1998

Rotten Altruists, Saccharine Altruists, and Saints: Altruism and Social Optimality

Martin Zelder

POLITICAL
ECONOMY
RESEARCH
GROUP

PAPERS IN POLITICAL ECONOMY

Paper No. 87

"Rotten Altruists, Saccharine Altruists, and Saints: Altruism and Social Optimality"

Martin Zelder

The UNIVERSITY *of* WESTERN ONTARIO

The Political Economy Research Group was established in the faculty of Social Science at the University of Western Ontario in 1988. Its purpose is to foster scholarship, teaching and interdisciplinary research in political economy, with a focus on:

1. the application of economic models and methods to the study of political processes and institutions,
2. the economic impact of political processes and institutions,
3. the influence of economic factors on the formation of public policy and on institutional change,
4. the politics of economic policy making,
5. the political, social, and economic effects of public policy.

Co-directors:
Ronald Wintrobe (Economics)
Robert Young (Political Science)

Board of Directors:
Peter Howitt (Economics)
B.B. Kymlicka (Political Science)
John N. McDougall (Political Science)
Peter Neary (History)
John Whalley (Economics)

Staff:
Jayne Dewar

For further information:
Political Economy Research Group,
Department of Economics,
Social Science Centre,
London, Ontario, Canada N6A 5C2
phone: (519) 661-3877
fax: (519) 661-3292

ISSN: 1181-7151
ISBN: 0-7714-2117-6

ROTTEN ALTRUISTS, SACCHARINE ALTRUISTS, AND SAINTS: ALTRUISM AND SOCIAL OPTIMALITY

Martin Zelder*
Michigan State University and Australian National University
November 1997

*Helpful comments were received from Gerhard Glomm. Luis Fernandez, Paul Thistle, and from participants in a session at the 1997 Canadian Law and Economics Association meetings in Toronto and in seminars at Western Michigan University and Oberlin College. Postal address: Department of Economics, 101 Marshall Hall, Michigan State University, East Lansing, MI 48824-1038; e-mail address: msuecon1@pilot.msu.edu; phone: (517) 432-3116; fax: (517) 432-1068.

1. Introduction

In attempting to make socially-optimal laws, lawmakers are often prominently concerned with *selfish* or *malevolent* decision-makers who impose negative externalities on others. For example, does a (selfish) firm produce and thus pollute too much? Or are (malevolent) assailants punished enough by the criminal law? Lawmakers, however, do not often worry about whether *altruistic* decision-makers should be governed by legal rules. Few would quibble if sexually abusive (obviously malevolent) parents were subject to legal constraint, but what about non-abusive parents who make tangible sacrifices (say, financial ones) for their children, a phenomenon which most economists view as indicative of parental altruism--should their altruistic acts, too, be subject to legal control?[1]

This same question, in a far more abstract sense, has also concerned economists for some time. The economists' abstract version is: do altruists who can transfer resources to their beneficiaries according to various mechanisms make socially-optimal decisions? The answer has been the subject of some disagreement. Depending on the formulation of the model, particularly with regard to the description of altruism, economists have in some cases concluded that altruism **does** generate a social optimum and in other cases that it **does not**. This paper thus attempts to summarize these varied conclusions and to articulate a series of new results on altruism. One major new result, robust to a variety of different specifications of altruism, is labeled the "rotten altruist" theorem: privately optimal resource allocations are not generally Bergson-Samuelson

[1]Anecdotal evidence of legal intervention in the affairs of (presumably) altruistic parents is reported by Hamburg [1997], who describes instances in which mothers were found guilty of neglect and temporarily lost custody of their children as a consequence of leaving the children alone for five minutes or less in seemingly innocuous situations.

social-welfare-maximizing, although they are Pareto-optimal. Attainment of Pareto-optimality will be seen, however, to require that society, paradoxically, places no independent weight on the preferences of the altruist's beneficiary.

2. Altruism and social optimality: a review

Following the seminal contributions to general equilibrium theory by Arrow and Debreu, economists began to address the implications for the fundamental welfare theorems arising from deviations from the axiomatic foundations laid by those and other general equilibrium theorists. One such deviation, the focus of this paper, was interdependent preferences. Unfortunately, much of the previous interdependent-preferences welfare literature contains results reliant on highly specific assumptions.

Winter [1969] presents an example of altruistic preferences which leads to a violation of the first fundamental theorem, i.e., a competitive equilibrium which is not a Pareto-optimum. Winter's characterization of altruism is not very general, however, as it embodies a lexicographic aspect: *any* allocation in which no individual lies below a given poverty line is preferred to *any* allocation in which one or more individuals lie below the line. Yet, it is obvious that for a non-trivial definition of the poverty line, there may be competitive equilibria which, despite the altruism of individuals, leave some individuals in poverty, although equilibria with no one in poverty are attainable.

Appearing roughly simultaneously was Hochman and Rodgers [1969], who justified state-coerced redistribution on the grounds that altruistic competitive equilibrium was not Pareto-optimal. Their formulation of altruism is general, unlike Winter's, although it sidesteps the distinction between paternalistic and non-paternalistic altruism[2] by depicting the altruist's utility as a function of two arguments: altruist's and beneficiary's *incomes*. Given this setup, Pareto-inefficiency due to altruism is non-existent in a (nonstrategic) single-altruist/single-beneficiary model. In order for altruistic preferences to generate inefficiency, Hochman and Rodgers are forced to assume a multiple-altruist (or multiple-beneficiary) model. With, say, multiple altruists and a single beneficiary, the transfer to the beneficiary becomes a public good, and the competitive equilibrium involves suboptimal contribution to the public good due to free-riding. This public-good problem provides the efficiency motivation for state-coerced taxation-and-redistribution.

Archibald and Donaldson [1976] generalize Winter's [1969] model such that preferences are interdependent without necessarily expressing altruism (i.e., envy is allowed) and by eliminating the lexicographic feature characterizing Winter's formulation. Although Archibald and Donaldson find, correctly, that the first welfare theorem fails for their preference specification, their result is fundamentally limited, as Kranich [1988] points out, by the fact that they do not allow for transfers. Restricting the model in this fashion substantially diminishes its relevance for policy.

Kranich [1988] *does* incorporate transfers in his paper, and nevertheless determines that the first fundamental theorem fails as well. The basis of his result is the form of altruism assumed. Specifically, he assumes altruism is paternalistic, and is a function of the beneficiary's *wealth*. This formulation inherently generates welfare conflicts between altruist and beneficiary,

[2]Non-paternalistic altruism exists when the beneficiary's **utility** enters into the altruist's utility function; paternalistic altruism exists when the beneficiary's **consumption (or some other non-utility measure)** enters into the altruist's utility function.

in that the beneficiary will, in some cases, receive at equilibrium a consumption bundle which requires higher wealth (and is thus preferable to the altruist) but which is also less desirable to the beneficiary than a cheaper bundle.

