J Risk Uncertain (2008) 37:187–197
DOI 10.1007/s11166-008-9054-9

Perfect and total altruism across the generations

Stephen Fels · Richard Zeckhauser

Published online: 15 October 2008

Abstract The traditional formulation of the altruism model has altruistic terms that relate solely to other parties' felicities from consumption. But if those others are altruistic as well, their altruism benefits are being neglected. *Total altruism* takes the total utilities of others, rather than their mere felicities, as the basis for altruistic valuations. We assess total altruism in an intergenerational world. *Perfect altruism*, a concept due to Ramsey (*Econ. J.* 38(152):543–559, 1928), requires that a generation value itself relative to its successor as it would any two consecutive generations. Total altruism and perfect altruism prove to be incompatible concepts. Total altruism is only meaningful when there is some generational selfishness. The analysis considers both forward-looking and forward and backward-looking altruism.

Keywords Altruism · Perfect altruism · Total altruism · Backward-looking · Discounting

JEL Classifications D64 · D90 · D62

Fels was physicist at Princeton before his untimely death in 1989. This paper draws heavily on Zeckhauser and Fels (1968). Fels was responsible for the mathematics in this paper, which utilizes Green's function, a key concept in a variety of fields in physics. Nils Wernerfelt provided able research assistance.

S. Fels
Princeton University, Princeton, NJ, USA

R. Zeckhauser (✉)
Kennedy School of Government, Harvard University,
79 John F. Kennedy Street, Cambridge, MA 02138, USA
e-mail: richard_zeckhauser@Harvard.Edu

We consider a dynamic situation in which generations follow one upon another. Each generation has a utility function of common form. One argument of a generation's utility function, U_t, is a felicity term relating only to its own consumption, f_t. The other arguments are altruistic terms relating to other generations' utilities.[1]

Our discussion is general and need not be related to the successive generations model. It is equally applicable, for example, to a model in which a single individual has a utility function for each unit period in his life that includes "altruistic" terms relating to other of his period utilities, as might someone contemplating saving or exercising for long-term well being. Similarly, it would capture a situation where individuals have utility functions that include altruistic terms relating to their neighbors' utilities. Perhaps surprisingly, Luttmer (2005) discovers that individuals tend to value their neighbors' income as a negative factor, suggesting envy rather than altruism.[2] We know of no analysis, however, where the welfares of other generations are not valued positively.

1 Total altruism

The traditional formulation of the altruism model has altruistic terms in a generation's utility function that relate solely to other generations' felicity terms. We would argue that a truly altruistic generation would let other generations evaluate their own welfares. If they, like the present generation have utility functions that include altruistic terms, these terms should not be neglected in the altruistic considerations of the present generation. In our formulation, therefore, we use total utilities rather than felicities as the basis for altruistic evaluations. It is evident that if altruism is a major factor driving current policies, say in curbing greenhouse gases to save the future from severe climate change, the impact of that altruism would be much greater if it incorporated other generations' total utilities.

Although we think that altruistic preferences should and do relate directly to total utilities, we would really admit that felicities from own consumption are the original source of all utilities. In this essay we are interested in what might be called the basic structure of a formulation using total utilities. We will think of this structure as a derived formulation that employs felicities as the primordial argument of altruistic terms.

If we discover that some apparently reasonable formulations using total utilities imply unreasonable basic structures, we will think further before em-

[1] In spirit, our terminology is borrowed from Gorman (1957). He uses felicity to apply to an individual's current satisfaction. He defines utility as an aggregate function of felicities. It is the highest achievable average level of felicities in a given situation.

[2] Fortunately, the model could be adapted to the negative altruism case. It is well known that individuals are most envious of the success of those close to them. See Tesser et al. (1988).

ploying these formulations. In some sense, the former is only truly reasonable if the latter makes sense as well.[3]

It is equally interesting to know whether traditional formulations relating to felicities are consistent with, and are the basic structure for, acceptable models with total utilities. If the answer is affirmative, we can think of these formulations as shorthand forms of writing what we regard to be the true utility functions. This is much in the way we might think of structural equations in relation to the behavioral equations of an economic model.

