DÉPARTEMENT DE SCIENCE ÉCONOMIQUE
DEPARTMENT OF ECONOMICS

CAHIERS DE RECHERCHE / WORKING PAPERS

0404E

Classical (Generalized) Utilitarianism and the Repugnant Conclusion

by
Leslie Shiell

ISSN: 0225-3860

uOttawa

Faculté des sciences sociales
Faculty of Social Sciences

CP 450 SUCC. A
OTTAWA (ONTARIO)
CANADA K1N 6N5

P.O. BOX 450 STN. A
OTTAWA, ONTARIO
CANADA K1N 6N5

The Repugnant Conclusion on Realistic Choice Sets

by

Leslie Shiell

May 2005

University of Ottawa
Department of Economics
200 Wilbrod Street
Ottawa, Canada

phone (613) 562-5800 x1693
fax (613) 562-5999
email: lshiell@uottawa.ca

Abstract

Both classical and critical-level generalized utilitarianism (CGU and CLGU) exhibit the repugnant conclusion on an unrestricted choice set when repugnance is defined in terms of the critical level. However, contrary to common belief, the repugnant conclusion is not an inherent feature of utilitarian population principles. Rather, it results from the interaction of the principles with certain choice sets. Choice sets can be assessed in terms of their realism, defined as conformity with universal properties of physics, biology and preferences. It is shown that these properties entail a particular structure of choice set on which CGU and CLGU do not exhibit the repugnant conclusion.

(JEL D63, D71; keywords: variable-population social choice, utilitarianism, repugnant conclusion)

I. **Introduction**

Like its static counterpart, variable-population social choice seeks to find rules for choosing among a number of social alternatives. The defining feature in this case is that the population may vary among alternatives. It is perhaps not surprising that the naïve application of static social choice rules in this context may lead to unexpected results. Therefore, a literature has developed to address these challenges and some impossibility results have emerged.

In order to make meaningful comparisons in the variable-population context, it is necessary to define a neutral level of well-being. Blackorby et al. (1995, 2002b, 2003) define neutrality such that an individual who experiences well-being at this level is indifferent between her actual life and a hypothetical life devoid of experiences. An individual who experiences well-being above this level prefers her life to a life without experiences, while an individual who experiences well-being below this level would prefer a life without experiences. In economists' vocabulary, well-being is referred to as utility, and it is conventional to fix the neutral level of utility at a value of zero.

Blackorby and Donaldson (1984) have also introduced the concept of the critical level which is the analogue to neutrality for the social planner. In particular, the critical level is the value of utility such that the addition to an existing population of an individual with utility at this level, when nothing else changes, leaves society indifferent between the new situation and the old.

The main impossibility results in this literature concern the repugnant conclusion and mere addition principle of Parfit (1976, 1982, 1984), the Pareto plus principle of Sikora (1978), and the strong sadistic conclusion of Arrhenius (2000). A social choice rule exhibits the repugnant conclusion if, given any population and utility allocation, there is always a better alternative in which everyone receives a utility level arbitrarily close to neutrality and the population is larger. The repugnant conclusion describes a choice rule which is always disposed to sacrifice well-being in exchange for some increase in population. The undesirability of this property seems evident and therefore most authors have accepted Parfit's position that a good social choice rule should avoid it.

The mere addition principle (Pareto plus principle) states that social welfare is not decreased (is increased) by the addition to an existing population of an individual with utility above neutrality, when nothing else changes. This principle conveys an aspect of non-paternalism which seems appealing, and therefore one may desire that a social choice rule exhibit it as well.

The strong sadistic conclusion entails that, for any population consisting of people who all have utility below neutrality, there is another population whose existence would be worse although all its members have utility above neutrality. The undesirability of this axiom also seems evident, and therefore one may hold that a social choice rule should avoid it.

The present paper focuses on welfarist social choice rules. A social choice rule is welfarist if it takes into account individuals' well-being or utility levels only and disregards how the utility levels are achieved. In this case, the rule is said to generate a single social ordering, defined over utility levels only, where an ordering is defined as a complete, reflexive and transitive set of binary relations which rank all social alternatives. In the variable-population context, social orderings are frequently referred to as population principles, since they involve ranking alternatives with different populations. If continuity is assumed, the social ordering can be represented by a social welfare function.

Blackorby, Bossert and Donaldson (2003) identify the following axioms, in addition to welfarism, which they argue a good variable-population social ordering should satisfy: strong Pareto, continuity, anonymity, existence of critical levels, existence independence,[1] non-negative critical levels, avoidance of the repugnant conclusion, and avoidance of the strong sadistic conclusion. They report that there is no population principle that satisfies all of these axioms. Similarly, Ng (1989) notes that the combination of the mere addition principle and a principle called non-anti-egalitarianism implies the repugnant conclusion.

Blackorby and Donaldson (1984) and Blackorby, Bossert and Donaldson (1995) develop the class of critical-level utilitarian welfare functions which aggregate deviations of utility from the critical level. The authors argue in favour of positive critical levels (i.e. above neutrality) in order to avoid the repugnant conclusion. Unfortunately, as noted in Blackorby et al. (2002a and 1997), positive critical levels are incompatible with mere addition and Pareto plus. A general investigation of this conflict is provided in Blackorby, Bossert, Donaldson and Fleurbaey (1998).

Comparison of the impossibility results is facilitated by noting the similarity of the Pareto-plus principle and avoidance of the strong sadistic conclusion. The two axioms are equivalent under generalized utilitarianism. Specifically, it can be shown that a generalized utilitarian welfare function that satisfies strong Pareto, anonymity, and non-negative critical levels exhibits the strong sadistic conclusion if and only if the critical level is positive, in which case the function also violates Pareto-plus. For simplicity, the discussion below emphasizes Pareto-plus rather than the strong sadistic conclusion.

Blackorby et al. (2002b and 2003) show that, of 13 welfarist population principles examined, only two – classical generalized utilitarianism (CGU) and critical-level generalized utilitarianism (CLGU) – satisfy the axioms of strong Pareto, continuity, anonymity, existence of

[1] Referred to as extended independence of the utilities of unconcerned individuals in Blackorby et al. (2002b).

critical levels, non-negative critical levels, and existence independence. Both CGU and CLGU welfare aggregate transformed utilities, in which the transformation allows for aversion to inequality of utility. The term classical indicates a unique critical level equal to individual neutrality, while for CLGU the critical level is strictly greater than individual neutrality.

