Systems of Benevolent Utility Interdependence

Ted Bergstrom

May 19, 1987
Number 87-4

DEPARTMENT OF ECONOMICS
University of Michigan
Ann Arbor, Michigan 48109
Abstract. This paper concerns the logic of benevolently related utility functions. A paradox of 'superbenevolence' is examined and defanged. For a finite set of benevolent consumers, the theory of dominant diagonal matrices is shown to be a powerful tool for the study of 'normal' benevolence. To treat intergenerational benevolence properly, the standard theory of dominant diagonal matrices has to be extended to denumerably infinite dominant diagonal matrices. We show that there is a nice extension that exactly serves our purposes. These results make it possible to generalize and clarify the results of Robert Barro and Miles Kimball on familial altruism. Questions of cardinality and uniqueness of representation are also resolved. Finally, some additional light is thrown on the problem of forward and backward intertemporal consistency which was discussed by Robert Pollak, John Burbridge and others.
If Romeo’s happiness depends on his own consumption and on Juliet’s happiness and if Juliet’s happiness depends on her own consumption and Romeo’s happiness, then we are going to have to study their preferences simultaneously. Utility representation will take the form of a pair of interconnected simultaneous equations.

This paper concerns the logic of benevolently interrelated systems of utility functions. Benevolent interaction is an important ingredient in any satisfactory theory of the family. It is also important for the analysis of forward and backward intergenerational transfers in the form of inheritance, national debt, and social security.

The discussion begins with the simplest possible example of benevolent interaction—where the interaction is between just two people. Studying this elemental case helps to focus the issues that arise in more general situations. This special case is followed by a general analysis of benevolent interactions and then by another class of special cases. In particular we study some of the special utility structures that arise in stationary models of intergenerational preferences. This last discussion was stimulated by and benefitted from a recent paper by Miles Kimball, called “Making Sense of Two-sided Altruism” (1986). The treatment of two-person benevolence draws on some earlier work by Bergstrom (1971).

1. The Case of Two Star-Crossed Lovers

Romeo and Juliet both like to consume goods, and each wants the other to be happy. Their utility functions are:

\[ U_R(c_R, U_J) = u_R(c_R) + aU_J(c_J) \]

\[ U_J(c_J, U_R) = u_J(c_J) + bU_R(c_R) \]  \hspace{1cm} (RJ.1)

where \( c_R \) and \( c_J \) are the consumption vectors of Romeo and Juliet and where \( a \) and \( b \) are positive constants.

In matrix form, these equations are written:

\[
\begin{pmatrix}
1 & -a \\
-b & 1
\end{pmatrix}
\begin{pmatrix}
U_R \\
U_J
\end{pmatrix}
=
\begin{pmatrix}
u_R \\
u_J
\end{pmatrix}
\]  \hspace{1cm} (RJ.2)

If \( ab \neq 1 \), the matrix in RJ.2 can be inverted so that:

\[
\begin{pmatrix}
U_R \\
U_J
\end{pmatrix}
=
\frac{1}{1-ab}
\begin{pmatrix}
1 & a \\
b & 1
\end{pmatrix}
\begin{pmatrix}
u_R \\
u_J
\end{pmatrix}
\]  \hspace{1cm} (RJ.3)

Therefore we can write:
\[
U_R = \left(\frac{1}{1 - ab}\right) u_R(c_R) + \left(\frac{a}{1 - ab}\right) u_J(c_J)
\]
\[
U_J = \left(\frac{1}{1 - ab}\right) u_J(c_J) + \left(\frac{b}{1 - ab}\right) u_R(c_R).
\] (RJ.4)

The system of interconnected utility functions in RJ.1 can in this way be disentangled to yield the two independent utility functions in RJ.4. The interdependent utilities, RJ.1, are said to induce independent preferences over allocations of consumption for each individual where these preferences are represented by the utility functions, RJ.4.

From the equations RJ.4, it is apparent that Romeo's private preferences over consumption bundles for himself, holding Juliet's consumption bundle constant, are represented by the utility function \(\tilde{u}_R(c_R) = \left(\frac{1}{1 - ab}\right) u_R(c_R)\). On the other hand, we see from equation RJ.1 that Romeo's preferences over alternative consumption bundles, holding Juliet's utility constant are represented by \(u_R(c_R)\). Similarly, Juliet's private preferences on consumption bundles for herself, holding Romeo's consumption constant, are represented by \(\tilde{u}_J(c_J) = \left(\frac{1}{1 - ab}\right) u_J(c_J)\) and her preferences on consumption bundles for herself, holding Romeo's utility constant are represented by \(u_J(c_J)\). Now if \(ab < 1\), there is nothing mysterious here. For each person, \(i\), the function \(\tilde{u}_i(c_i)\) is a positive multiple of \(u_i(c_i)\) and hence both functions represent the same preferences over private allocations.

