),(t^i)\big) = .s^1_1 t^1_1 s^1_2 t^1_2 s^2_1 t^2_1 s^2_2 t^2_2 \ldots$$

Let F be the set of elements of $[0,1]$ whose binary expansions are eventually constant (i.e., end in an infinite string of 0's or an infinite string of 1's); note that F is a countable set. Set $X^* = ([0,1] \setminus F)^{\mathbb{N}}$, $Y^* = \Phi(X^* \times X^*)$, and $G = [0,1] \setminus Y^*$; let Φ^* be the restriction of Φ to $X^* \times X^*$. Note that G consists of those elements $s \in [0,1]$ for which one of a countable number of particular subsequences of the binary expansion of s ends in an infinite string of 0's or an infinite string of 1's; hence $\lambda(G) = 0$ and $\lambda(Y^*) = 1$. Because F is countable, $\Lambda(F) = 0$ and $\mathbf{\Lambda}((X \times X) \setminus (X^* \times X^*)) = 0$.

It is easily seen that Φ^* is one-to-one, onto, measurable with a measurable inverse, and measure-preserving. If $\mathscr{G} \subset X \times X$ is not measurable then neither is $\mathscr{G}^* =$

the existence of a model of one set of axioms and constructs a model of another set of axioms. Such relativism is unavoidable, because a theorem of Gödel shows that no reasonably strong system of axioms can prove its own consistency. In this case, Solovay assumes the existence of a model in which the Zermelo–Fraenkel Axioms, the Axiom of Choice, and a "large cardinal axiom" are true and deduces the existence of a model in which the Zermelo–Fraenkel Axioms and the Axiom of Dependent Choice are true and in which all sets of real numbers are Lebesgue measurable.

$\mathscr{G} \cap ((X \times X) \setminus (X^* \times X^*))$. Hence $\Phi^*(\mathscr{G}^*)$ is a non-measurable subset of Y^*, and thus of $[0,1]$.

To summarize: the existence of an ethical preference relation on $X = [0,1]^{\mathbb{N}}$ entails the existence of a non-measurable subset of $[0,1]$. Hence, in Solovay's model $\mathbb{M}_1$, no such preference relation exists, as was asserted. □

What does it mean to say that no ethical preference relation on $[0,1]^{\mathbb{N}}$ can be "explicitly described"? An informal interpretation is that any proof of the existence of such a preference relation must rely on the Axiom of Choice—but a formal interpretation of the same idea can be based on the notion of definability.

A set $A \subset \mathbb{R}$ is *definable* if there is a set-theoretic formula $\Phi(t, r, \alpha_1, \ldots,)$ in which r is a real number, $\alpha_1, \ldots,$ is a sequence of ordinals, and t is the only free variable, such that

$$A = \{t \in \mathbb{R} : \Phi(t, r, \alpha_1, \ldots)\}.$$

Solovay (1970) shows that there is a model $\mathbb{M}_2$ of the Zermelo–Fraenkel Axioms together with the Axiom of Choice in which every set of reals that is definable is Lebesgue measurable. In the model $\mathbb{M}_2$ the Axiom of Choice is true, so in $\mathbb{M}_2$ there are subsets of $\mathbb{R}$ that are non measurable—but such sets are not definable. Because there is a model in which the Zermelo–Fraenkel Axioms together with the Axiom of Choice are true and all definable sets are Lebesgue measurable, it follows that the proposition "there exists a definable set that is not Lebesgue measurable" is not provable from the Zermelo–Fraenkel Axioms together with the Axiom of Choice. In different words: no definable set can be proved (on the basis of the Zermelo–Fraenkel Axioms and the Axiom of Choice) to be non-measurable.

Similarly, say that a set $B \subset [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}}$ is *definable* if there is a set-theoretic formula $\Phi(t, r, \alpha_1, \ldots)$ in which r is a real number, $\alpha_1, \ldots$ is a sequence of ordinals, and t is the only free variable, such that

$$B = \{(x, y) \in [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}} : \Phi((x, y), r, \alpha_1, \ldots)\}.$$

Say that a preference relation $\succeq$ is *definable* if its graph

$$\mathscr{G} = \{(x, y) \in [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}} : y \succeq x\}$$

is a definable set. The next result shows that no definable preference relation on $X = [0,1]^{\mathbb{N}}$ can be proved to be ethical; in this sense, at least, no ethical preference relation on $X = [0,1]^{\mathbb{N}}$ can be "explicitly constructed."