Becker [1974,1981;1991], on the other hand, claimed that a non-paternalistic altruist will "internalize all 'exernalities' affecting" a selfish beneficiary (Becker [1991,p. 284]). Friedman [1988] qualified this claim by illustrating that its truth depended on which definition of efficiency was employed--in his terms, either "Marshall" or "Kaldor" efficiency. A "Marshall" efficient reallocation is one where the compensating variation gained by the gainers exceeds the equivalent variation lost by the losers, while a "Kaldor" efficient reallocation is one where the gainers could, in theory, compensate the losers such that both parties were left better off post-reallocation. Friedman's somewhat instructive point is that altruism allocates resources efficiently in the "Marshall" sense but not in the "Kaldor" sense. Ordinarily, in non-altruism contexts, notes Friedman, the two definitions provide the same answer on whether a particular reallocation is efficient or not. The paradox created by altruism is that "Kaldor" compensation, even in the abstract, makes both parties *worse* off because it effectively undoes the altruistic transfer.

That altruism was not Bergson-Samuelson social-welfare-maximizing in a general single-altruist/single-beneficiary model was first discovered by Zelder [1993] as his "rotten altruist" theorem. Kaplow [1995] restated this result with particular functional forms for utility and social welfare. The task of the present paper is to examine and extend the original Zelder [1993] result and assess its generality.

3. Theorems on altruism

The social optimality properties of standard private choice mechanisms which incorporate altruistic preferences can be examined in a variety of contexts. Consider the conventional single-altruist/single-beneficiary problem extensively analyzed by Becker [1981;1991]. In it, the altruist maximizes her utility function subject to a simple resource constraint. The precise form of the problem depends on assumptions along two dimensions: whether the altruist is *paternalistic* or *non-paternalistic*, and whether altruism is *mutual* or *unilateral*. Utility for a *paternalistic* altruist is not a function of the beneficiary's utility per se, but of some objectifiable (potentially partial) determinant of beneficiary utility, such as beneficiary consumption or income. Alternatively, utility for a *non-paternalistic* altruist contains as a direct argument the beneficiary's utility function. *Unilateral* altruism means the altruist indirectly gains utility from the beneficiary (either paternally or non-paternally) but the beneficiary does not indirectly gain utility from the altruist. *Mutual* altruism means that while the altruist benefits indirectly (either paternally or non-paternally) from the "beneficiary", the "beneficiary" also benefits indirectly from the altruist.

Thus, there are four relevant depictions of altruism: **unilateral paternalistic** altruism, **unilateral non-paternalistic** altruism, **mutual paternalistic** altruism, and **mutual non-paternalistic** altruism. Within each of these four models of altruism, we can derive and compare four allocations: two private allocation mechanisms (the "dog and master" 'equilibrium', and the Nash equilibrium), and two socially-optimal allocations (the Pareto-optimum, and the Bergson-Samuelson social welfare maximum). What we will discover is that "dog and master" altruists, regardless of the form of altruism, choose Pareto-optimal allocations, Nash altruists do not attain Pareto-optimality in some cases, and neither "master" nor Nash altruists maximize Bergson-Samuelson social welfare, except in special cases.

3.1 UNILATERAL PATERNALISTIC ALTRUISM

3.1.1 PRIVATE OPTIMA FOR UNILATERAL PATERNALISTIC ALTRUISM

3.1.1.1 The "dog and master" 'equilibrium'

A model of altruism in which the beneficiary is completely passive, i.e., the altruist chooses on the beneficiary's behalf, has been referred to as the "Dog and Master model" by Archibald and Donaldson [1976]. If the beneficiary is the passive "dog" in this sense, and the altruist is paternalistic, the private maximization problem which generates 'equilibrium'[3] is:

$$(1) \qquad \max_{x_A, x_B} U_A(x_A, x_B) \text{ s.t. } x_A + x_B \leq \hat{x},$$

where x_A and x_B are, respectively, allocations of the single composite consumption good to altruist (A) and beneficiary (B), and price of each good is normalized to one. This implies a first-order condition for a private optimum of:

$$(2) \qquad \partial U_A/\partial x_A = \partial U_A/\partial x_B$$

[3] 'Equilibrium' is a euphemism here, in that while a final allocation to more than one consumer is derived, the allocation does not result from simultaneous optimization by both consumers.

3.1.1.2 The Nash equilibrium

If we instead assume that the beneficiary is not passive, but chooses a best-response decision rule to the altruist's best response decision rule, an alternative private outcome of unilateral paternalistic altruism can be derived, i.e., Nash equilibrium. In such a model, we can imagine each consumer dividing her or his *own* endowment of x between personal consumption and a transfer to the other: thus, the altruist's original endowment ($\hat{x}_A$) is divided between what she retains ($\bar{x}_A$) and what she transfers to the beneficiary ($\bar{x}_B$), meaning that $\bar{x}_A + \bar{x}_B \leq \hat{x}_A$. The analogous constraint for the beneficiary is $\tilde{x}_A + \tilde{x}_B \leq \hat{x}_B$. Corresponding to this representation, the altruist's own (post-transfer) consumption is $x_A = \bar{x}_A + \tilde{x}_A$, and the beneficiary's own (post-transfer) consumption is $x_B = \bar{x}_B + \tilde{x}_B$; the economy-wide budget constraint is $\hat{x}_A + \hat{x}_B = \hat{x}$.

The reaction function for A is then her first-order condition from:

(3) $$\max_{\bar{x}_A, \bar{x}_B} U_A(\bar{x}_A + \tilde{x}_A, \bar{x}_B + \tilde{x}_B) \quad \text{s.t. } \bar{x}_A + \bar{x}_B \leq \hat{x}_A$$

That first-order condition (for an interior solution) is:

(4) $$(\partial U_A/\partial x_A)(dx_A/d\bar{x}_A) = (\partial U_A/\partial x_B)(dx_B/d\bar{x}_B),$$

which can be rewritten as:

(4′) $$\partial U_A/\partial x_A = \partial U_A/\partial x_B$$

Analogously, the reaction function for B can be derived as the first-order condition from:

(5) $$\max_{\tilde{x}_A, \tilde{x}_B} U_B(\bar{x}_B + \tilde{x}_B) \quad \text{s.t. } \tilde{x}_A + \tilde{x}_B \leq \hat{x}_B$$

Because altruism is unilateral, B's reaction function reveals his dominant strategy, i.e.,

(6) $x_B = \hat{x}_B$,

where the following inequality holds:

(6′) $dU_B/dx_B > dU_B/dx_A = 0$

The Nash equilibrium, then is the simultaneous solution to equations (4′) and (6).