The case where \(ab > 1\) is a little harder to understand. In this case, the larger \(u_i(c_i)\) is, the smaller will be \(\tilde{u}_i(c_i)\). This means that a change in his own consumption that Romeo would favor if Juliet’s consumption bundle were held constant is a change that he would oppose if Juliet's consumption were to be altered so as to hold her utility constant when Romeo's consumption changes. For example, suppose that there is just one good, spaghetti, and suppose that \(ab > 1\). If Romeo prefers more spaghetti to less, given Juliet’s spaghetti consumption, then he must prefer less spaghetti for himself rather than more, if Juliet's utility is held constant.

With a bit of reflection, we see why this is the case. Suppose, for instance, that \(a > 1\) and \(b > 1\). Then each lover is more concerned about the other's consumption utility, \(\tilde{u}_i(c_i)\) than about his own. Now consider two allocations \((c_R, c_J)\) and \((c'_R, c'_J)\) such that Juliet is indifferent between the two allocations and such that \(c'_R > c_R\). Since Juliet is more interested in Romeo's consumption utility than in her own, but is indifferent between the two allocations, it must be that \(c'_J < c_J\) and indeed \(\tilde{u}_J(c_J) - \tilde{u}_J(c'_J) > \tilde{u}_R(c'_R) - u_R(c_R)\). But Romeo is more interested in Juliet's consumption utility than in his own. Therefore the reduction in Juliet's spaghetti consumption that was required to hold her utility constant will reduce his utility by more than he gained from the increase in his own spaghetti consumption.

Let us call the case where \(a \geq 0, b \geq 0\) and \(ab < 1\) the case of normal benevolence. This is the case where conventional demand theory works most easily. All that needs to be added to the theory is to allow the possibility that people may give each other gifts of desirable goods. But it is not unreasonable to believe that there are couples whose concern for each other is so strong that \(ab > 1\). This is the case of superbenevolence. If there is
superbenevolence, then demand theory encounters extra difficulties since these affectionate pairs will want to give each other nice gifts but will be unwilling to accept them. In order for the analysis to proceed, some assumptions have to be made about how conflicts of this type are resolved. For example, “property rights” could be interpreted to mean that you are not allowed to give somebody a present unless the recipient accepts it willingly.

A First Remark on Scaling of Utility Functions

Another question that arises in the case of simultaneous systems of utility functions is the matter of cardinality of utility. For example, What kinds of transformations of utility functions represent the same preferences? What properties of this system of equations are invariant under transformations that represent the same preferences?

In this introductory section, we simply look at an example. Suppose that in RJ.1, we have \( a = 2 \) and \( b = 1/4 \). Then

\[
U_R = u_R(c_R) + 2U_J
\]

\[
U_J = u_J(c_J) + \frac{1}{4} U_R. \tag{RJ.6}
\]

Looking at RJ.6, one might be tempted to conclude that Romeo loves Juliet more than Juliet loves Romeo. But, in fact, no such inference can be made. To see this, let us consider another system of utility functions that represent the same preferences. Define \( U_J^* = 3U_J \) and \( U_R^* = U_R \). Then, by simple arithmetic, we find that the system of equations RJ.6 is equivalent to the following:

\[
U_R^* = u_R(c_R) + \frac{2}{3} U_J^*
\]

\[
\frac{1}{3} U_J^* = \frac{1}{3} u_J(c_J) + \frac{1}{4} U_R^*. \tag{RJ.6'}
\]

If we define \( u_J^*(c_J) = 3u_J(c_J) \) and multiply both sides of the second equation in RJ.6' by 3, we can write the following system of equations which is equivalent to RJ.6.

\[
U_R^* = u_R(c_R) + \frac{2}{3} U_J^*
\]

\[
U_J^* = u_J^*(c_J) + \frac{3}{4} U_R^*. \tag{RJ.6'‘}
\]

Therefore, although the systems of equations RJ.6 and RJ.6' are alternative representations of the same system of preferences, in the former system, \( a > b \) and in the latter system, \( b > a \). The relative weights that Romeo and Juliet put on each other’s “utility” are not invariant under transformations that represent the same preferences. On the other hand, notice that in both systems, it happens that \( ab = 1/2 \). In the next section we will develop general results on invariance under alternative representations of preferences. An implication of that theory is that in the case of Romeo and Juliet, the product \( ab \) is preserved under all equivalent representations.
2. General Systems of Benevolent Interaction

The issues that arise in the case of Romeo and Juliet appear in more intricate form when the affections of several people entangle. Furthermore, there are some additional mathematical problems to be sorted out because in the study of intergenerational preferences it is natural to deal with interaction among a countably infinite set of consumers.