THEOREM 4. *No definable preference relation on $X = [0,1]^{\mathbb{N}}$ can be proved (on the basis of the Zermelo–Fraenkel Axioms and the Axiom of Choice) to be ethical.*

PROOF. Suppose that $\succeq$ is a definable preference relation on $X = [0,1]^{\mathbb{N}}$ and that $\succeq$ can be proved (on the basis of the Zermelo–Fraenkel Axioms and the Axiom of Choice) to be ethical. By Theorem 2, the graph $\mathscr{G}$ of $\succeq$ is a non-measurable subset of $X \times X$. The argument of Theorem 3 together with a proof that $\succeq$ is ethical and the proof of Theorems 2 and 3 yield a definable subset $\Phi^*(\mathscr{G}^*)$ of $\mathbb{R}$ *and* a proof that that $\Phi^*(\mathscr{G}^*)$ is not Lebesgue measurable. As we have noted above, if follows from Solovay (1970) that this is impossible. We have reached a contradiction so the proof is complete. □

5. Restricted domain

A natural response to impossibility results in social choice is to look for restricted domains on which possibility is restored. In the present instance, one might require that each generation's utility lies in a finite set, say $\{0,1\}$; see Basu and Mitra (2003) and Basu and Mitra (forthcoming) for instance. If we make this restriction, then the relevant space of utility streams is $Z = \{0,1\}^{\mathbb{N}}$. However, the weak Pareto ordering is almost degenerate on Z: if $w, z \in Z$ and $w \gg z$ then $w = \mathbf{1}$ and $z = \mathbf{0}$ (the constant sequences of all 1's and all 0's, respectively). Hence a complete transitive preference relation $\succeq$ on Z that displays intergenerational equity and respects the weak Pareto ordering can be defined by

$$\mathbf{1} \succ z \text{ for } z \neq \mathbf{1}$$
$$w \sim z \text{ for } w, z \neq \mathbf{1}.$$

However, if we insist, as seems natural in this context, that preferences respect the *strong* Pareto ordering, then conclusions parallel to those of Theorems 1–4 again emerge.

As for preference relations on X, say that an irreflexive (incomplete) preference relation $\succ$ on Z *displays intergenerational equity* if it is invariant under permutations in $\mathbb{F}$, in the sense that for all $z, w \in Z$ and all $\sigma, \tau \in \mathbb{F}$ we have

$$w \succ z \iff T_\tau(w) \succ T_\sigma(z).$$

As before, a transitive irreflexive relation is invariant under permutations in $\mathbb{F}$ exactly when it is invariant under permutations in $\mathbb{F}_2$. Say that a complete transitive preference relation $\succeq$ on Z *displays intergenerational equity* exactly when its irreflexive part $\succ$ does. Equivalently, $\succeq$ *displays intergenerational equity* if and only if it has the property

$$z \in Z, \sigma \in \mathbb{F}_2 \Rightarrow z \sim T_\sigma(z).$$

The preference relation $\succeq$ *respects the strong Pareto ordering* if $y > x$ implies $y \succ x$. Say that a complete transitive preference relation on X is *ethical* if it displays intergenerational equity and respects the strong Pareto ordering.

Let γ be normalized counting measure on $\{0,1\}$, let Γ be the infinite product measure on $Z = \{0,1\}^{\mathbb{N}}$, and let $\mathbf{\Gamma} = \Gamma \times \Gamma$ be the product measure on $Z \times Z$. As before, say that a preference relation $\succeq$ on Z is *measurable* if its graph

$$\mathscr{G} = \{(z, w) \in Z \times Z : w \succeq z\}$$

is measurable with respect to $\mathbf{\Gamma}$ (i.e., its inner and outer measure coincide).

The following results are the parallels to Theorems 1–4 for relations on the restricted domain $Z = \{0,1\}^{\mathbb{N}}$; because the proofs closely parallel the proofs of Theorems 1–4, I give only sketches.

THEOREM 1′. *If $\succ$ is an irreflexive preference relation on $Z = \{0,1\}^{\mathbb{N}}$ that displays intergenerational equity then*

$$\mathbf{\Gamma}_{\text{out}}\big(\{(z,w) : z \not\succ w \text{ and } w \not\succ z\}\big) = 1.$$

PROOF. The proof is almost the same as the proof of Theorem 1. Define an inversion

$$\iota : Z \times Z \to Z \times Z$$

by $\iota(z,w) = (w,z)$. Note that $\iota \circ \iota$ is the identity, and that ι is measurable and measure-preserving. Write

$$\begin{aligned} R &= \{(z,w) : w \succ z\} \\ L &= \{(z,w) : w \succ z\} \\ I &= \{(z,w) : z \not\succ w \text{ and } w \not\succ z\}. \end{aligned}$$

Evidently, R, L, and I are disjoint and their union is $Z \times Z$. As in the proof of Theorem 1, it is enough to prove that $\mathbf{\Lambda}_{\text{in}}(R) = 0$.