3.1.2 SOCIAL OPTIMA FOR UNILATERAL PATERNALISTIC ALTRUISM

3.1.2.1 The Pareto-optimum

The condition for Pareto-optimality with unilateral paternalistic altruism is derived from the following problem:

(7) $\max_{x_A,x_B} U_A(x_A,x_B)$ s.t. $U_B(x_B) \geq \bar{U}_B$ and $x_A + x_B \leq \hat{x}$,

which can be rewritten as:

(7′) $\max_{x_A} U_A(x_A,\hat{x}-x_A)$ s.t. $U_B(\hat{x}-x_A) \geq \bar{U}_B$

The first-order condition for the revised problem is:

(8) $\partial U_A/\partial x_A + (\partial U_A/\partial x_B)(dx_B/dx_A) - \lambda(dU_B/dx_B)(dx_B/dx_A) = 0$,

where λ represents the Lagrange multiplier, and (8) can be rewritten as:

(8′) $\partial U_A/\partial x_A = \partial U_A/\partial x_B - \lambda(dU_B/dx_B)$

3.1.2.2 The Bergson-Samuelson social-welfare-maximum

For the corresponding social optimum of a Bergson-Samuelson social-welfare function, however, we must solve:

(9) $$\max_{x_A, x_B} W(U_A(x_A,x_B),U_B(x_B)) \quad \text{s.t.} \quad x_A + x_B \leq \hat{x}$$

This social-welfare function is chosen because it is the most general (least restrictive) possibility, in that it is weakly separable in individual utilities. This generates the first-order condition for the social optimum:

(10) $$(\partial W/\partial U_A)(\partial U_A/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial x_B) + (\partial W/\partial U_B)(dU_B/dx_B)$$

3.1.3 COMPARISON OF PRIVATE, SOCIAL OPTIMA FOR UNILATERAL PATERNALISTIC ALTRUISM

The purpose of the preceding calculations is to determine whether either private allocation mechanism--"dog and master", or Nash--generates allocations which are socially-optimal in either a Pareto-optimal or Bergson-Samuelson-maximizing sense (and if they do not, in what direction each private optimum deviates from each social optimum).

"Dog and master" compared

Consider the "dog and master" solution. The first-order condition describing it is equation (2), while the Pareto-optimum and Bergson-Samuelson-maximum conditions are, respectively, described by equations (8′) and (10):

(2) $\partial U_A/\partial x_A = \partial U_A/\partial x_B$,

(8′) $\partial U_A/\partial x_A = \partial U_A/\partial x_B - \lambda(dU_B/dx_B)$

(10) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial x_B) + (\partial W/\partial U_B)(dU_B/dx_B)$

From these, we can derive:

Proposition 1: Privately optimal *interior* consumption choices made by *unilateral paternalistic "master"* altruists are Pareto-optimal, but are not Bergson-Samuelson social-welfare-maximizing in that the altruist chooses more consumption for herself and less consumption for her beneficiary than is Bergson-Samuelson social-welfare-maximizing.

Proof:
"Dog and master" optimum is Pareto-optimum
Observe that (2) and (8′) are identical iff $\lambda = 0$; we can prove that λ must equal 0 at the private optimum. The allocation at which $\lambda = 0$ is the allocation at which the slope of the utility possibility frontier (i.e., $1/\lambda$) is $-\infty$. When $1/\lambda = -\infty$, U_A is maximized given the minimum utility constraint $U_B \geq \hat{U}_B$. But this same point is also the unconstrained maximum for A because at that point A provides B with exactly $\hat{U}_B$, the minimum guarantee level, which is confirmed by the fact that the utility constraint is non-binding when $\lambda = 0$. Thus, the "dog and master" private optimum must be Pareto-optimal.
"Dog and master" optimum is not Bergson-Samuelson social-welfare maximum
Because the slope of the utility possibility frontier at the private optimum is $-\infty$, the only way a social indifference curve could be tangent to the UPF at such a point would be if the social indifference curve were vertical (i.e., also had infinite slope). This can be confirmed by finding the condition under which (2) and (10) simultaneously hold, which is:
$(\partial W/\partial U_A)/(\partial W/\partial U_B) = (dU_B/dx_B)/(\partial U_A/\partial x_A - \partial U_A/\partial x_B)$
But if (2) is valid, the numerator of the RHS of this expression is 0, meaning that the social MRS must be infinite for the private optimum to be Bergson-Samuelson social-welfare-maximizing, and the social MRS cannot be infinite given that Bergson-Samuelson SWF satisfies the "Pareto principle" (Boadway and Bruce [1984]), i.e., in this case that $\partial W/\partial U_B > 0$. In fact, in order to satisfy the Pareto principle, we must have $\partial U_A/\partial x_A - \partial U_A/\partial x_B > 0$, which implies $\partial U_A/\partial x_A > \partial U_A/\partial x_B$, at the social optimum. In contrast, at the private optimum, $\partial U_A/\partial x_A = \partial U_A/\partial x_B$. The assumption of diminishing private MRS for A between x_B and x_A thus implies that x_A is too high and x_B too low relative to the Bergson-Samuelson welfare-maximum allocation.

Proposition 1 is valid as long as the altruist regards her own consumption as desirable enough to induce her to choose an interior solution (or a corner solution which is a tangency).

If, however, she attaches lower marginal utility to her own consumption than to the beneficiary's at the allocation where the beneficiary receives the entire consumption endowment, then the "master" altruist's private optimum is the point $(0,\hat{X})$. Operationally, we might refer to an altruist who possesses this property as a "saint". We then have an interesting corollary to accompany Proposition 1:

Corollary 1: "Saints" are not "rotten" unilateral paternal altruists: they make social-welfare-maximizing decisions in both Pareto-optimal and Bergson-Samuelson social-welfare-maximizing senses.

Proof:
If $(0,\hat{X})$ is the (non-tangent) private optimum, then $\partial U_A/\partial x_A < \partial U_A/\partial x_B$ at this optimum. This inequality implies that (8′) and (10) cannot hold as equalities, but rather can only hold as inequalities:

(8″) $\partial U_A/\partial x_A < \partial U_A/\partial x_B - \lambda(dU_B/dx_B)$ (Pareto)

(10′) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) < (\partial W/\partial U_A)(\partial U_A/\partial x_B) + (\partial W/\partial U_B)(dU_B/dx_B)$ (SWF)

This means that when the private optimum is a non-tangent corner solution, the social optima are the same corner solution: $(0,\hat{X})$.

Nash compared

The Nash equilibrium, characterized by the reaction functions (4′) and (6), can be compared with the two social-optimality conditions:

(4′) $\partial U_A/\partial x_A = \partial U_A/\partial x_B$, and
(6) $x_B = \hat{x}_B$ (Nash)

(8′) $\partial U_A/\partial x_A = \partial U_A/\partial x_B - \lambda(dU_B/dx_B)$ (Pareto)

(10) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial x_B) + (\partial W/\partial U_B)(dU_B/dx_B)$ (SWF)

This gives rise to:

Proposition 2: Privately optimal *interior* equilibria for *unilateral paternalistic "Nash"* altruists are Pareto-optimal, but are not Bergson-Samuelson social-welfare-maximizing in that the altruist chooses more consumption for herself and less consumption for her beneficiary than is Bergson-Samuelson social-welfare-maximizing.

Proof: A's reaction function, (4′), and the first-order condition when A is the "master", (2), are identical. In order for (4′) to hold with equality, A's equilibrium transfer to B, $\dot{x}_B$, must be ≥ 0. Therefore, the Nash equilibrium is identical to the master's private optimum, meaning that $\lambda = 0$ at the <u>Nash</u> equilibrium. Thus, this Nash equilibrium is Pareto-optimal. Because of this identity between the Nash equilibrium and the corresponding "dog and master" optimum, it is also the case that the Nash equilibrium is not Bergson-Samuelson social-welfare-maximizing, and in particular, that more x_A and less x_B are consumed at Nash equilibrium than at the Bergson-Samuelson social optimum. The proof of this is identical to the analogous portion of the proof of Proposition 1.