2 Forward-looking model

Let us start with a model that is forward-looking; the felicities and utilities of interest are those of the present and future generations. It is traditional in economic models to weight future returns of a proximity basis. That is, the closer to the current period returns are received the greater the weight they receive in the objective function. Formulating our model in the classical manner, we have

$$U_t = f_t + bU_t + d\sum_{j=1}^{\infty} a^j U_{t+j}, \tag{1}$$

where b and d are weighting constants and a gives the relative weights assigned to the utilities of generations n+1 and n periods into the future. If $a < 0$, we have oscillating weights, half of which are negative, placed on future generations' utilities. If $a > 1$, the summation in Eq. 1 does not converge. In accord with our economic intuition we will consider the case $0 < a < 1$.[4]

We solve Eq. 1 to get the recursive relation

$$U_t = \left[\frac{a(1-b+d)}{1-b}\right] U_{t+1} + \left[\frac{d}{1-b}\right]\left[\frac{f_t}{a} - f_{t+1}\right], \tag{2}$$

which can be shown to have the solution

$$U_t = \left[\frac{1}{1-b} f_t\right] + \frac{d}{(1-b)(1-b+d)} \sum_{j=1}^{\infty} \left[\frac{a(1-b+d)}{1-b}\right]^j f_{t+j} + P\left[\frac{1-b}{a(1-b+d)}\right]^j, \tag{3}$$

[3]This stipulation may seem strange, particularly in view of the claims we made for our approach. However, we would argue that a decision maker may know what form he wishes his preference function to take without being able to specify its exact configuration. Much of the interesting work in decision theory relates to the point of making expressed personal preferences accord with self-imposed standards of rationality.

[4]Koopmans (1960) provides an axiomatic justification for such a value.

where P is any arbitrary constant. The last addend in Eq. 3 can be neglected as it contains none of the f_t's, the parameters of interest; alternatively, we can set $P = 0$. Equation 3 is of the form

$$U_t = C\left[f_t + \delta \sum_{j=1}^{\infty} \lambda^j f_{t+j}\right], \tag{4}$$

where $\delta = \dfrac{d}{1-b+d}$ and $\lambda = \dfrac{a(1-b+d)}{1-b}$.

3 Perfect altruism

The case where $\delta = 1$ was the one considered by Ramsey (1928) in his pioneering paper, "A Mathematical Theory of Savings." Phelps and Pollak (1968) have labeled this case perfect altruism.

> ...By this we mean that each generation's preference for their own consumption relative to the next generation's consumption is not different from their preference for any future generation's consumption relative to the succeeding generation. This is a stationary postulate; the present generation's preference ordering of consumption streams is invariant to changes in their timing. (p. 185)

4 Result with forward-looking model

With $\delta = 1$, $1 - b$ must equal zero and Eq. 3 is meaningless. This rules out perfect altruism in a forward-looking model that relates altruistic preferences to total utilities rather than felicities. Here, total altruism and perfect altruism are incompatible with each other.

In the equivalent formulations Eqs. 1 and 4, the discount factor for total utilities, a, is below the factor, λ, that in effect is applied to felicities. If we feel that the latter should be less than unity, that the weights assigned to generations' felicities should decline regularly as we move forward away from the current generation, then we must have

$$0 < a < \frac{1-b}{1-b+d}. \tag{5}$$

5 Forward and backward-looking model

If altruism crosses generations, if we have a positive concern for individuals in other generations, it would seem that this concern should go backward as well as forward. Framed on an ethical basis, it is hard to argue that altruism should stretch forward but not backward. Apart from the fact that past generations may not get to see what we do, it would seem that their utilities should get weight. Thus, for example, if past generations willed us a clean planet or a proud religious tradition, or if our ancestors passed down the family silver, we surely expect that they would like to see us keep those assets proudly, and pass them forward. Following what would be their wishes probably gives us a warm glow; violating them would probably bring pain. That is what tradition is all about, and is part of the reason why older generations fiercely try to instill the warmth-and-guilt surrounding tradition in their successors.

We will now derive an equivalent result for a model that allows utilities of past generations as well as those of the present and future generations to enter as altruistic arguments of the present generation's utility function.

One might argue that for decision-making purposes it is hardly possible to consider the forward-looking altruistic feelings of the past. These feelings can only be based on expectations and any divergence between what is expected and what happens is not of consequence. However, if this argument ever gets established no generation will be able to expect that its altruistic feelings will be taken into account in future periods. Its utility will likely decline, and this in turn will diminish the altruistic arguments of the utility functions of backward-looking present and future generations. In such a situation it is quite likely that there would be a benefit to all if there were some means by which each generation could assure generations previous to itself that their altruistic feelings would be taken into account. As is often the case, establishing a mechanism for commitment benefits those who bind themselves as well as those to whom they are bound.