As in the proof of Theorem 1, construct a Borel set $A \subset R$ that is invariant, in the sense that

$$(z,w) \in A,\ \sigma, \tau \in \mathbb{F} \Rightarrow (T_\sigma(z), T_\tau(w)) \in A.$$

As in the proof of Theorem 1, $\mathbf{\Lambda}(R_{\text{in}}) = \mathbf{\Lambda}(A) = \mathbf{\Lambda}(R)$.

For each $z \in Z$, write $A_z = \{w \in X : (z,w) \in A\}$ for the vertical section. By construction each vertical section is invariant under each T_σ. Use Fubini's theorem and the Hewitt–Savage 0–1 law to show that each A_z has measure either 0 or 1.

Now write

$$Z_0 = \{z \in Z : \Gamma(A_z) = 0\}, \qquad Z_1 = \{z \in Z : \Gamma(A_z) = 1\}$$

The sets Z_0, Z_1 are disjoint Borel sets and their union is Z, so

$$\mathbf{\Gamma}(A) = \int_{Z_0} \Gamma(A_z)\, d\Gamma(z) + \int_{Z_1} \Gamma(A_z)\, d\Gamma(z)$$

The first integral is 0 (because the integrand is identically 0) and the second integral is $\Gamma(Z_1)$ (because the integrand is identically 1), so $\Gamma(Z_1) = \mathbf{\Gamma}(A)$.

Now, arguing exactly as in Theorem 1, we can show that Z_1 is invariant under T_σ for each $\sigma \in \mathbb{F}$. Another application of the Hewitt–Savage 0–1 law implies that Z_1 has measure either 0 or 1, and hence that R has inner measure either 0 or 1. Because R and L have the same inner measure and Γ is a probability measure, it follows that R has inner measure 1, as asserted. □

THEOREM 2′. *No ethical preference relation on $Z = \{0,1\}^{\mathbb{N}}$ is measurable.*

PROOF. Suppose to the contrary that $\succeq$ is a measurable, ethical preference relation on Z. Write

$$I = \{(z,w) \in Z \times Z : w \sim z\}.$$

Arguing exactly as in the proof of Theorem 2, we can construct a Borel subset $J \subset I$ such that $\mathbf{\Gamma}(J) = \mathbf{\Gamma}(I)$ and $(w,z) \in J$ implies $(T_\sigma(w), T_\tau(w)) \in J$ for each $\sigma, \tau \in \mathbb{F}$. Continuing to argue as in the proof of Theorem 2, we can show that $\mathbf{\Gamma}(J) = 1$, and then that there is some $z^* \in Z$ for which $\Gamma(J_{z^*}) = 1$.

Define a map $g : Z \to Z$ by

$$g(z)_n = \begin{cases} 0 & \text{if } n = 1 \text{ and } z_1 = 1 \\ 1 & \text{if } n = 1 \text{ and } z_1 = 0 \\ z_n & \text{if } n \neq 1. \end{cases}$$

(That is, g reverses the first component of z.) It is evident that g is one-to-one, measurable and measure-preserving. By assumption, $\succeq$ respects the strong Pareto ordering, so if $z_1 = 0$ then $g(z) > z$ and $g(z) \succ z$, while if $z_1 = 1$ then $g(z) < z$ and $z \succ g(z)$; in either case, $g(z) \not\sim z$. Hence, if $z \in J_{z^*}$ then $g(z) \not\sim z^*$, so $g(z) \notin J_{z^*}$. It follows that $g(J_{z^*}) \cap J_{z^*} = \emptyset$. Because g is measure-preserving and $\Gamma(J_{z^*}) = 1$, this is absurd, so we have reached a contradiction and the proof is complete. □

THEOREM 3′. *The existence of an ethical preference relation on $Z = \{0,1\}^{\mathbb{N}}$ is independent of the Zermelo–Fraenkel Axioms and the Axiom of Dependent Choice.*

PROOF. As in the proof of Theorem 3, it suffices to show that the existence of a non-measurable subset of $Z \times Z$ entails the existence of a non-measurable subset of $[0,1]$.

Note first that the map $(z,w) \mapsto (z_1, w_1, z_2, w_2, \ldots)$ is a one-to-one, onto, and measure-preserving map of $Z \times Z$ to Z. Hence the existence of a non-measurable subset of $Z \times Z$ entails the existence of a non-measurable subset of Z. Now let $F \subset Z$ be the set of sequences that are eventually constant (either eventually 0 or eventually 1) and let $G \subset [0,1]$ be the set of real numbers that have a binary expansion that ends in all 0's or all 1's. Set $Z^* = Z \setminus F$, and define a map $h : Z^* \to [0,1] \setminus G$ by

$$h(z) = \sum_{n=1}^{\infty} z_n 2^{-n}.$$

That is, h maps the sequence z of 0's and 1's to the real number having z as its binary expansion. It is easily seen that h is one-to-one, measurable (in fact continuous), and measure-preserving, and that its inverse is also measurable. Because F and G are countable, $\Gamma(F) = 0$ and $\lambda(G) = 0$. Hence the existence of a non-measurable subset of Z entails the existence of a non-measurable subset of $[0,1]$, as required. □