3.2 UNILATERAL *NON*-PATERNALISTIC ALTRUISM

3.2.1 PRIVATE OPTIMA FOR UNILATERAL *NON*-PATERNALISTIC ALTRUISM
3.2.1.1 "Dog and master" 'equilibrium'

For the *non-paternalistic* altruist with a passive selfish beneficiary, the maximization problem is:

$$(11) \quad \max_{x_A, x_B} U_A(x_A, U_B(x_B)) \text{ s.t. } x_A + x_B \leq \hat{x}$$

For this formulation, the first-order condition for an interior private optimum is:

$$(12) \quad \partial U_A/\partial x_A = (\partial U_A/\partial U_B)(dU_B/dx_B)$$

3.2.1.2 Nash equilibrium

For *non-paternalistic* altruism, the reaction function for A is derived as the first-order condition from:

$$(13) \quad \max_{\dot{x}_A, \dot{x}_B} U_A(\dot{x}_A + \tilde{x}_A, U_B(\dot{x}_B + \tilde{x}_B)) \text{ s.t. } \dot{x}_A + \dot{x}_B \leq \hat{x}_A$$

That first-order condition is:

(14) $(\partial U_A/\partial x_A)(dx_A/d\bar{x}_A) = (\partial U_A/\partial U_B)(\partial U_B/\partial x_B)(dx_B/d\bar{x}_B)$,

which can be rewritten as:

(14′) $\partial U_A/\partial x_A = (\partial U_A/\partial U_B)(\partial U_A/\partial x_B)$

Analogously, the reaction function for B can be derived as the first-order condition from:

(15) $\underset{\breve{x}_A,\breve{x}_B}{\text{Max.}}\ U_B(\bar{x}_B+\breve{x}_B)\ \text{ s.t. }\ \breve{x}_A + \breve{x}_B \leq \hat{x}_B$

Because altruism is again unilateral (although now non-paternal), B's reaction function again reflects his dominant strategy:

(16) $\breve{x}_B = \hat{x}_B$, where the following inequality holds:

(16′) $dU_B/dx_B > dU_B/dx_A = 0$

The Nash equilibrium is the simultaneous solution to (14′) and (16).

3.2.2 SOCIAL OPTIMA FOR UNILATERAL *NON*-PATERNALISTIC ALTRUISM

3.2.2.1 The Pareto-optimum

The condition for Pareto-optimality with unilateral *non*-paternalistic altruism is derived from the following problem:

(17) $\underset{x_A,x_B}{\text{Max.}}\ U_A(x_A,U_B(x_B))\ \text{ s.t. }\ U_B(x_B) \geq \hat{U}_B\ \text{ and }\ x_A + x_B \leq \hat{x}$,

which can be rewritten as:

(17′) $\max_{x_A} U_A(x_A, U_B(\hat{x}-x_A))$ s.t. $U_B(\hat{x}-x_A) \geq \bar{U}_B$

The first-order condition for the revised problem is:

(18) $\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(dU_B/dx_B)(dx_B/dx_A) - \lambda(dU_B/dx_B)(dx_B/dx_A) = 0,$

which can be rewritten as:

(18′) $\partial U_A/\partial x_A = (\partial U_A/\partial U_B)(dU_B/dx_B) - \lambda(dU_B/dx_B)$

3.2.2.2 The Bergson-Samuelson social-welfare-maximum

The Bergson-Samuelson social-welfare optimum for the same problem is given by the solution to:

(19) $\max_{x_A, x_B} W(U_A(x_A, U_B(x_B)), U_B(x_B))$ s.t. $x_A + x_B \leq \hat{x},$

where W is the Bergson-Samuelson social-welfare function. Solving this optimization problem yields the following first-order condition for an interior solution:

(20) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial U_B)(dU_B/dx_B) + (\partial W/\partial U_B)(dU_B/dx_B)$

3.2.3 COMPARISON OF PRIVATE, SOCIAL OPTIMA FOR UNILATERAL *NON-*PATERNALISTIC ALTRUISM

"Dog and master" compared

The "dog and master" optimum and the two social optima are, respectively:

(12) $\partial U_A/\partial x_A = (\partial U_A/\partial U_B)(dU_B/dx_B)$ ("dog and master")

(18′) $\partial U_A/\partial x_A = (\partial U_A/\partial U_B)(dU_B/dx_B) - \lambda(dU_B/dx_B)$ (Pareto-optimum)

(20) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial U_B)(dU_B/dx_B) + (\partial W/\partial U_B)(dU_B/dx_B)$ (SWF)

Qualitatively, the comparison of the "dog and master" optimum and the two social optima given unilateral *non-paternal* altruism is identical to the previous comparison for unilateral *paternalistic* altruism, leading to:

Proposition 3: Privately optimal *interior* consumption choices made by *unilateral non-paternalistic "master"* altruists are Pareto-optimal, but are not Bergson-Samuelson social-welfare-maximizing in that the altruist chooses more consumption for herself and less consumption for her beneficiary than is Bergson-Samuelson social-welfare-maximizing.

Proof: Follows same steps as proof of Proposition 1.

Corollary 2: Corner-solution-choosing "saints" are not "rotten" unilateral non-paternal altruists: they make social-welfare-maximizing decisions according to both Pareto-optimality and Bergson-Samuelson social-welfare-maximization.

Proof: Follows proof of Corollary 1.

Nash compared

Similarly, the comparison of the Nash equilibrium and the social optima is the same for unilateral *non-paternal* altruism as it was for unilateral *paternal* altruism:

(14′) $\partial U_A/\partial x_A = (\partial U_A/\partial U_B)(\partial U_B/\partial x_B)$, and
(16) $\bar{x}_B = \hat{x}_B$ (Nash)

(18′) $\partial U_A/\partial x_A = (\partial U_A/\partial U_B)(dU_B/dx_B) - \lambda(dU_B/dx_B)$ (Pareto-optimality)

(20) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial U_B)(dU_B/dx_B) + (\partial W/\partial U_B)(dU_B/dx_B)$
(B-S S-W-max.)

From these equations we can derive:
Proposition 4: Privately optimal *interior* equilibria for *unilateral non-paternalistic "Nash"* altruists are Pareto-optimal, but are not Bergson-Samuelson social-welfare-maximizing in that the altruist chooses more consumption for herself and less consumption for her beneficiary than is Bergson-Samuelson social-welfare-maximizing.

Proof: Follows same line as proof of Proposition 2.