We proceed on the assumption that the altruistic feelings of past generations matter, and that where they are based on the expected utilities of future generations these will not be different from the actual utilities that are achieved. We will use proximity to the present generation as the basis for discounting the utilities of past generations. This means, quite reasonably we believe, that in the utility function of the present generation, the utility of the most recent generation will be weighted more highly than one further past. This also means that we will be discounting backwards rather than forwards in time, an anomaly which is not unreasonable given the structure of backward-looking altruistic preferences.

We define the utility function of a generation in a manner parallel to that used in the forward-looking model. We have,

$$U_t = f_t + bU_t + d\sum_{j=1}^{\infty} a^j (U_{t+j} + U_{t-j})\,. \tag{6}$$

For the same economics arguments as before, we consider only the case $0 < a < 1$.

The solution of the infinite set of simultaneous linear equations given in Eq. 6 is not uncomplicated. The primarily mathematical solution is given in the Appendix. There it is shown that the final structure is:

$$U_t = C_1 \left[C_2 f_t + C_3 \sum_{j=1}^{\infty} (\lambda_-)^j (f_{t+j} + f_{t-j}) \right], \tag{7}$$

which reproduces Eq. 35 from the Appendix. Equation 7 is parallel to Eq. 4 of the forward-looking model. For perfect altruism to exist, we must have $C_2 = C_3$, which requires that $\lambda_- = K/2$. Unfortunately, this renders meaningless the overall multiplicative constant, C_1.

6 Result with forward and backward-looking model

Let us return to economics. We have seen that a model that allows for backward as well as forward-looking altruistic utility terms can be consistent with a reasonable model employing only felicities, the original source of all utilities. It will be consistent if the inequality in Eq. 16 is satisfied. From Eq. 16, as from Eq. 5 in the forward-looking model, it can be seen that the range of the values of a for which the derived structural model is meaningful contracts toward 0 as b or d increases.

7 Conclusion

We have found that models with total altruism, models in which generations' total utilities rather than their felicities from own consumption enter as the altruistic arguments of any specific generation's utility function, can be reduced to a structural form in which only felicities enter as arguments of the specific generation's utility function.

Total altruism has been shown to be inconsistent with perfect altruism. If we believe that total altruism is justified or required on some ethical or logical basis, we will have to reconsider our feelings about Ramsey-type perfect altruism. The two concepts are incompatible.

Appendix

Algebraic derivation of result with forward and backward-looking model

After a bit of algebra, one may show that U_t satisfies the recursion relation

$$U_{t+1} + U_{t-1} - KU_t = F_t\,, \tag{8}$$

where

$$K = \left[\frac{2ad + (1-b)(a+\frac{1}{a})}{1-b+d} \right] \tag{8.a}$$

and

$$F = \left[\frac{f_{t+1} + f_{t-1} - (a+\frac{1}{a}) f_t}{1-b+d} \right]. \tag{8.b}$$

We wish to find the general solution to Eq. 8. We observe that if we have any two linearly independent solutions $U_t^{(1)}$ and $U_t^{(2)}$ which satisfy the homogeneous equation

$$U_{t+1}^{(i)} + U_{t-1}^{(i)} - KU_t^{(i)} = 0, \ \ (i = 1, 2), \tag{9}$$

and if we are able to find any particular solution $U_t^{(p)}$, of the full Eq. 8, the most general solution of Eq. 8 is

$$U_t = AU_t^{(1)} + BU_t^{(2)} + U_t^{(p)}, \tag{10}$$

where A and B are arbitrary constants. (The above remarks may seem more reasonable if the reader thinks of Eq. 8 as a finite difference approximation to a second order ordinary differential equation.)

Let us first dispose of the homogeneous equation. We make the ansatz,

$$U_t = \lambda^t U_0, \tag{11}$$

and find that Eq. 9 becomes

$$\left(\lambda + \frac{1}{\lambda} - K \right) \lambda^t U_0 = 0,$$

or

$$\lambda_{\pm} = \frac{K \pm \sqrt{K^2 - 4}}{2}. \tag{12}$$

It is evident that if $|K| > 2$, we have two real roots, while if $|K| < 2$, the roots are complex and mutually conjugate. If the two roots of a quadratic $ax^2 + bx + c = 0$ satisfy $r_1 r_2 = c$, we may conclude that $\lambda_+ \lambda_- = 1$ in the case at hand. Now if $|K| > 2$, one of the λ's is greater that 1, and the other is less. In this case, therefore, we have two solutions of Eq. 9;

$$U_t^{(1)} = \left[\frac{K + \sqrt{K^2 - 4}}{2} \right]^t U_0^{(1)} = (\lambda_+)^t U_0^{(1)} \tag{13.a}$$

and

$$U_t^{(2)} = \left[\frac{K - \sqrt{K^2 - 4}}{2} \right]^t U_0^{(2)} = (\lambda_-)^t U_0^{(2)}. \tag{13.b}$$

The first grows exponentially with increasing t, while the second dies exponentially.