Classical generalized utilitarianism avoids the strong sadistic conclusion (satisfies Pareto-plus) but fails to avoid the repugnant conclusion. In contrast, critical-level generalized utilitarianism avoids the repugnant conclusion but exhibits the strong sadistic conclusion (fails to satisfy Pareto plus). Faced with this dilemma, Blackorby et al. argue that avoiding the repugnant conclusion is more important than avoiding the strong sadistic conclusion (satisfying Pareto-plus), and therefore they advocate CLGU.

The situation is changed if the repugnant conclusion is based on the critical level rather than neutrality. In this case, a social choice rule exhibits the repugnant conclusion if there is always a better alternative in which everyone receives a utility level arbitrarily close to the *critical level* and population is larger. It will be argued that this definition is compelling due to the similarity between the critical level and neutrality.

Under this revised definition, neither CGU nor CLGU avoid the repugnant conclusion. One might be led to reject all forms of utilitarianism as a consequence, if one believes that the repugnant conclusion is truly repugnant. At the other extreme, those who do not share this belief will find this development unimportant. The present paper begins from an intermediate position, with the view that there is value in probing into the deeper structure of repugnance before making a judgement as to its validity.

In the process, the paper argues that, contrary to common belief, the repugnant conclusion is not inherent in CGU and CLGU. Rather, it results from the interaction of the rule with the set of feasible alternatives. On some choice sets, CGU and CLGU do not exhibit the repugnant conclusion, while on others they do.

In order to assess these outcomes, it is proposed that choice sets should be evaluated in terms of their realism, where realism is defined in terms of conformity with elementary laws of physics and biology, and universal norms of preferences. Five restrictions are proposed to reflect such universal properties: essentiality of material consumption, positive subsistence consumption, upper bounds on resources and non-material goods, and the law of conservation of matter. It is shown that these restrictions entail a particular form of choice set, on which neither CGU nor CLGU exhibit the repugnant conclusion. Thus, if one is willing to ignore unrealistic choice sets, there is no impossibility result.

The paper is structured as follows. Section II provides basic definitions and notation. Section III examines the incidence of the repugnant conclusion under CGU and CLGU for six key choice sets which are representative of relevant outcomes. Section IV defines the restrictions for universal properties – i.e. the realism restrictions – and assesses the six choice sets in terms of these properties. Section V investigates the role of positive subsistence consumption in the main result. Section VI concludes.

II. Notation and definitions

Let $\Re$, $\Re_+$, and $\Re_{++}$ denote the set of real numbers, non-negative real numbers, and positive real numbers, respectively. Define $x_{ij} \in \Re_+$ as the consumption by individual i of good j, with $j = 1,...,m$. Then $x_i = (x_{i1}, x_{i2},..., x_{im})$ represents the consumption vector of individual i. The paper abstracts from the natural discreteness of population, opting instead for real-valued population $N \in \Re_+$ in order to facilitate the use of calculus in the sequel. Thus the reference to individual i is really a shorthand for a position on the interval $[0, N]$. A social alternative x is characterized by population size N_x and the consumption vectors of individuals; i.e. $x = \{N_x, \{x_i\}_{i \in [0, N_x]}\}$. Finally, X represents the set of all social alternatives.

An individual's lifetime utility is given by the real-valued function $u(x_i)$. For simplicity, it is assumed that u is the same for all individuals and that this function is strictly concave and continuously differentiable. A comprehensive notion of utility is intended here, in which well-being is derived from both material and non-material goods, where the latter may include things such as health, longevity, family, friendships, career, education, recreation, or participation in the community. The focus on lifetime measures of utility means that time does not enter explicitly into the analysis; rather, each alternative x represents a complete history of the world, which is to be socially evaluated. The convention is followed here of normalizing lifetime utility such that a value of zero represents neutrality, as discussed in the introduction. Denote the utility realization corresponding with alternative x as $U_x = \{u(x_i)\}_{i \in [0, N_x]}$ and the set of all such realizations on X as $U = \{U_{x \in X}\}$.

An ordering R generates a complete, reflexive and transitive set of binary relations on X, which can be used to socially rank the elements in X. The statement xRx' is translated as "x is at least as good as x'". Completeness entails that for all $x, x' \in X$, xRx' or $x'Rx$ and possibly both. Reflexivity entails xRx for all $x \in X$. Transitivity entails for any triplet $x, x', x'' \in X$:

$$xRx' \quad \text{and} \quad x'Rx'' \quad \Rightarrow \quad xRx''.$$

The ordering has symmetric and asymmetric parts, I and P, corresponding with indifference and strict preference respectively, which are defined as follows:

$$xIx' \quad \Leftrightarrow \quad xRx' \quad \text{and} \quad x'Rx$$

$$xPx' \quad \Leftrightarrow \quad xRx' \quad \text{and not } x'Rx$$

For welfarist social orderings, the relevant information about alternative x is summarized by the utility realization U_x. In other words, for all $x, x' \in X$

$$xRx' \quad \Leftrightarrow \quad U_x R\, U_{x'}.$$

The axioms of continuity, weak and strong Pareto, and anonymity refer to comparisons of alternatives with the same population size. For reference, denote the set of alternatives with the same population size $X_N = \{x \in X \mid N_x = N\}$. The following definitions are adapted from Blackory et al. (2002b).

Continuity: For all $N \in \Re_{++}$ and all $x \in X_N$, the sets $\{x' \in X_N \mid x'Rx\}$ and $\{x' \in X_N \mid xRx'\}$ are closed.

This property states that all at-least-as-good-as sets and all no-better-than sets include their boundaries, which implies that tradeoffs under the ordering R "are gradual, without sudden jumps from better to worse" (Blackorby et al. 2002b, p. 4).

Weak Pareto: For all $N \in \Re_{++}$ and all $x, x' \in X_N$, if $u(x_i) > u(x_i')$ for all $i \in [0, N]$, then xPx'.

Strong Pareto: For all $N \in \Re_{++}$ and all $x, x' \in X_N$, if $u(x_i) \geq u(x_i')$ for all $i \in [0, N]$ and there exists at least one $i \in [0, N]$ such that $u(x_i) > u(x_i')$, then xPx'.

Anonymity: For all $N \in \Re_{++}$, all $x, x' \in X_N$, and all bijective mappings $b : [0, N] \to [0, N]$, if $x_i = x'_{b(i)}$ for all $i \in [0, N]$, then xIx'.