We consider a set $I$ of persons with interrelated preferences. Let $c_i$ denote the consumption vector of person $i$, let $U_i$ denote $i$'s utility level, and let $U_{-i}$ denote the vector of utilities of every person other than $i$. Each consumer $i$ has preferences defined over vectors $(c_i, U_{-i})$ that can be represented in the weakly separable form:

$$U_i = F_i(u(c_i), U_{-i}). \quad (GS.1)$$

where $F_i$ is a non-decreasing function of all of its arguments.

One question that we will want to ask is: "When can a system of interconnected utility functions like $GS.1$ be disentangled to yield an equivalent system of independent utility functions of the form:

$$U_i = G_i(u_i(c_i), u_{-i}(c_{-i})) \quad (GS.2)$$

where $u_{-i}(c_{-i})$ is the vector whose components are the consumption utilities $u_j(c_j)$ for $j \neq i$ and where $G_i$ is a non-decreasing function of all of its arguments?" Recall that in the case of Romeo and Juliet, the answer turned out to be "This will happen as long as there is not 'too much benevolence'." We will find a similar answer for the general case.

More formally, we will say that the set of interdependent utility functions $GS.1$ induces the set of independent utility functions $GS.2$ if the functions $U_i = G_i(u_i(c_i), u_{-i}(c_{-i}))$ are determined implicitly as a solution to the system of equations $GS.1$. Let us define the system $GS.1$ to be normally benevolent if the functions $F_i$ are nondecreasing in all of their arguments and if $GS.1$ induces a system of independent utility functions, $U_i = G_i(u_i(c_i), u_{-i}(c_{-i}))$, such that the functions $G_i$ are nondecreasing in all of their arguments. In the special case of Romeo and Juliet, the system of utility functions is normally benevolent if and only if $a > 0, b > 0$, and $ab < 1$. In this new vocabulary, the question of the previous paragraph becomes "When is a system of utility functions of the form $GS.1$ a normally benevolent system?"

In the current discussion we simplify our problem by looking at the additively separable case where

$$U_i = F_i(u_i, U_{-i}) = u_i + \sum_{j \neq i} a_{ij}U_j. \quad (GS.3)$$

But the results that we find extend directly to the general case where $F_i(u_i, U_{-i})$ is continuously differentiable but not necessarily linear.\(^1\) The system of equations $GS.3$ can be

\(^1\) A local version of our results follows from straightforward application of the implicit function theorem. There is also a global theorem which can be estabished using the Gale-Nikaido theorem on the existence of global inverses for nonlinear functions on convex sets (Nikaido, 1968). The conditions we impose on the linear function imply that the Jacobian of $U-F(u,U)$ as a function of $U$ is a $P$-matrix and hence the Gale-Nikaido can be applied.
written as a matrix equation:

\[ U = u + AU \]  \hspace{1cm} (GS.4)

where \( A \geq 0 \). When \((I - A)^{-1}\) exists, the system GS.4 induces the independent system of utility functions represented by the matrix equation

\[ U = (I - A)^{-1}u. \]  \hspace{1cm} (GS.5)

Therefore the system GS.4 will be normally benevolent if and only if the \((I - A)^{-1}\) is nonnegative in every element. Evidently, what we want to know is "What conditions on a non-negative matrix \( A \), will guarantee that \((I - A)^{-1}\) exists and \((I - A)^{-1} \geq 0\)? The answer comes straight from the theory of dominant diagonal matrices with positive diagonals and non-positive off-diagonals. This theory as applied to finite dimensional matrices (McKenzie, 1960) is a familiar economists' tool.

In order to treat intergenerational preferences in an elegant and convenient way, it is useful to be able to deal with a doubly infinite sequence of generations—where we consider ancestors running back to an infinite past and descendants running forward to an infinite future. Even though we may believe that the "true" model is one with a long but finite past and a long but finite future, approximation by an infinite horizon model is an attractive method. This procedure allows us to state crisp, clean results which exclude complicated but inessential details. Of course the approximation of long finite horizons models by infinite horizon models is only appropriate if the distant future and the distant past turn out to have "little influence" on current preferences. Conveniently for the purposes of this paper, the theory of dominant diagonal matrices extends very nicely to the case of denumerable matrices.