If on the other hand, $|K| < 2$, $\lambda_+^* = \lambda_-$, and $|\lambda_+| = |\lambda_-| = 1$. Hence we may write

$$\lambda_+ = e^{i\phi} \text{ and } \lambda_- = e^{-i\phi}, \tag{14.a}$$

and

$$U_t^{(1)} = e^{it\phi} U_0^{(1)}, \tag{14.b}$$

and

$$U_t^{(2)} = e^{-it\phi} U_0^{(2)}. \tag{14.c}$$

These solutions oscillate with constant amplitude. If we wish the solutions to be real, we can choose

$$U_t^{(3)} = U_0^{(3)} \sin(t\phi) \quad \text{and} \quad U_t^{(4)} = U_0^{(4)} \cos(t\phi). \tag{15}$$

We note that if

$$0 < a < \frac{1-b}{1-b+2d}, \tag{16}$$

then $2 < |K|$, and the solutions are real exponentials, while if

$$\frac{1-b}{1-b+2d} < a < 1, \tag{17}$$

then $|K| < 2$, and we have oscillatory behavior.

We now turn to the task of finding a particular solution of

$$U_{t+1} + U_{t-1} - KU_t = F_t.$$

This is most easily done by means of what are called Green's functions.

Let us suppose that we are able to find a function $G_{t,j}$ which satisfies

$$G_{t+1,j} + G_{t-1,j} - KG_{t,j} = \delta_{t,j}, \tag{18}$$

where

$$\delta_{t,j} = 1 \text{ if } t = j \tag{18.a}$$

$$= 0 \text{ if } t \neq j. \tag{18.b}$$

We now claim that if we set

$$U_t = \sum_{j=-\infty}^{\infty} G_{t,j} F_j, \tag{19}$$

then U_t satisfies the inhomogeneous equation. This is easily proved:

$$\begin{aligned} U_{t+1} + U_{t-1} - KU_t &= \sum_{j=-\infty}^{\infty} \{G_{t+1,j} + G_{t-1,j} - KG_{t,j}\} F_j \\ &= \sum_{j=-\infty}^{\infty} \delta_{t,j} F_j \\ &= F_t. \end{aligned} \tag{20}$$

We therefore need only find an appropriate $G_{t,j}$. To do so we note that if $t \neq j$,

$$G_{t+1,j} + G_{t-1,j} - KG_{t,j} = 0\,. \tag{21}$$

Thus we may write (c.f. Eqs. 13.a, 13.b)

$$G_{t,j} = A(\lambda_+)^t + B(\lambda_-)^t \quad \text{for} \quad t < j, \tag{22.a}$$

and

$$G_{t,j} = C(\lambda_+)^t + D(\lambda_-)^t \quad \text{for} \quad t > j. \tag{22.b}$$

We cannot have $A = C$ and $B = D$, for this would imply that

$$G_{t+1,j} + G_{t-1,j} - KG_{t,j} = 0 \text{ for all } t,\ j. \tag{23}$$

To determine the constants, let us first use the fact that

$$G_{j+2,j} + G_{j,j} - KG_{j+1,j} = 0, \tag{24.a}$$

from which it follows that

$$G_{j,j} = A(\lambda_+)^j + B(\lambda_-)^j. \tag{24.b}$$

In the same way, it is easy to show that

$$G_{j,j} = C(\lambda_+)^j + D(\lambda_-)^j; \tag{25.a}$$

thus,

$$(A - C)(\lambda_+)^j = (B - D)(\lambda_-)^j. \tag{25.b}$$

To proceed, we recall that $G_{j+1,j} + G_{j-1,j} - KG_{j,j} = 1$ by definition; therefore,

$$C(\lambda_+)^{j+1} + D(\lambda_-)^{j+1} + A(\lambda_+)^{j-1} + B(\lambda_-)^{j-1} - K\left[A(\lambda_+)^j + B(\lambda_-)^j\right] = 1. \tag{26}$$

We now have two equations in 4 unknowns; the two remaining degrees of freedom can (and later will) be shown to arise from the fact that there is no unique particular solution to the inhomogeneous equation.