This property indicates that the ordering R is impartial with respect to the identities of individuals. In particular, simply reassigning allocations among a given population yields an alternative with the same rank as the original.

For comparisons of alternatives with different population sizes, the following axiom requires that the ordering be independent of the utilities of individuals who are equally well off in both alternatives.

Existence independence: For all $x, x', x'' \in X$, $(x, x'')R(x', x'') \quad \Leftrightarrow \quad xRx'$.

A welfarist social ordering which satisfies continuity can be represented by a real-valued function $f : U \to \Re$. The at-least-as-good-as set defined on U, for a given $U_{x'} \in U$, is represented $\left\{U_x \in U \middle| f(U_x) \geq f(U_{x'})\right\}$.

Inequality aversion: Given a welfarist social-evaluation function f and a convex domain U, the at-least-as-good-as set is also convex.

This property rules out a social preference for inequality in utilities.

Blackorby et al. (2002b and 2003) demonstrate that, of 13 welfarist population principles examined, only classical generalized utilitarianism (CGU) and critical-level generalized utilitarianism (CLGU) satisfy continuity, strong Pareto, anonymity, existence of critical levels, non-negative critical levels, and existence independence. As inequality aversion also seems compelling, one may wish to add it to this list. It will be shown presently that this property has implications for the functional forms of CGU and CLGU.

CGU compares alternatives on the basis of transformed utility, $V(x_i) = g[u(x_i)]$, where g is an increasing function, assumed continuously twice differentiable, with $g[0] = 0$ to preserve neutrality. Anonymity ensures that g is identical for all individuals. Inequality aversion entails that g is concave; i.e. $g' > 0$ and $g'' \le 0$.[2] The combination of strict concavity of u and concavity of g entails that V is also strictly concave.

The CGU social ordering is defined by

$$xRx' \quad \Leftrightarrow \quad \int_0^{N_x} g[u(x_i)]di \ge \int_0^{N_{x'}} g[u(x_i')]di$$

If g represents a multiplicative constant, then the ordering is classical utilitarian.

CLGU requires the specification of a critical level, $u_c \in \Re$, and then aggregates the difference between transformed utility and the transformed critical level. The CLGU social ordering is defined by

$$xRx' \quad \Leftrightarrow \quad \int_0^{N_x} \left[g[u(x_i)] - g[u_c]\right]di \ge \int_0^{N_{x'}} \left[g[u(x_i')] - g[u_c]\right]di$$

The axiom of non-negative critical levels requires $u_c \ge 0$. Clearly CGU is just a special case of CLGU, with $u_c = 0$.

Comparison of population principles is facilitated by the following representation theorem, due to Blackorby and Donaldson (1984). Define representative utility as a value

[2] See the Appendix for the proof.

$\upsilon(x) \in [\min\{U_x\}, \max\{U_x\}]$ such that x is socially indifferent to an alternative with the same population in which every member receives utility of $\upsilon(x)$.

Representation theorem: For every variable-population ordering R that satisfies continuity, anonymity and weak Pareto, there exists a real-valued function, $W(N,\upsilon)$, which represents the ordering in terms of population size and representative utility only; i.e.

$$xRx' \quad \Leftrightarrow \quad W(N_x, \upsilon(x)) \geq W(N_{x'}, \upsilon(x')) . \blacksquare$$

The characterization of an ordering in terms of $W(N,\upsilon)$ simplifies analysis considerably, as υ provides a mapping of X onto the real number line. Thus let $D \subseteq (\Re_+ \times \Re)$ denote the domain of $W(N,\upsilon)$. Henceforth $W(N,\upsilon)$ is referred to as a variable population welfare function.

Under both CGU and CLGU, representative utility is given by the formula

$$\upsilon(x) = g^{-1}\left[\frac{1}{N_x}\int_0^{N_x} g[u(x_i)]di\right], \tag{1}$$

which reduces to average utility in the non-generalized case, i.e. when g is the identity transformation. It follows that the functions

$$\overline{W}(N,\upsilon) = N \cdot g(\upsilon) \tag{2}$$

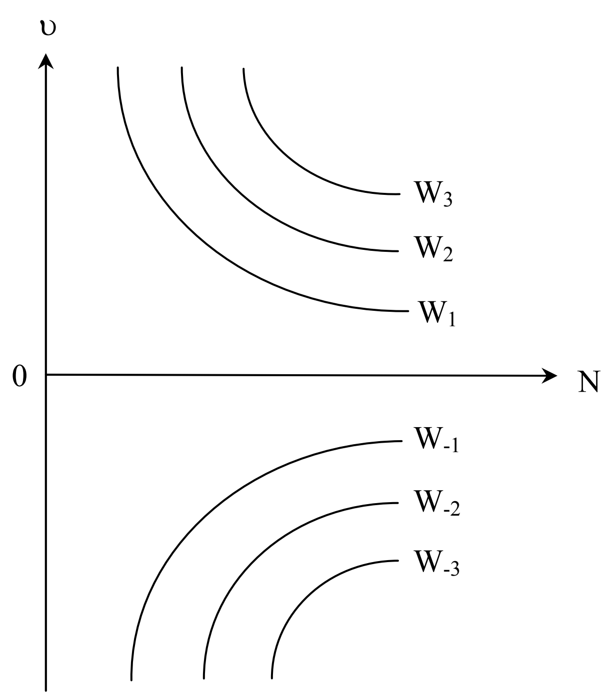
and

$$\hat{W}(N,\upsilon) = N[g(\upsilon) - g(u_c)]$$

provide representations of the CGU and CLGU orderings respectively. These functions have the same domain, as they share a common definition of $\upsilon(x)$, and have the same differentiability properties as $g[\cdot]$, which under present assumptions means they are continuously twice differentiable

Figure 1 presents social indifference curves in (N, υ) space under CGU, according to (2). By construction, $\overline{W}(N, \upsilon)$ is increasing (decreasing) in N for $\upsilon > 0$ ($\upsilon < 0$) and constant at a value of zero along both axes. The indifference map under CLGU, which has not been shown, is simply a translation of the map for CGU, with the origin shifted up by the positive critical value u_c.