**On dominant diagonal matrices—finite and denumerable**

An \( n \times n \) matrix \( M \) with positive diagonals and negative off-diagonal elements is said to be **dominant diagonal** if there exists an \( n \times n \) diagonal matrix \( D \) with \( D_{ii} > 0 \) for all \( i \), such that the matrix \( MD \) has all of its row sums positive. The result that we will want to use is the following.

**Lemma DDM.** Let \( I - A \) be an \( n \times n \) dominant diagonal matrix such that \( A \geq 0 \). Then the matrix \( B = \sum_{t=0}^{\infty} A^t \) is well-defined and finite, \( B \geq 0 \), and \( B = (I - A)^{-1} \).

The notation of linear algebra and its operations extend in the obvious way to denumerably infinite matrices and vectors. While many of the fundamental results of finite dimensional linear algebra carry over to this environment, there are some dramatic surprises. Among the surprises are the fact that matrix multiplication is not in general associative and the fact that a matrix may have more than one inverse (For a good exposition of this theory, see Kemeny et. al. (1966). Fortunately for us, in the case of denumerable dominant diagonal matrices, both of these monstrous facts can be defanged.

For our analysis, a vector \( x \) will have denumerable infinity of components, \( x_i \), where the index \( i \) runs from \(-\infty \) to \( \infty \). A matrix \( M \) has components \( M_{ij} \) where \( i \) and \( j \) run from \(-\infty \) to \( \infty \). Matrix multiplication is defined in exact analogy to the case of finite matrices, with the elements of the product being the appropriate sums of an infinite series. We need to introduce one new concept before we can state our extension of Lemma DDM to
denumerable matrices. A matrix $M$ is said to be bounded if there exists some real number $b$ such that $|M_{ij}| \leq b$ for all $i$ and $j$. Obviously, every finite matrix is bounded, but, for example, a denumerable matrix in which $M_{ij} = i$ is not bounded. We define a denumerable matrix $M \geq 0$ to be dominant diagonal if there exists a bounded diagonal matrix $D \geq 0$ such that the row sums of the matrix $MD$ are all positive.

With the above definitions, Lemma DDM extends to denumerable matrices. But there is one more matter that must be taken care of. Everybody knows that if a finite matrix has an inverse, then it has only one inverse. But this is not true in general of infinite matrices. In fact it is not even true of dominant diagonal matrices. In the appendix we demonstrate these claims by counterexamples. But it is true that denumerable dominant diagonal matrices as we have defined them have only one bounded inverse.

**Lemma DDDM.** Let $I - A$ be a denumerable dominant diagonal matrix such that $A \geq 0$. Then the matrix $B = \sum_{t=0}^{\infty} A^t$ is well-defined and bounded and $B \geq 0$. Furthermore $B$ is the unique bounded inverse matrix for $(I - A)$.

We will also have use for a kind of converse to Lemma DDDM. This is the following.

**Lemma DDDMC.** Let $B \geq 0$ be a finite or denumerable matrix and suppose that $B = (I - A)^{-1}$ for some matrix $A \geq 0$. Then $I - A$ is a dominant diagonal matrix and $B = \sum_{t=0}^{\infty} A^t$.

**Different systems of utility functions that represent the same preferences over allocations**

A thorough understanding of interdependent utility functions and of the interpretation of the dominant diagonal property requires an analysis of the invariance of preferences under transformations of the utility functions. The analogue in the case of independent utility functions to this matter is the invariance of preferences under monotone transformations of utility. But the story here is complicated by the fact one can't simply transform utility representations of individuals "one at a time" since the utility of one individual enters as an argument of other individuals' utility function. Consider two systems of interdependent utility functions defined by the systems of equations $U_i = F_i(u_i(c_i), U_{-i})$ for all $i$, and $U^*_i = F^*_i(u^*_i, U^*_{-i})$. The two systems are said to be equivalent, if they induce systems of independent utility functions that represent the same preferences over allocations. The result that will be most useful for our purposes is the following.

**Lemma ET.** If $(I - A)^{-1}$ exists, then the system of preferences $U = u + AU$ where $u_i(c_i)$ represents $i$'s private preferences is equivalent to the system $U^* = u^* + A^*U^*$ where $u^*_i(c_i)$ represents $i$'s private preferences if and only if $A^* = D^{-1}AD$ for some diagonal matrix $D$ with positive diagonals and $u^* = D^{-1}u + b^*$ for some vector of constants $b^*$.