3.3 MUTUAL PATERNALISTIC ALTRUISM

3.3.1 PRIVATE OPTIMA FOR MUTUAL PATERNALISTIC ALTRUISM
3.3.1.1 "Dog and master" 'equilibrium'

Mutual paternalistic altruism means that not only does B's consumption enter into A's utility function, but **A's** consumption also enters into **B's** utility function. For the "dog and master" optimization, however, B's mutual altruism is irrelevant, and the optimization problem and its result are identical to the corresponding equations for the previously-analyzed unilateral paternal altruism "dog and master" problem, equations (1) and (2):

(1) $\max_{x_A, x_B} U_A(x_A, x_B)$ s.t. $x_A + x_B \leq \hat{x}$,

which implies a first-order condition for a private optimum of:

(2) $\partial U_A/\partial x_A = \partial U_A/\partial x_B$

3.3.1.2 Nash equilibrium

A's reaction function is determined by the first-order condition corresponding to equation (3), the problem solved previously for the <u>unilateral</u> paternalistic altruist:

(3) $$\max_{\dot{x}_A,\dot{x}_B} U_A(\dot{x}_A+\bar{x}_A,\dot{x}_B+\bar{x}_B) \text{ s.t. } \dot{x}_A + \dot{x}_B \leq \hat{x}_A,$$

which yields the corresponding first-order condition, equation (4′):

(4′) $\partial U_A/\partial x_A = \partial U_A/\partial x_B$

B's maximization problem differs, however, from its analog in the unilateral paternal altruist context, as B now cares about A's consumption. Thus, B solves:

(21) $$\max_{\bar{x}_A,\bar{x}_B} U_B(\dot{x}_A+\bar{x}_A,\dot{x}_B+\bar{x}_B) \text{ s.t. } \bar{x}_A + \bar{x}_B \leq \hat{x}_B$$

The relevant first-order condition derived from this is:

(22) $\partial U_B/\partial x_A = \partial U_B/\partial x_B$

Solving (4′) and (22) simultaneously generates the condition for Nash equilibrium:

(23) $\partial U_A/\partial x_A + \partial U_B/\partial x_A = \partial U_A/\partial x_B + \partial U_B/\partial x_B$

3.3.2 SOCIAL OPTIMA FOR MUTUAL PATERNALISTIC ALTRUISM

3.3.2.1 The Pareto-optimum

With mutual paternal altruism, the social planner's algorithm for Pareto-optimality is:

(24) $$\max_{x_A,x_B} U_A(x_A,x_B) \text{ s.t. } U_B(x_A,x_B) \geq \hat{U}_B \text{ and } x_A + x_B \leq \hat{x},$$

which can be rewritten as:

(24′) $$\max_{x_A} U_A(x_A,\hat{x}-x_A) \text{ s.t. } U_B(x_A,\hat{x}-x_A) \geq \hat{U}_B$$

The first-order condition which satisfies this is:

(25) $$\partial U_A/\partial x_A + (\partial U_A/\partial x_B)(dx_B/dx_A) - \lambda(\partial U_B/\partial x_A) - \lambda(\partial U_B/\partial x_B)(dx_B/dx_A) = 0,$$

which can be rewritten as:

(25′) $$\partial U_A/\partial x_A - \lambda(\partial U_B/\partial x_A) = \partial U_A/\partial x_B - \lambda(\partial U_B/\partial x_B)$$

3.3.2.2 The Bergson-Samuelson social-welfare-optimum

If instead the social planner takes Bergson-Samuelson social welfare as the social goal, the problem becomes:

(26) $$\max_{x_A,x_B} W[U_A(x_A,x_B),U_B(x_A,x_B)] \text{ s.t. } x_A + x_B \leq \hat{x}$$

The associated first-order condition is:

(27) $$(\partial W/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_B)(\partial U_B/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial x_B)$$

3.3.3 COMPARISON OF PRIVATE, SOCIAL OPTIMA FOR MUTUAL PATERNALISTIC ALTRUISM

"Dog and master" compared

The first-order conditions for the "dog and master" optimum, the Pareto-optimum, and the Bergson-Samuelson-maximum are, respectively are:

(2) $\partial U_A/\partial x_A = \partial U_A/\partial x_B$ ("dog and master")

(25′) $\partial U_A/\partial x_A - \lambda(\partial U_B/\partial x_A) = \partial U_A/\partial x_B - \lambda(\partial U_B/\partial x_B)$ (Pareto-optimum)

(27) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_B)(\partial U_B/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial x_B)$
(B-S social-welfare max.)

From these, we can derive:

Proposition 5: Privately optimal interior consumption choices made by mutual paternalistic "master" altruists *are Pareto-optimal and may be Bergson-Samuelson social-welfare-maximizing*; deviations from Bergson-Samuelson social-welfare-maximization may involve the altruist choosing *more or less* consumption for herself and (respectively) *less or more* consumption for her beneficiary than the Bergson-Samuelson social-welfare-maximum recommends.

Proof:
Pareto-optimality
(2) and (25′) are identical if:
$\partial U_A/\partial x_A - \partial U_A/\partial x_B = 0 = \lambda\{(\partial U_B/\partial x_A) - (\partial U_B/\partial x_B)\}$, which holds if $\lambda = 0$, which it must at the private optimum. Therefore, the private optimum and the Pareto-optimum must coincide.

Bergson-Samuelson social-welfare-maximization
(2) and (27) are identical if:
$\partial U_A/\partial x_A - \partial U_A/\partial x_B = 0 = (\partial W/\partial U_B)/(\partial W/\partial U_A)\{(\partial U_B/\partial x_B) - (\partial U_B/\partial x_A)\}$, which holds if: $(\partial U_B/\partial x_B) - (\partial U_B/\partial x_A) = 0$, given that $\partial W/\partial U_A$ and $\partial W/\partial U_B$ are each > 0 according to the Pareto principle. For $(\partial U_B/\partial x_B) - (\partial U_B/\partial x_A) = 0$ to hold, it means that the allocation chosen by "master" A exactly coincides with the allocation "dog" B would choose were he the "master". This coincidence is possible but not necessary. If A's private optimum does not coincide with B's hypothetical private optimum, A might choose more x_A and less x_B than is Bergson-Samuelson-maximizing, or the reverse. To see this, divide both sides of (27), the Bergson-Samuelson condition, by $(\partial W/\partial U_A)(\partial U_A/\partial x_B)$. Then, we can rewrite (27) as:
(27′) $(\partial U_A/\partial x_A)/(\partial U_A/\partial x_B) = 1 + [(\partial W/\partial U_B)/(\partial W/\partial U_A)]\{(\partial U_B/\partial x_B - \partial U_B/\partial x_A)/(\partial U_A/\partial x_B)\}$
This condition implies that if $\partial U_B/\partial x_B - \partial U_B/\partial x_A > 0$ at the social optimum, then

$(\partial U_A/\partial x_A)/(\partial U_A/\partial x_B)$, A's private MRS at the social optimum, is > 1, meaning that A consumes less x_A and more x_B at the social optimum than at the private optimum (where her private MRS = 1). In other words, she is a "rotten" altruist. Alternatively, if $\partial U_B/\partial x_B - \partial U_B/\partial x_A < 0$ at the social optimum (while still keeping the RHS of (27') positive), then $(\partial U_A/\partial x_A)/(\partial U_A/\partial x_B)$, A's private MRS at the social optimum, is < 1, meaning that A consumes **more** x_A and **less** x_B at the social optimum than at the private optimum (where her private MRS = 1). In this case, in other words, she is a *"saccharine"* altruist, too altruistic from society's perspective.

A's private optimum may also socially-optimal (in both senses) if A is a corner-solution-choosing saint.

Corollary 3: "Saints" might not be "rotten" mutual paternal altruists: they may make social-welfare-maximizing decisions.