Figure 1



The social indifference curves are asymptotic to both axes and have slope

$$\left.\frac{d\upsilon}{dN}\right|_{\overline{W}} = -\frac{g(\upsilon)}{Ng'(\upsilon)}$$

for CGU and

$$\left.\frac{d\upsilon}{dN}\right|_{\hat{W}} = -\frac{g(\upsilon)-g(u_c)}{Ng'(\upsilon)}$$

for CLGU.[3] The slope is always negative in the positive orthant (above $\upsilon = u_c$), which characterizes the trade-off embodied in utilitarianism between N and υ.

The following definition of the repugnant conclusion is based on the critical level rather than neutrality.

Repugnant conclusion (RC): A variable-population welfare function $W(N,\upsilon)$ exhibits the repugnant conclusion if, for every $(N,\upsilon) \in D$ with $\upsilon > u_c$, and for any intermediate value $\varepsilon \in (u_c, \upsilon)$, there exists a population size $M > N$ such that $W(M,\varepsilon) > W(N,\upsilon)$.

This definition seems compelling given the similarity between the critical level and neutrality. To see this point, consider any social alternative in which everyone receives utility equal to the critical level. By definition, this alternative describes a society which is not worth having, since the planner is indifferent between it and the null alternative in which no one lives. Such a society may be referred to as a null-equivalent, although there may be any number of people alive in it. Now, the revised definition of the repugnant conclusion entails that there is always a better alternative arbitrarily close to a null-equivalent, regardless of the starting point. But by virtue of proximity to a null-equivalent, this better alternative must be scarcely worth having. It seems, therefore, that this outcome is no less repugnant than one based on individual neutrality.

[3] For CLGU, the indifference curves are asymptotic to the horizontal line $\upsilon = u_c$.

III. Choice sets and the repugnant conclusion under utilitarianism

Contrary to popular belief, the repugnant conclusion is not an intrinsic property of utilitarianism. Rather, it reflects the interaction of utilitarianism with certain choice sets. This section investigates the incidence of repugnance for six key choice sets. The first set is wholly unbounded, while the remaining five represent variations of a minimally restricted structure. The results obtained for these six sets cover all relevant outcomes.

The unbounded choice set is represented by the welfare domain $D = (\Re_+ \times \Re)$. Both CGU and CLGU exhibit the repugnant conclusion on this set. The proof follows directly from the asymptotic nature of the social indifference curves and is not presented formally.

The five restricted sets which will be considered are based on the specification of an upper bound on representative utility conditional upon the size of the population. This upper bound is denoted $\overline{\upsilon}(N)$. An upper bound may also be specified for population, $\overline{N}$, which may be finite or infinite, depending on the case. With this notation, a minimally restricted choice set is represented as $D = \left\{(N,\upsilon) \,\middle|\, 0 \le N \le \overline{N},\ \upsilon \le \overline{\upsilon}(N)\right\}$. For simplicity, D is represented here as a closed set. All the results would go through for open sets as well, with additional effort.

It will be useful to consider the possibility of a socially optimal choice. Let $(N^*, \upsilon^*) \in D$ represent the optimum according to the social welfare function $W(N, \upsilon)$; i.e. $W(N^*, \upsilon^*) > W(N, \upsilon) \quad \forall (N, \upsilon) \in D,\ N \ne N^*,\ \upsilon \ne \upsilon^*$. Such an optimum may or may not exist, depending upon the choice set.

Now, by virtue of weak Pareto, $\overline{\upsilon}(N)$ is also the locus of conditional optima, i.e. the optimal choices of υ conditional on the corresponding values of N. An unconditional optimum, if it exists, will therefore be located on $\overline{\upsilon}(N)$; i.e. $(N^*, \upsilon^*) = (N^*, \overline{\upsilon}(N^*))$.

As the first of the restricted choice sets, consider a finite population bound $\overline{N}$ and a constant utility bound $\overline{\upsilon} > u_c$. This set is represented in Figure 2 (a.) by the rectangular shaded area. Since the set is closed and bounded in both N and υ, a social optimum exists. The optimum is the corner solution $(\overline{N}, \overline{\upsilon})$, by virtue of the weak Pareto principle and $\overline{W}$ increasing in N. Since an optimum exists, CGU and CLGU do not exhibit the repugnant conclusion on this set. Similarly, the repugnant conclusion does not hold on any other set with positive upper bounds on both N and υ.

A variation on this case would be to keep N bounded but let υ be unbounded; i.e. $\upsilon \in \Re$. CGU and CLGU do not exhibit the repugnant conclusion on this set either, since any point on the population bound, $(\overline{N}, \upsilon)$, can be used to falsify the definition.

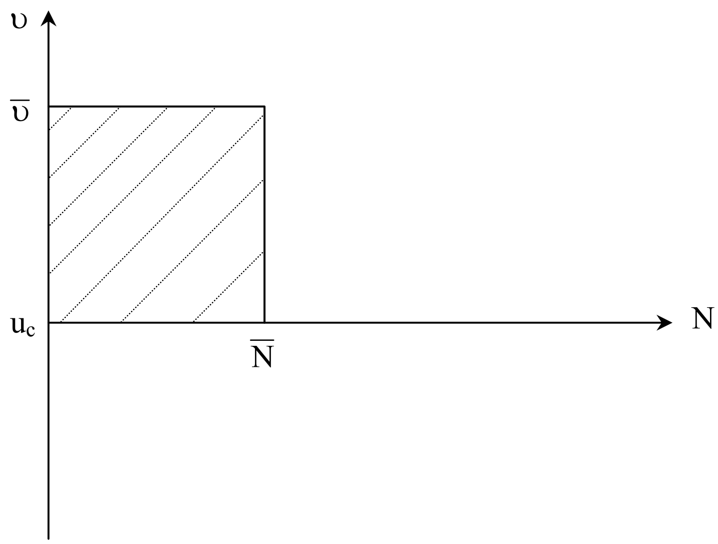
The remaining four choice sets, displayed in parts (b.) – (e.) of Figure 2, are characterized by bounds on utility but not on population. Part (b.) depicts a positive lower bound υ_ℓ on $\overline{\upsilon}(N)$. Part (c.) shows $\overline{\upsilon}(N)$ increasing in N, cutting the horizontal axis ($\upsilon = u_c$) from below. Part (d.) shows $\overline{\upsilon}(N)$ decreasing in N, cutting the axis from above. Finally, Part (e.) depicts $\overline{\upsilon}(N)$ decreasing in N, approaching the horizontal axis asymptotically from above.