With the aid of Lemma ET, we can now ask of a system of utility functions $U = u + AU$. "What properties of the matrix $A$ are preserved under equivalent utility representations?" Since any equivalent system represented in separable form will have $U^* = u^* + A^*U^*$ where $A^* = D^{-1}AD$ for some diagonal matrix $D$ with positive diagonal entries, this question is equivalent to asking "What properties of the matrix $A$ are invariant under the transformation $D^{-1}AD$?" Lemma IP states some useful properties of the matrix $A$ that are invariant under equivalent representations.
Lemma IP. If the system of utility functions $U = u + AU$ is equivalent to the system of utility functions $U^* = u^* + A^*U^*$, then for all $i$ and $j$, the sign of the $ij$th entry of $A^*$ will be the same as the sign of the $ij$th entry of $A$. The determinant of $A^*$ will equal the determinant of $A$. The eigenvalues of $A^*$ will be the same as the eigenvalues of $A$. The matrix $I - A$ will be dominant diagonal if and only if the matrix $I - A^*$ is dominant diagonal.

Normally benevolent systems

Recall that an additively separable system of utility functions $U = u + AU$ is normally benevolent if and only if this system induces preferences over allocations which can be represented by a system of equations of the form $U = Bu$ where $B \geq 0$. The following result is an immediate consequence of our earlier discussion and of Lemmas DDM and DDDM.

Proposition 1. A finite or denumerably infinite system of interrelated utility functions of the form $U_i = u_i(c_i) + \sum_{j\neq i} a_{ij}U_j$ where $a_{ij} \geq 0$ for all $i$ and $j$ is normally benevolent if and only if $I - A$ is a dominant diagonal matrix where $A$ is the matrix whose $ij$th entry is $a_{ij}$.

Recall the example of Romeo and Juliet. In that case the matrix $A$ is just the two-by-two matrix with zeroes on the diagonal and $a$ and $b$ on the off-diagonals. The relative size of $a$ and $b$ was not invariant under transformations to equivalent utility systems. But, according to Lemma IP, the determinant of $A$, which is equal to $ab$, is invariant to transformations of the form $D^{-1}AD$. Likewise, the matrix $A$ can be shown to be dominant diagonal if and only if $ab < 1$. Indeed, let $D$ be the two-by-two diagonal matrix with $D_{11} = 1 + a$ and $D_{22} = 1 + b$. Then it is easily verified that the row sums of both the first and second rows of the matrix $I - D^{-1}AD$ are positive.

This observation generalizes to a more useful form of Proposition 1, which follows immediately from Proposition 1 and Lemmas ET and IP.

Proposition 2. A finite or denumerably infinite system of interrelated utility functions $U = u + AU$ is normally benevolent if and only if this system is equivalent to a system $U^* = u^* + A^*U^*$ where the row sums of the matrix $I - A^*$ are all positive.

So far we have discussed ways to determine whether an interrelated system of utility functions is normally benevolent. There is a converse problem that is also important in applications. “Suppose that we observe a system of independent utility functions over allocations that take the form: $U = Bu$ where $B \geq 0$. When does there exist a matrix $A \geq 0$ such that the interrelated system $U = u + AU$ induces the system $U = Bu$?” From Lemma DDDMC we can deduce the following result about interrelated preferences.

Proposition 3. Suppose that preferences of consumers over allocations can be represented by utility functions $U_i = \sum_j b_{ij}u_j(c_j)$ where $b_{ij} > 0$ for all $i$ and $j$. These utility functions are induced by a normally benevolent system of interrelated utility functions of the form $U = u + AU$ if and only if $B = \sum_{t=0}^{\infty} A^t$ where $B$ is the matrix whose $ij$th element is $b_{ij}$.
Can more than one set of preferences over allocations be represented by the same system of utility functions?

In the case of independent utility functions this is such a trivially answered question that it has probably never been explicitly asked. But if preferences are represented by a system of interdependent utility functions, then utility functions over allocations are obtained by "inverting the system". Of course a general system of equations of the form GS.1 needn't have a unique inverse. So it is easy to find systems of interdependent utility functions that have more than one "solution". One such example is \( U_1 = u_1(c_1) + aU_2^2 \) and \( U_2 = u_2(c_2) + U_1^2 \). But if there are a finite number of individuals with utility functions of the separable form \( U = u + AU \), then we know from matrix algebra that if \( I - A \) has an inverse, then it has a unique inverse. Therefore the system of independent utility functions \( U = Bu \) induced by \( U = u + AU \) is uniquely determined.