We may thus choose $B = C = 0$. We then have

$$A(\lambda_+)^j = D(\lambda_-)^j. \tag{27.a}$$

and

$$D(\lambda_-)^{j+1} = A(\lambda_+)^{j-1} - KA(\lambda_+)^j = 1. \tag{27.b}$$

These show that

$$A = \frac{\lambda_+^{-j}}{(\frac{1}{\lambda_+} + \lambda_- - K)}. \tag{28}$$

Hence,

$$G_{t,j} = \frac{1}{(\frac{1}{\lambda_+} + \lambda_- - K)}(\lambda_+)^{t-j} \text{ for } \; t \leq j, \tag{29.a}$$

and

$$G_{t,j} = \frac{1}{(\frac{1}{\lambda_+} + \lambda_- - K)}(\lambda_-)^{t-j} \text{ for } \; t \geq j. \tag{29.b}$$

We then may easily see that

$$\begin{aligned} U_t &= \frac{1}{(\frac{1}{\lambda_+} + \lambda_- - K)}\left[\sum_{j=-\infty}^{t-1}(\lambda_-)^{t-j}F_j + \sum_{j=t+1}^{\infty}(\lambda_+)^{t-j}F_j + F_t\right] \\ &= \frac{1}{(\frac{1}{\lambda_+} + \lambda_- - K)}\left[F_t + \sum_{j=1}^{\infty}(\lambda_-)^j F_{t-j} + \sum_{j=1}^{\infty}(\lambda_+)^{-j}F_{t+j}\right]. \end{aligned} \tag{30}$$

If we remember that $\lambda_+\lambda_- = 1$, and add the general solution of the homogeneous equation, we find that the most general solution of the full equation is

$$U_t = P(\lambda_+)^t + Q(\lambda-)^t + \frac{1}{(2\lambda_- - K)}\left[F_t + \sum_{j=1}^{\infty}\lambda_-^j\left[F_{t+j} + F_{t-j}\right]\right], \tag{31}$$

where P and Q are arbitrary constants. When expressed in terms of f_t rather than F_t, this becomes

$$\begin{aligned} U_t = P(\lambda_+)^t + Q(\lambda-)^t + \frac{1}{(1+d-b)(2\lambda_- - K)} \\ \times\left[\left(2\lambda_- - \left(a + \frac{1}{a}\right)\right)f_t + \left(K - \left(a + \frac{1}{a}\right)\right)\sum_{j=1}^{\infty}(\lambda_-)^j[f_{t+j} + f_{t-j}]\right] \end{aligned} \tag{32}$$

It remains only to redeem the promise made earlier, concerning the indeterminacy in $G_{t,j}$. Consider, therefore, two different $G_{t,j}$'s, both of which satisfy

$$G_{t+1,j} + G_{t-1,j} - KG_{t,j} = \delta_{t,j}; \tag{33.a}$$

call them $G^1_{t,j}$ and $G^2_{t,j}$. Let

$$U^1_t = \sum_{j=-\infty}^{\infty} G^1_{t,j} F_j \quad \text{and} \quad U^2_t = \sum_{j=-\infty}^{\infty} G^2_{t,j} F_j; \tag{33.b}$$

finally set

$$U^3_t = U^1_t - U^2_t. \tag{33.c}$$

It is trivial to show that

$$U^3_{t+1} + U^3_{t-1} - KU^3_t = 0, \tag{34}$$

i.e., U^3 is a solution of the homogeneous equation. It follows that any Green's function $G_{t,j}$ may be used in the particular solution.

Equation 32 can be rewritten in the form

$$U_t = C_1 \left[C_2 f_t + C_3 \sum_{j=1}^{\infty} (\lambda_-)^j (f_{t+j} + f_{t-j}) \right], \tag{35}$$

where the arbitrary constants, P and Q, are set equal to zero.

References

Gorman, W. M. (1957). Convex indifference curves and diminishing marginal utility. *Journal of Political Economy, 65*(1), 40–50.
Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica, 28*(2), 287–309.
Luttmer, E. F. P. (2005). Neighbors as negatives: relative earnings and well-being. *The Quarterly Journal of Economics, 120*(3), 963–1002.
Phelps, E. S., & Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies, 35*(2), 185–199.
Ramsey, F. P. (1928). A mathematical theory of saving. *Economic Journal, 38*(152), 543–559.
Tesser, A., Millar, M., & Moore, J. (1988). Some affective consequences of social comparison and reflection processes: The pain and pleasure of being close. *Journal of Personality and Social Psychology, 54*(1), 49–61.
Zeckhauser, R., & Fels, S. (1968). Discounting for proximity with perfect and total altruism. Discussion Paper Number 50, Harvard Institute of Economic Research.