CGU and CLGU exhibit the repugnant conclusion on the set depicted in Part (b.). As in the unbounded case, the proof follows from the asymptotic nature of the social indifference curves (not shown in the diagram). Similarly, both CGU and CLGU exhibit the repugnant conclusion on the set depicted in Part (c.).

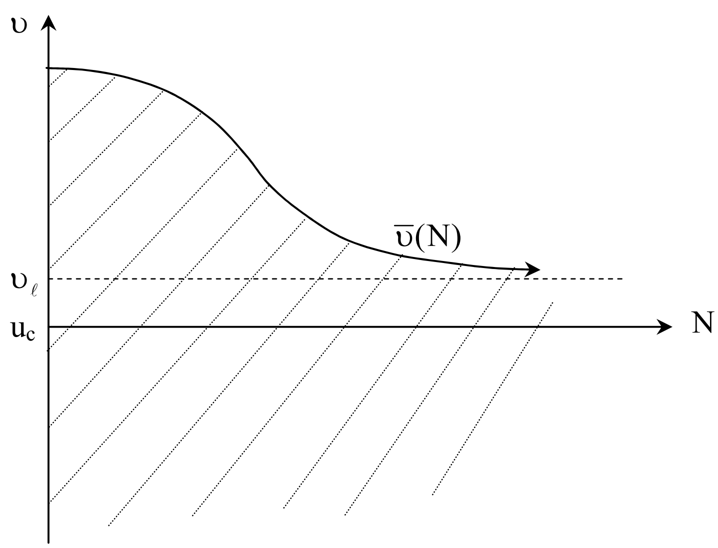
CGU and CLGU do not exhibit the repugnant conclusion on the set depicted in Part (d.). As drawn, $\overline{\upsilon}(N)$ intersects both the vertical and horizontal axes; however, only the horizontal intercept is essential for the result. As drawn, a social optimum exists at a point α on the

Figure 2

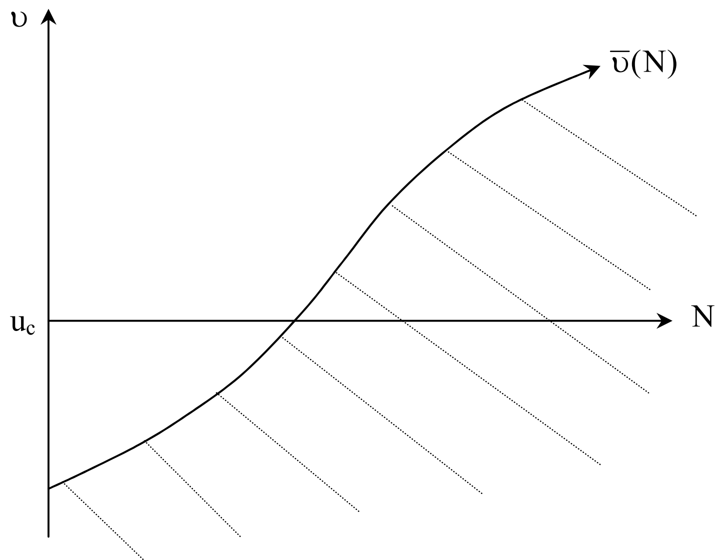
(a.)



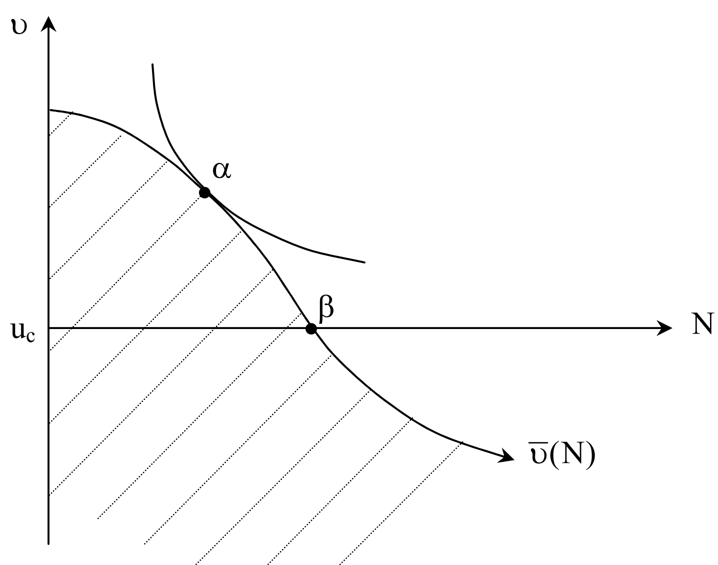
(b.)



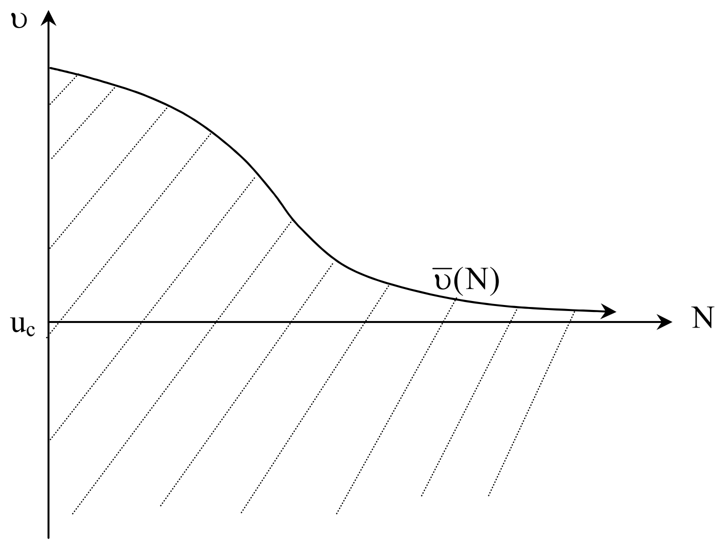
(c.)



(d.)



(e.)



boundary; if $\bar{\upsilon}(N)$ were asymptotic to the vertical axis, the social optimum could instead be a degenerate solution with N arbitrarily close to zero. Note however that the optimum can never occur at the horizontal intercept β, since $\overline{W} = 0$ at this point. In contrast, $\overline{W} > 0$ for alternatives with lower N and positive υ, and therefore the planner will always prefer such points over β. The failure of the repugnant conclusion on this set follows from the existence of an optimum at α (or a degenerate optimum in the neighbourhood of the vertical axis).