Proof (for nontangency corner):
Observe that if the private optimum is $(0,\hat{x})$ and a non-tangency, then $\partial U_A/\partial x_A < \partial U_A/\partial x_B$ at that optimum. $\partial U_A/\partial x_A < \partial U_A/\partial x_B$ is consistent with but does not necessarily imply that the Pareto-optimality condition (25') and the SWF-max. condition (27) will hold as inequalities (i.e., corners):

(25') $\partial U_A/\partial x_A - \lambda(\partial U_B/\partial x_A) < \partial U_A/\partial x_B - \lambda(\partial U_B/\partial x_B)$ (Pareto-optimum)

(27) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_B)(\partial U_B/\partial x_A) < (\partial W/\partial U_A)(\partial U_A/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial x_B)$
(B-S social-welfare-max.)

Note also that $\partial U_A/\partial x_A < \partial U_A/\partial x_B$ is consistent with but does not necessarily imply that the two social-optimality conditions hold as equalities (and thus, possible tangency corner solutions).

Nash compared
Nash equilibrium in the mutual paternal altruism problem requires:

(23) $\partial U_A/\partial x_A + \partial U_B/\partial x_A = \partial U_A/\partial x_B + \partial U_B/\partial x_B$ (Nash)

This condition closely resembles each of the corresponding social-optimality conditions, (25') and (27):

(25') $\partial U_A/\partial x_A - \lambda(\partial U_B/\partial x_A) = \partial U_A/\partial x_B - \lambda(\partial U_B/\partial x_B)$ (Pareto-optimality)

(27) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_B)(\partial U_B/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial x_B)$
(B-S SW-max.)

From these we can derive:

Proposition 6: Privately optimal *interior* consumption by mutual paternalistic Nash altruists *may or may not be* social-welfare-maximizing (i.e., coincide with the social optimum recommended by either Pareto-optimality or Bergson-Samuelson maximization) in that the altruist may consume *more, less, or the same* while the beneficiary consumes, respectively, *less, more, or the same* compared to either the Pareto-optimal or Bergson-Samuelson social-welfare-maximizing solutions.

Proof:

Pareto-optimality

Observe that (23) and (25′) are identical when $\lambda = -1$, i.e., the slope of the utility possibility frontier $(1/\lambda)$ equals -1. For other reasonable (i.e., negative) values of λ, the private and social optimum conditions differ. To see this rewrite (25′) as:

(25″) $[(\partial U_A/\partial x_A)/(\partial U_A/\partial x_B)] = 1 - \lambda[\partial U_B/\partial x_B - \partial U_B/\partial x_A]/(\partial U_A/\partial x_B)$

The Nash condition, (23), can correspondingly be rewritten as:

(23′) $[(\partial U_A/\partial x_A)/(\partial U_A/\partial x_B)] = 1 + [\partial U_B/\partial x_B - \partial U_B/\partial x_A]/(\partial U_A/\partial x_B)$
$= 1$

Then, if $\partial U_B/\partial x_B - \partial U_B/\partial x_A > 0$ at the Pareto-optimum, A's private MRS at the Pareto-optimum is > 1, meaning that (given diminishing MRS) A consumes too much and B consumes too little at the Nash equilibrium; A is a "rotten" altruist. Conversely, if $\partial U_B/\partial x_B - \partial U_B/\partial x_A < 0$ at the Pareto-optimum, A's private MRS at the Pareto-optimum is < 1, meaning (given diminishing MRS) that A consumes too *little* and B consumes too *much* at the Nash equilibrium; A is a *"saccharine"* altruist.

Bergson-Samuelson maximization

Observe that (23) and (27) are identical when $(\partial W/\partial U_B)/(\partial W/\partial U_A) = 1$ at the social optimum. Clearly, for other reasonable optimal values of $(\partial W/\partial U_B)/(\partial W/\partial U_A)$, which is the reciprocal of the social MRS, the private and social optimum conditions differ. Specifically, from the proof of Proposition 5, we saw that at the Bergson-Samuelson social-welfare maximum:

(27′) $(\partial U_A/\partial x_A)/(\partial U_A/\partial x_B) = 1 + [(\partial W/\partial U_B)/(\partial W/\partial U_A)]\{(\partial U_B/\partial x_B - \partial U_B/\partial x_A)/(\partial U_A/\partial x_B)\}$,

whereas, at Nash equilibrium:

(23′) $[(\partial U_A/\partial x_A)/(\partial U_A/\partial x_B)] = 1 + [\partial U_B/\partial x_B - \partial U_B/\partial x_A]/(\partial U_A/\partial x_B)$
$= 1$

Thus, if $\partial U_B/\partial x_B - \partial U_B/\partial x_A > 0$ at the Bergson-Samuelson maximum, A's private MRS at the Bergson-Samuelson maximum is > 1, meaning that (given diminishing MRS) A consumes too much and B consumes too little at the Nash equilibrium; A is a "rotten" altruist. Conversely, if $\partial U_B/\partial x_B - \partial U_B/\partial x_A < 0$ at the Pareto-optimum, A's private MRS at the Pareto-optimum is $<$ 1, meaning that (given diminishing MRS) A consumes too *little* and B consumes too *much* at the Nash equilibrium; A is a *"saccharine"* altruist.

3.4 MUTUAL NON-PATERNALISTIC ALTRUISM

3.4.1 PRIVATE OPTIMA FOR MUTUAL NON-PATERNALISTIC ALTRUISM
3.4.1.1 "Dog and master" 'equilibrium'

With mutual **non**-paternalistic altruism and a passive beneficiary, the altruist solves:

(28) $$\max_{x_A, x_B} U_A(x_A, U_B(U_A, x_B)) \quad \text{s.t.} \quad x_A + x_B \leq \hat{x}$$

The solution to this problem is given by:

(29) $$\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + [(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_A/\partial x_A] = [(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_B/\partial x_B] + (\partial U_A/\partial U_B)(\partial U_B/\partial x_B)$$

In order for the condition stated in (29) to be meaningful, it must be the case that the infinitely-recursive terms $[(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}$ be finite. This is only possible in two circumstances: $(\partial U_A/\partial U_B)(\partial U_B/\partial U_A) = 1$, or $0 < (\partial U_A/\partial U_B)(\partial U_B/\partial U_A) < 1$; in the former case, $[(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty} = 1$, and in the latter case, $[(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty} \rightarrow 0$. Suppose, for simplicity, that $0 < (\partial U_A/\partial U_B)(\partial U_B/\partial U_A) < 1$, and thus that $[(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty} \rightarrow 0$. Call this:

Assumption 1: $0 < (\partial U_A/\partial U_B)(\partial U_B/\partial U_A) < 1$; therefore $[(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty} \rightarrow 0$.