THEOREM 4′. *No definable preference relation on $Z = \{0,1\}^{\mathbb{N}}$ can be proved (on the basis of the Zermelo–Fraenkel Axioms and the Axiom of Choice) to be ethical.*

PROOF. Suppose $\succeq$ is a definable preference preference relation on $Z = \{0,1\}^{\mathbb{N}}$. By Theorem 2′, the graph of $\succeq$ is a non-measurable subset of $Z \times Z$. A proof that the graph of $\succeq$ is definable, together with the proofs of Theorems 2′ and 3′ provide a definable subset of $\mathbb{R}$ *and* a proof that this set is non-measurable. In view of Solovay (1970), this is impossible. □

References

Anderson, Robert M. (1991), "Non-standard analysis with applications to economics." In *Handbook of Mathematical Economics, Volume 4* (Werner Hildenbrand and Hugo Sonnenschein, eds.), 2145–2208, North-Holland, Amsterdam. [188]

Anderson, Robert M. and Roberto C. Raimondo (2005), "Market clearing and derivative pricing." *Economic Theory*, 25, 21–34. [188]

Aumann, Robert J. and Lloyd S. Shapley (1994), "Long-term competition—a game-theoretic analysis." In *Essays in Game Theory* (Nimrod Megiddo, ed.), 1–15, Springer-Verlag, New York. [190]

Basu, Kaushik and Tapan Mitra (2003), "Aggregating infinite utility streams with intergenerational equity: The impossibility of being Paretian." *Econometrica*, 71, 1557–1563. [188, 198]

Basu, Kaushik and Tapan Mitra (forthcoming), "Possibility theorems for equitably aggregating infinite utility systems." In *Intergenerational Equity and Sustainability* (John E. Roemer and Kotaro Suzumura, eds.), Palgrave Macmillan, Basingstoke. [188, 198]

Billingsley, Patrick (1995), *Probability and Measure*, third edition. Wiley, New York. [192]

Bossert, Walter, Yves Sprumont, and Kotaro Suzumura (forthcoming), "Ordering infinite utility streams." *Journal of Economic Theory*. [188]

Chang, Chen Chung and Jerome H. Keisler (1992), *Model Theory*, third edition. North-Holland, Amsterdam. [194]

Diamond, Peter A. (1965), "The evaluation of infinite utility streams." *Econometrica*, 33, 170–177. [188]

Fleurbaey, Marc and Philippe Michel (2003), "Intertemporal equity and the extension of the Ramsey criterion." *Journal of Mathematical Economics*, 39, 777–802. [188, 189]

Hara, Chiaki, Tomoichi Shinotsuka, Kotaro Suzumura, and Yongsheng Xu (2006), "On the possibility of continuous, Paretian and egalitarian evaluation of infinite utility streams." Unpublished paper, Kyoto Institute of Economic Research, Kyoto University. [188]

Jech, Thomas J. (1978), *Set Theory*. Academic Press, New York. [194]

Lauwers, Luc (2006), "Ordering infinite utility streams: Completeness at the cost of a non-Ramsey set." Unpublished paper, Katholieke Universiteit Leuven. [189]

Maharam, Dorothy (1942), "On homogeneous measure algebras." *Proceedings of the National Academy of Sciences of the United States of America*, 28, 108–111. [196]

Ramsey, Frank P. (1928), "A mathematical theory of saving." *Economic Journal*, 38, 543–559. [187]

Rawls, John (1971), *A Theory of Justice*. Harvard University Press, Cambridge, Massachusetts. [187]

Rubinstein, Ariel (1979), "Equilibrium in supergames with the overtaking criterion." *Journal of Economic Theory*, 21, 1–9. [190]

Sidgwick, Henry (1907), *The Methods of Ethics*, seventh edition. Macmillan, London. [187]

Solovay, Robert M. (1970), "A model of set-theory in which every set of reals is lebesgue measurable." *Annals of Mathematics*, 92, 1–56. [195, 196, 197, 200]

Svensson, Lars-Gunnar (1980), "Equity among generations." *Econometrica*, 48, 1251–1256. [188]

Szpilrajn, Edward (1930), "Sur l'extension de l'ordre partiel." *Fundamenta Mathematicae*, 16, 386–389. [189]

Weiss, William and Cherie D'Mello (1997), *Fundamentals of Model Theory*. Unpublished book, Department of Mathematics, University of Toronto. [194]

Submitted 2006-12-28. Final version accepted 2007-5-17. Available online 2007-5-17.