The asymptotic case shown in Part (e.) raises the possibility of a degenerate optimum with N^* arbitrarily large and $\bar{\upsilon}(N^*)$ arbitrarily close to the horizontal axis, in which case the repugnant conclusion would hold. A sufficient condition for this result would be that the slope of $\bar{\upsilon}(N)$ is less steep than the indifference map at all points of intersection; i.e. $\left|\frac{d\bar{\upsilon}(N)}{dN}\right| < \left|\frac{d\upsilon}{dN}\Big|_{\overline{W}}\right|$.

Conversely, having $\bar{\upsilon}(N)$ steeper than the indifference map in the neighbourhood of the horizontal axis would mean that the repugnant conclusion does not hold.

IV. Realistic Constraints

The previous section considered the incidence of the repugnance conclusion for utilitarian welfare on six different choice sets. These sets are possible mathematically, but they have not been examined to determine whether they are realistic in the sense of conforming with elementary physical and biological laws or with universal norms concerning individual utility. The present section considers five restrictions to reflect such laws and examines the implications of these restrictions for choice sets and the repugnant conclusion. The five restrictions are: essentiality of material consumption, bounds on non-material goods, positive subsistence consumption, finite material resources, and conservation of matter.

Essentiality of material consumption refers to the fact that biological beings require positive inputs of food and other material goods for survival. For simplicity, this section assumes a single material consumption good, identified as good 1. The results generalize in a straightforward way to cases with more than one such good. The remaining goods, indexed by $j = 2,...,m$, are considered non-material.

Bounds on non-material goods mean that the individual allocation of each non-material good has a quantitative upper bound, denoted $\overline{x}_{ij}$ $j = 2,...,m$. These bounds reflect the reality that all goods are restricted by either biology, physics, or time. For example, biology places an upper limit on how much health an individual can enjoy, since it is not possible to have more than perfect health. Time places restrictions on how much family, career, education or recreation one can enjoy. Since the laws of nature apply equally to everyone, it is reasonable to assume that the bounds are the same for everyone, i.e. $\overline{x}_{ij} = \overline{x}_j \ \forall i$.

An alternative for non-material goods would be to require utility to be bounded above in each good, rather than quantities. This restriction has some appeal, as it ensures diminishing marginal rates of substitution among goods. However, this property seems more an expression of preferences than a universal norm. In contrast, the argument for quantitative bounds seems compelling.

Subsistence consumption $c_s \in \Re_+$ is defined by the relationship

$u(c_s, x_2, ..., x_m) = 0$;

i.e. c_s is the level which yields neutral utility for a given allocation of the non-material goods. This relationship implies a function $c_s = c_s(x_2, ..., x_m)$, which is assumed to be decreasing in the levels of the non-material goods; i.e. $\partial c_s / \partial x_j < 0$ for $j = 2, ..., m$.

The possibility of $c_s = 0$ is ruled out because it violates essentiality of material consumption. Furthermore, it is taken as self-evident that very low levels of consumption will always be judged as worse than a life with no experiences, since low consumption brings hunger, disease, indignity and pain. Therefore c_s must be strictly positive for any non-material allocation $(x_2, ..., x_m)$. It is debatable whether this last point springs from biological necessity or preferences. If it reflects preferences only, it may nonetheless be considered a universal norm. Strict concavity of utility ensures that c_s is unique.

The requirement of finite material resources reflects the relationship between consumption and the environment. Material consumption is derived from material resources taken from the environment, either consumed directly or produced in conjunction with labour and other inputs. At any location, material resources are in finite supply and the possibility of transferring resources from one location to another is limited by the cost and feasibility of transportation.

The limitation on transportation is not simply a reflection of current technology. If it were, one could simply hypothesize the existence of a different social alternative with better technology. Rather, it reflects underlying physical constraints imposed by thermodynamics, among other things. In particular, minimums thresholds of energy and matter are required to transport a given mass a certain distance under ideal conditions. These thresholds establish a maximum distance for which transportation to a certain location would be economic. Beyond this distance, the cost of transportation would always exceed the value of the resources transported, regardless of the technology available. Thus, for example, the transfer of materials from outer space to Earth (or of people to outer space) will remain quite limited. This discussion suggests that a realistic choice set must be defined in terms of a specific location, and, at any given location, the available resources must be considered limited. In what follows, material resources are denoted R and the relevant upper bound is $\overline{R}$.

The law of conservation of matter states that matter can neither be created nor destroyed in an isolated system. This law has implications for the production of the consumption good from a given stock of material resources. In particular, given R, the production of the consumption good must be bounded in all other inputs.

In the present section, the supply of the consumption good is given by the production function $S = S(N, R)$, with N interpreted interchangeably as labour or total population and $R \leq \overline{R}$ representing material resources. In what follows, the notation S_N, S_R and S_{NN}, S_{RR} shall be used for first and second-order partial derivatives respectively, with respect to the noted inputs. In the case of a fixed aggregate supply, the function reduces to the simple form $S = R$

and $S_N = S_{NN} = 0$. Blackorby and Donaldson (1984) refer to this case as the pure population problem. In the case of production, it is assumed $S_N > 0$ and the law of conservation of matter entails eventually diminishing returns to N; i.e. $S_{NN} \leq 0$ beyond some threshold level N', with $\lim_{N\to\infty} S_N = 0$. For simplicity, the paper assumes strictly diminishing returns for all values of N. Note that the vanishing marginal product also entails a vanishing average product; i.e. $\lim_{N\to\infty} S(N,R)/N = 0$. It is also assumed that $S_R > 0$ but no restriction is placed on the second derivative.

Taken together, these five restrictions entail an upper feasibility boundary, $\overline{\upsilon}(N)$, conditional on N. The boundary corresponds with the maximum use of resources, $\overline{R}$, maximal consumption of non-material goods $(\overline{x}_2,...,\overline{x}_m)$ by every individual, and the equal distribution of aggregate consumption, such that everyone receives the average product, S/N. To be precise, each individual receives the allocation $\overline{x}_i = \left(S(N,\overline{R})/N, \overline{x}_2,...,\overline{x}_m\right)$, contingent on N. The maximal consumption of resources and non-material goods on the boundary is obvious. Equal distribution follows from the strict concavity of utility, for, if one individual received less than S/N and another received more, total utility – and therefore representative utility – could be increased by moving both individuals to the average.