Under Assumption 1, the "dog and master" optimum condition becomes:

(29a) $$\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial U_A/\partial U_B)(\partial U_B/\partial x_B)$$

3.4.1.2 Nash equilibrium

For non-paternalistic altruist A, the Nash maximization problem is:

(30) $$\max_{\dot{x}_A, \dot{x}_B} U_A(\dot{x}_A + \bar{x}_A, U_B(U_A, \dot{x}_B + \bar{x}_B)) \quad \text{s.t.} \quad \dot{x}_A + \dot{x}_B \leq \hat{x}_A,$$

which generates a reaction function identical to the first-order condition for the "master", equation (29):

$$(29)\quad \partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + [(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_A/\partial x_A] = [(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_B/\partial x_B] + (\partial U_A/\partial U_B)(\partial U_B/\partial x_B)$$

Applying Assumption 1, this can again be rewritten as:

$$(29a)\quad \partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial U_A/\partial U_B)(\partial U_B/\partial x_B)$$

Analogously, mutually non-paternalistic altruist B solves the Nash maximization problem:

$$(31)\quad \underset{\tilde{x}_A,\tilde{x}_B}{\text{Max.}}\ U_B(U_A(\bar{x}_A+\tilde{x}_A,U_B),\bar{x}_B+\tilde{x}_B) \text{ s.t. } \tilde{x}_A + \tilde{x}_B \leq \hat{x}_B$$

This implies a reaction function for B of the form:

$$(32)\quad (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + [(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_A/\partial x_A] = [(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_B/\partial x_B] + (\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B$$

Applying Assumption 1, (32) can be rewritten as:

$$(32a)\quad (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B$$

The Nash equilibrium, then, is the simultaneous solution to (29a) and (32a):

$$(33)\quad \partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B$$

3.4.2 SOCIAL OPTIMA FOR MUTUAL NON-PATERNALISTIC ALTRUISM

3.4.2.1 The Pareto-optimum

With mutual non-paternal altruists, the Pareto-optimality-seeking social planner's problem

is:

$$(34)\quad \underset{x_A}{\text{Max.}}\ U_A(x_A,U_B(U_A,\hat{x}-x_A)) \text{ s.t. } U_B(U_A,\hat{x}-x_A) \geq \hat{U}_B$$

This implies a Pareto-optimality condition of:

(35) $\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + [(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_A/\partial x_A] - (\partial U_A/\partial U_B)(\partial U_B/\partial x_B) = \lambda\{(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + (\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + [(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_A/\partial x_A] - (\partial U_B/\partial x_B)\}$

Applying Assumption 1 and rearranging terms, (35) becomes:

(35a) $\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) - \lambda(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial U_A/\partial U_B)(\partial U_B/\partial x_B) - \lambda\{(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B\}$

3.4.2.2 Bergson-Samuelson social-welfare-maximum

The maximum problem in this case is:

(36) $$\underset{x_A,x_B}{\text{Max.}}\ W[U_A(x_A,U_B(U_A,x_B)),U_B(U_A(x_A,U_B),x_B)] \text{ s.t. } x_A + x_B \leq \hat{x}$$

This yields the first-order condition:

(37) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_A)[(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_A/\partial x_A] + (\partial W/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_B)[(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_A/\partial x_A] = (\partial W/\partial U_A)[(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_B/\partial x_B] + (\partial W/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)[(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]^{\infty}[\partial U_B/\partial x_B] + (\partial W/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial x_B)$

Applying Assumption 1, (37) can be rewritten as:

(37a) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial x_B)$

3.4.3 COMPARISON OF PRIVATE, SOCIAL OPTIMA FOR MUTUAL NON-PATERNALISTIC ALTRUISM

"Dog and master" compared

After applying Assumption 1 to the "dog and master" optimum, and the Pareto-optimum and Bergson-Samuelson maximum, we have, respectively:

(29a) $\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial U_A/\partial U_B)(\partial U_B/\partial x_B)$ ("dog and master")

(35a) $\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) - \lambda(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial U_A/\partial U_B)(\partial U_B/\partial x_B) - \lambda\{(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B\}$ (Pareto-optimum)

(37a) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial x_B)$ (B-S social-welfare maximum)

From these we can derive:

Proposition 7: Privately optimal <u>interior</u> consumption choices made by <u>mutual *non-*</u> <u>paternalistic "master"</u> altruists *are Pareto-optimal but may or may not be Bergson-Samuelson social-welfare-maximizing* in that the altruist may choose more, less, or the same consumption for herself and (respectively) less, more, or the same consumption for her beneficiary compared to what the Bergson-Samuelson standard recommends.

Proof:
Pareto-optimality
The "master" allocates consumption so that:
$\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) -(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) = 0$
The Pareto-optimal social planner allocates consumption so that:
$\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) -(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) = -\lambda\{(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A)\}$

The two conditions are identical if $\lambda = 0$, which must be the case at the private optimum.

Bergson-Samuelson social-welfare-maximization
The private optimum and the Bergson-Samuelson maximum are the same if:
$[\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) -(\partial U_A/\partial U_B)(\partial U_B/\partial x_B)] = (\partial W/\partial U_B)/(\partial W/\partial U_A)[(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A)] = 0$,
which is true if
$(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = 0$
For this condition to hold, "master" A's private optimum must be identical to the allocation "dog" B would choose if <u>she</u> were the master. To analyze the cases where the two conditions do <u>not</u> coincide, we can rewrite (29a) and (37a) as:

(29a') $(\partial U_A/\partial x_A)/(\partial U_A/\partial x_B) = [(\partial U_A/\partial U_B)(\partial U_B/\partial x_B)]/[(\partial U_A/\partial x_B)[1 + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]$ and

(37a') $(\partial U_A/\partial x_A)/(\partial U_A/\partial x_B) = \{(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)/(\partial W/\partial U_A)[(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A)]\}/\{(\partial U_A/\partial x_B)[1 + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)\}$

Then, if $(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) > 0$ at the social optimum, (37a'), which is A's private MRS at the social optimum, is greater than (29a'), which is A's private MRS at the *private* optimum. In this case, A is a "rotten" altruist, choosing too much x_A and too little x_B. Conversely, if $(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) < 0$ at the social optimum, (37a'), which is A's private MRS at the social optimum, is less than (29a'), which is A's private MRS at the *private* optimum. In this case, A is a *"saccharine"* altruist, choosing too *little* x_A and too *much* x_B.

A saint corollary also applies:

Corollary 4: Mutual non-paternalistic "master" saints might make Pareto-optimal choices, and might make Bergson-Samuelson social-welfare-maximizing choices.
Proof:
If the private optimum is a non-tangency corner solution, then
$\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) < (\partial U_A/\partial U_B)(\partial U_B/\partial x_B)$
This is consistent with but does not necessarily imply that the two social-optimality conditions to hold with inequality, i.e.:

$\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) - \lambda(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) < (\partial U_A/\partial U_B)(\partial U_B/\partial x_B) - \lambda\{(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial U_B/\partial x_B)\}$ (Pareto-optimality)

$(\partial W/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) < (\partial W/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial x_B)$ (B-S social-welfare maximum)

Note also that $\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) < (\partial U_A/\partial U_B)(\partial U_B/\partial x_B)$ is consistent with both social-optimality conditions holding as <u>equalities</u> (which could thus be <u>tangency</u> corner solutions).

Nash compared

Finally, the Nash equilibrium condition and the two social-optimality conditions are, respectively:

(33) $\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B$ (Nash)

(35a) $\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) - \lambda(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial U_A/\partial U_B)(\partial U_B/\partial x_B) - \lambda\{(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B\}$ (Pareto-optimum)

(37a) $(\partial W/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) + (\partial W/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) = (\partial W/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)(\partial U_B/\partial x_B)$ (B-S social-welfare maximum)

From these we can derive:

Proposition 8: Privately optimal *interior* consumption by <u>mutual non-paternalistic Nash</u> altruists *may or may not be* social-welfare-maximizing (i.e., coincide with the social optimum recommended by either Pareto-optimality or Bergson-Samuelson maximization) in that the altruist may consume *more, less, or the same* while the beneficiary consumes, respectively, *less, more, or the same* compared to either the Pareto-optimal or Bergson-Samuelson social-welfare-maximizing solutions.