Substituting $\overline{x}_i$ into (1) yields

$$\overline{\upsilon}(N) = u\left(\frac{S(N,\overline{R})}{N}, \overline{x}_2,...,\overline{x}_m\right), \tag{3}$$

with slope

$$\frac{d\overline{\upsilon}}{dN} = \frac{u_1(\cdot)}{N}\left[S_N - \frac{S}{N}\right] < 0.$$

The negative sign is obvious in the special case of $S_N = 0$ (pure population problem). In the more general case of $S_N > 0$ and $S_{NN} < 0$, the negative sign follows from the fact that $S_N < S/N$ (strictly diminishing returns to N).

The limiting behaviour of $\overline{\upsilon}(N)$ in this model is determined from (3) as

$$\begin{aligned}\lim_{N\to\infty} \overline{\upsilon}(N) &= u\left(\lim_{N\to\infty} \frac{S(N,\overline{R})}{N}, \overline{x}_2,...,\overline{x}_m\right) \\ &= u\left(0,\overline{x}_2,...,\overline{x}_m\right)\end{aligned} \tag{4}$$

where the last line follows from conservation of matter. Now positive subsistence consumption entails $u(c_s,\overline{x}_2,...,\overline{x}_m) = 0$ corresponding with $c_s > 0$. Therefore $u\left(0,\overline{x}_2,...,\overline{x}_m\right) < 0$, which means $\lim_{N\to\infty} \overline{\upsilon}(N) < 0$ by (4).

The result $\lim_{N\to\infty} \overline{\upsilon}(N) < 0$ means that $\overline{\upsilon}(N)$ cuts the horizontal axis from above, as in Part (d.) of Figure 2. As demonstrated in the previous section, the repugnant conclusion does not hold in this case. In fact, this case is stronger than the one illustrated in Figure 2. There, $\overline{\upsilon}(N)$ intersects the axis defined by $\upsilon = u_c$, whereas here it intersects the neutrality axis $\upsilon = 0$.

All of the other choice sets discussed in the previous section are inconsistent with the restrictions presented in this section. These sets must therefore be regarded as representing imaginary constructs. It follows that CGU and CLGU avoid the repugnant conclusion under realistic restrictions on choice sets.

V. The role of subsistence consumption

All five reality restrictions are essential for the result in the previous section under CGU. In contrast, positive subsistence consumption is not required under CLGU.

To demonstrate, consider the case where $c_s = 0$. By the definition of subsistence, $U(0, \bar{x}_2, \ldots, \bar{x}_m) = 0$ in this case. It follows then from (4) that $\lim_{N \to \infty} \bar{\upsilon}(N) = 0$. Thus $\bar{\upsilon}(N)$ is asymptotic with the neutrality axis. It follows that $\bar{\upsilon}(N)$ cuts the u_c axis from above under CLGU, since $u_c > 0$. This outcome corresponds with the previous section.

In contrast, under CGU, $\bar{\upsilon}(N)$ is asymptotic with the u_c axis, since $u_c = 0$. This outcome is consistent with Part (e.) of Figure 2. It will be shown that the repugnant conclusion holds in this case.

On the boundary $\bar{\upsilon}(N)$, CGU welfare is given by

$$\overline{W}(N, \bar{\upsilon}(N)) = N \cdot g(\bar{\upsilon}(N)) = N \cdot g\left[u\left(\frac{S(N, \overline{R})}{N}\right)\right] = N \cdot V\left(\frac{S(N, \overline{R})}{N}\right),$$

where the notation representing the constant allocation of non-material goods has been suppressed for clarity. The first and second derivatives are:

$$\overline{W}_N = V\left(\frac{S}{N}\right) + V'\left(\frac{S}{N}\right)\left[S_N - \frac{S}{N}\right] \tag{5}$$

and

$$\overline{W}_{NN} = \frac{V''\left(\frac{S}{N}\right)}{N}\left[S_N - \frac{S}{N}\right]^2 + V'\left(\frac{S}{N}\right) S_{NN} < 0.$$

The negative sign of the second derivative follows from strict concavity of transformed utility $(V'' < 0)$ and diminishing or constant marginal product of labour $(S_{NN} \le 0)$, and thus establishes the concavity of $\overline{W}$ in N. An optimum N^* is obtained as the solution of the first-order condition $\overline{W}_N = 0$. The concavity of $\overline{W}$ ensures that N^*, if it exists, is a global optimum.

For ease of exposition, define $\theta \in (0,1)$ as the ratio of marginal to average product; i.e. $\theta \equiv \frac{S_N}{S/N}$. Then the first-order derivative (5) can be rewritten

$$\overline{W}_N = V\left(\frac{S}{N}\right) - (1-\theta)\frac{S}{N}V'\left(\frac{S}{N}\right), \tag{6}$$

which reduces to

$$\overline{W}_N = V\left(\frac{S}{N}\right) - \frac{S}{N}V'\left(\frac{S}{N}\right) \tag{6'}$$

when $S_N = \theta = 0$ (pure population problem).

The domain of V in this case corresponds with the range of S/N. No assumptions have been made about the behaviour of S/N as N becomes arbitrarily small. For simplicity, let δ denote an arbitrarily large upper bound corresponding with N arbitrarily small. The range of S/N is then given by the interval $[0,\delta]$, which is also the domain of V. Neutrality is defined by $c_s = 0$ in this case; thus $V(0) = 0$. V is strictly positive for all other points in its domain.

The following proposition characterizes the sign of $\overline{W}_N$ under current assumptions.

Proposition: In the production model defined by $S = S(N,\overline{R})$, $S_N \geq 0$, $S_{NN} \leq 0$, and $\lim_{N\to\infty} S_N = 0$, if $c_s = 0$ and V is strictly concave, then $\overline{W}_N > 0 \ \ \forall N > 0$.

Proof: Consider first the case where $\theta = 0$. Consider two points $\frac{S}{N} > 0$ and 0 in the domain of V. By strict concavity

$$V(0) - V\left(\frac{S}{N}\right) < V'\left(\frac{S}{N}\right)\left[0 - \frac{S}{N}\right],$$

which reduces to

$$V\left(\frac{S}{N}\right) > \frac{S}{N}V'\left(\frac{S}{N}\right) \tag{7}$$

since $V(0)=0$. This result holds $\forall N > 0$ since $\frac{S(N,\overline{R})}{N}$ is a one-to-one mapping of N into $[0,\delta]$. Combining (7) with $(6')$ establishes that $\overline{W}_N > 0\ \forall N > 0$. The result when $\theta \in (0,1)$ follows immediately from the fact that

$$V\left(\frac{S}{N}\right) - (1-\theta)\frac{S}{N}V'\left(\frac{S}{N}\right) > V\left(\frac{S}{N}\right) - \frac{S}{N}V'\left(\frac{S}{N}\right). \blacksquare$$

The proposition establishes that, regardless of the starting position on $\overline{\upsilon}(N)$, $\overline{W}$ can always be increased by increasing N. It follows that CGU exhibits the repugnant conclusion under these assumptions.

VI. Conclusion

The paper has attempted to shed some light on the apparent incompatibility of avoiding Parfit's repugnant conclusion and satisfying the Pareto plus principle (alternatively the mere addition principle) in variable population social choice. The present state of the literature holds that critical-level generalized utilitarianism with a positive critical level avoids the repugnant conclusion but fails to satisfy Pareto plus, while in contrast classical generalized utilitarianism satisfies Pareto plus but exhibits the repugnant conclusion.

The paper argues for a revised definition of the repugnant conclusion based on the critical level rather than neutrality. The reasonableness of this definition follows from the analogous definitions of neutrality and the critical level, with one defined for the individual and the other for the planner. This change would appear to make the impossibility result worse, since both

CGU and CLGU exhibit the repugnant conclusion under this definition on an unrestricted choice set.

As a possible solution, the paper observes that the repugnant conclusion is not an inherent property of utilitarianism. Rather, it results from the interaction of utilitarianism with certain choice sets. Five restrictions are proposed which reflect universal properties of physics, biology and preferences: essentiality of material consumption, positive subsistence consumption, upper bounds on resources and non-material goods, and the law of conservation of matter. These restrictions entail a particular form of choice set on which neither CGU nor CLGU exhibit the repugnant conclusion.

Thus at most utilitarianism can be said to exhibit the repugnant conclusion on imaginary choice sets – sets which do not conform with universal properties of the physical world or accepted norms regarding preferences. Such an outcome does not provide a compelling basis for rejecting utilitarianism.

Appendix

Inequality aversion and the g transformation.

The CGU welfare function aggregates transformed utilities, $g\left[u(x_i)\right]$. Without loss of generality, consider welfare defined over the utilities of two individuals, 1 and 2; i.e. $W(u_1, u_2) = g[u_1] + g[u_2]$.

A social indifference curve in utility space is characterized by the slope

$$\frac{du_2}{du_1} = -\frac{g'[u_1]}{g'[u_2]} < 0, \tag{A1}$$

and each indifference curve provides the lower bound of an at-least-as-good set $\left\{(u_1, u_2) \,\middle|\, W(u_1, u_2) \geq W(u_1', u_2')\right\}$, where (u_1', u_2') is a reference point on the indifference curve.

Inequality aversion requires that at-least-as-good sets be convex, which will be true if and only if

$$\frac{d^2 u_2}{du_1^2} \geq 0 \tag{A2}$$

on the indifference curve. From (A1),

$$\frac{d^2 u_2}{du_1^2} = -\frac{g''[u_1]}{g'[u_2]} - \frac{\{g'[u_1]\}^2}{\{g'[u_2]\}^3} g''[u_2],$$

which entails

$$g''[u_1] \leq -\left[\frac{g'[u_1]}{g'[u_2]}\right]^2 g''[u_2] \tag{A3}$$

under condition (A2).

Now consider an equal allocation $u_1 = u_2$. For such an allocation, (A3) holds only if $g''[\cdot] \leq 0$. (A3) must hold for any such allocation; therefore $g''[\cdot] \leq 0$ for any point in its domain. It follows that g is concave, since $g' > 0$ and $g'' \leq 0$. This result also holds for CLGU.

Non-generalized utilitarianism satisfies inequality aversion automatically, since $\frac{du_2}{du_1} = -1$ and $\frac{d^2u_2}{du_1^2} = 0$ in this case.

Bibliography

Arrhenius, G. (2000), "An Impossibility Theorem for Welfarist Axiologies," Economics and Philosophy 16, 247-66.

Blackorby, C. and D. Donaldson (1984), "Social Criteria for Evaluating Population Change," Journal of Public Economics 25(1/2), 13-33.

Blackorby, C., W. Bossert and D. Donaldson (2003), "The Axiomatic Approach to Population Ethics," Politics, Philosophy and Economics 2(3), 342-381. Also UBC Discussion Paper 01-03, <http://www.econ.ubc.ca/dpintro.htm>.

Blackorby, C., W. Bossert and D. Donaldson (2002a), "Utilitarianism and the Theory of Justice," ch. 11 in Arrow, K., A. Sen and K. Suzumura, eds., Handbook of Social Choice and Welfare Vol. 1 (Amsterdam: Elsevier). Also UBC Discussion Paper 99-20, <http://www.econ.ubc.ca/dpintro.htm>.

Blackorby, C., W. Bossert and D. Donaldson (2002b), "Population Principles with Number-Dependent Critical Levels," Journal of Public Economic Theory 4(3), 347-368.

Blackorby, C., W. Bossert, and D. Donaldson (1997), "Critical-Level Utilitarianism and the Population-Ethics Dilemma," Economics and Philosophy 13, 197-230.

Blackorby, C., W. Bossert and D. Donaldson (1995), "Intertemporal Population Ethics: Critical-Level Utilitarian Principles," Econometrica 63(6), 1303-20.

Blackorby, C., W. Bossert, D. Donaldson and M. Fleurbaey (1998), "Critical Levels and the (Reverse) Repugnant Conclusion," Journal of Economics (Zeitschrift fur Nationalokonomie) 67(1), 1-15.

Ng, Y.-K. (1989), "What Should We Do About Future Generations? Impossibility of Parfit's Theory X," Economics and Philosophy 5, 1989, 235-53.

Parfit, D. (1976), "On Doing the Best For Our Children," in M. Bayles, ed., Ethics and Population (Cambridge: Schenkman, 1976), 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).

Sikora, R. (1978), "Is It Wrong to Prevent the Existence of Future Generations?" in R. Sikora and B. Barry, eds., Obligations to Future Generations (Philadelphia: Temple), 112-166.

esrc-draft-05.doc