Proof:

Pareto-optimality

At the Nash equilibrium,

$$\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) - (\partial U_A/\partial U_B)(\partial U_B/\partial x_B) = (\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A)$$

At the Pareto-optimum,

$$\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) - (\partial U_A/\partial U_B)(\partial U_B/\partial x_B) = -\lambda\{(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A)\}$$

The two conditions are identical if $\lambda = -1$, which is plausible but not necessary. For the cases where $-\lambda \neq 1$, then we can compare the two conditions by rewriting (33) and (35a) as:

$$(33')\ (\partial U_A/\partial x_A)/(\partial U_A/\partial x_B) = \{(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A)\}/\{(\partial U_A/\partial x_B)[1 + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)\}$$
$$= [(\partial U_A/\partial U_B)(\partial U_B/\partial x_B)]/[(\partial U_A/\partial x_B)[1 + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)]$$

$$(35a')\ (\partial U_A/\partial x_A)/(\partial U_A/\partial x_B) = \{(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) - \lambda[(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A)]\}/\{(\partial U_A/\partial x_B)[1 + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)\}$$

Consequently, if $(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) > 0$, then (35a'), A's private MRS at the *social* optimum, is greater than (33'), A's private MRS at the *private* optimum, meaning that in this circumstance, A is a "rotten" altruist in Pareto terms. If instead $(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) < 0$, A is a "saccharine" altruist in Pareto terms.

Bergson-Samuelson social-welfare-maximization

Similarly, at the Bergson-Samuelson SW maximum,
$$(\partial W/\partial U_A)/(\partial W/\partial U_B)[\partial U_A/\partial x_A + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)(\partial U_A/\partial x_A) - (\partial U_A/\partial U_B)(\partial U_B/\partial x_B)] = (\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A),$$
which is identical to the Nash condition if $(\partial W/\partial U_A)/(\partial W/\partial U_B) = 1$. We saw in the proof of Proposition 7 that solving the Bergson-Samuelson condition for A's private MRS yields:

$$(37a') \quad (\partial U_A/\partial x_A)/(\partial U_A/\partial x_B) = \{(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + (\partial W/\partial U_B)/(\partial W/\partial U_A)[(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A)]\}/\{(\partial U_A/\partial x_B)[1 + (\partial U_A/\partial U_B)(\partial U_B/\partial U_A)\}$$

If $(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) > 0$, then (37a') is larger than A's private MRS from the Nash problem, namely (33'). In this circumstance, A is a "rotten" altruist in Bergson-Samuelson terms. If the reverse condition holds, i.e., $(\partial U_B/\partial U_A)(\partial U_A/\partial U_B)(\partial U_B/\partial x_B) + \partial U_B/\partial x_B - (\partial U_B/\partial U_A)(\partial U_A/\partial x_A) < 0$, then A is a "saccharine" altruist in Bergson-Samuelson terms.

4. Discussion and applications

4.1 The "rotten altruist" theorem and the Coase theorem

A skeptical reader might question whether the social-welfare suboptimum found in some cases still obtains given the abiding possibility that altruist and beneficiary may be able to effect Coasean transactions. As discussed in section 2, however, Friedman [1988] has identified the inherent defect in any attempt to remedy suboptimal altruistic transfers via private transactions between altruist and beneficiary: private transactions (in which the beneficiary 'bribes' the altruist) effectively reduce the size of the altruistic transfer, exacerbating the inefficiency. Because the trading of property rights does not eliminate (or even reduce) the social-welfare loss, a corrective subsidy (to the altruist) is, therefore, the only effective state intervention in this instance (Kaplow [1995]).

4.2 Why and where the "rotten altruist" theorem matters for law-and-economics

The main theoretical result of the paper--altruists' private decisions are rarely social-welfare-maximizing if the social welfare calculation places any non-zero weight on B's utility--implies a possibly substantial and previously unrecognized role for state intervention. The state's prevailing attitude towards intrafamilial matters has traditionally been one of presumptive nonintervention. Only in demonstrable instances of abuse or criminality has the state generally been willing to interfere in intact-family decision-making.[4] Implicitly underlying such a policy is that parental/spousal **malevolence** (or at least selfishness) is a necessary condition for legal control. What the "rotten altruist" theorem suggests is that the scope for efficient state regulation within the family (among other places) may be far wider than this previously-understood guideline of malevolence/selfishness control.

One example is in relation to adoption law. While concerns about abusive adoptive (or biological) parents have dominated the discussion of parental motivation and fitness, the finding of this paper suggests that attention should perhaps also be directed to 'ordinary' altruistic (adoptive and biological) parents. In particular, efficient adoption regulations would subsidize (or otherwise encourage) altruistic adoptive and biological parents to transact so as to reduce the welfare loss inherently associated with their private transactions. Designing such regulations is complicated, however, although a first step was taken by Zelder [1993]. Presumably, many other important applications of the "rotten altruist" theorem exist, within family law and elsewhere in law-and-economics.

5. Conclusion

Economists have recently come to suspect that there are problems with altruism, even in theory. This paper confirms this suspicion by describing a "rotten altruist" theorem, which implies that, in some circumstances, no altruist who cares about herself even a little more than she cares about her beneficiary makes social-welfare-maximizing choices. Only saints are guaranteed to maximize social welfare, and only then in certain circumstances. In a world with few saints, a role for law is even more crucial than was previously realized.

---

[4]But see Hamburg [1997] for reports of a growing propensity to intervene even in cases of minimal neglect.

REFERENCES

Archibald, G.C., and Donaldson, David. "Non-Paternalism and the Basic Theorems of Welfare Economics." **Canadian Journal of Economics** 9:492-507 [1976]

Becker, Gary S. "A Theory of Social Interactions." **Journal of Political Economy** 82:1063-1093 [1974]

Boadway, Robin, and Bruce, Neil. **Welfare Economics**. Oxford: Basil Blackwell [1984]

________. **A Treatise on the Family**. Cambridge: Harvard University Press [1981; 1991, enlarged edition]

Friedman, David D. "Does Altruism Produce Efficient Outcomes? Marshall Versus Kaldor." **Journal of Legal Studies** 17:1-13 [1988]

Hamburg, Gail Vida. "When Parents Fail." **Chicago Tribune** November 9 [1997]

Hochman, Harold M., and Rodgers, James D. "Pareto Optimal Redistribution." **American Economic Review** 59:542-557 [1969]

Kaplow, Louis. "A Note on Subsidizing Gifts." **Journal of Public Economics** 58:469-477 [1995]

Kranich, Laurence J. "Altruism and Efficiency: A Welfare Analysis of the Walrasian Mechanism with Transfers." **Journal of Public Economics** 36:369-386 [1988]

Winter, Sidney G., Jr. "A Simple Remark on the Second Optimality Theorem of Welfare Economics." **Journal of Economic Theory** 1:99-103 [1969]

Zelder, Martin. "Could the American Adoption Market Possibly Be Efficient? Empirical and Theoretical Analysis." Presented at the 1993 American Law and Economics Association annual meetings, Chicago [1993]