As we remarked previously, denumerable matrices need not have unique inverses. Therefore there may be more than one linear system \( U = Bu \) that is consistent with the system of interdependent utilities \( U = u + AU \). But according to Lemma DDDM, if \( A \geq 0 \) and \( I - A \) is dominant diagonal, then there is only one bounded matrix \( B \) that is consistent with \( I - A \).

While it is possible to discuss utility functions of the form \( U = Bu \) where \( B \) is unbounded, the preference relation they induce is very odd. For example, there will be time paths of consumption in which consumption is arbitrarily low for arbitrarily long after some generation dies to which that generation assigns an infinite utility and which therefore he finds indifferent to a program that is the same except that he gets an arbitrary high consumption during his own life. These preferences are therefore wildly discontinuous in the kinds of topologies we would like to use for measuring "nearness" of commodity paths. The issues surrounding unbounded utility representations are well-summarized in Koopmans (1972).

Let us define preferences over allocations to regular if they can be represented by a utility function that takes a finite value for every constant allocation \((c, c, \ldots, c)\). Clearly, preferences that can be represented by a utility function of the form \( U = Bu \) will be regular if and only if \( B \) is bounded. From our discussion above, we can conclude that:

**Proposition 4.** If the matrix \( A \geq 0 \) and \( I - A \) is dominant diagonal, then there is only one set of regular preferences over allocations induced by the interdependent utility system \( U = u + AU \).

3. Some Interesting Special Cases

**Two-sided Altruism Across Generations**

A familiar example of a system of interrelated utility functions is the case where there is a sequence of generations such that each generation cares about its own consumption and about the utility of the next generation. Preferences are assumed to be stationary across generations and the utility function of the \( t \)th generation is assumed to be:

\[
U_t = u_t(c_t) + aU_{t+1}
\]

\( (IU.1) \)
where \( 0 < a < 1 \). See, for example, (Barro (1974)). It is well-known that this system of interrelated utility functions induces independent utility functions for generation \( t \) which take the form 
\[
U_t = u_t(c_t) + \sum_{i \geq t} a^{i-t} u_i(c_i).
\]

A more complicated and interesting structure arises if each generation cares not only about its own consumption and the utility of its successor, but each generation also cares about the utility of its parent generation. This is the case that Miles Kimball calls “two-sided altruism”. Let us suppose that preferences are additively separable and stationary over time. The utility function of generation \( t \) is assumed to take the form:
\[
U_t = u_t(c_t) + aU_{t-1} + bU_{t+1}.
\]

We want to know, for what values of \( a \) and \( b \) the system \( IU.2 \) is normally benevolent. As it turns out, not only can we answer this question using the general theorems of the preceding section, but we can actually invert this system of equations to solve explicitly for the independent utility functions over allocations that are induced by \( IU.2 \). The solution that I find is equivalent to that found by Kimball but expressed in a different, and I think, slightly more transparent way. \(^2\)

Where benevolent preferences look forward, as in the system \( IU.1 \), in order to have a truly stationary model, it is necessary to have an infinite future horizon. Otherwise there will be a “last generation” that is different from all other generations in that it has no descendents. Furthermore the next to the last generation will be different from earlier generations since its descendent will be the last generation—and so on. A system of this kind can be solved using matrix methods, but the answers are a bit messy. For generations that are far from the end, the solutions are close to the solutions for the stationary case. When benevolent preferences look back to the preceding generation as well as forward to the next generation, in order to have a stationary model we need an infinite past as well as an infinite future. Otherwise their will be a “first generation” that is different from subsequent generations because it has no parents. Moreover, subsequent generations will differ because of their differing proximity to the first generation.

The system \( IU.2 \) can be written as a matrix equation \( U = u + AU \) where \( A \) is a doubly infinite matrix with the entry \( a \)'s on the first subdiagonal and with \( b \)'s on the first superdiagonal. Let \( J \) be the matrix with ones on the first subdiagonal (that is, \( J_{ij} = 1 \) if \( j = i + 1 \) and with zeroes everywhere else. Similarly, let \( J^{-1} \) be the matrix with ones on the first subdiagonal \( (J_{ij} = 1 \) if \( j = i - 1 \) and zeroes everywhere else). Then \( I - A = I - aJ^{-1} - bJ \). It can be seen that \( I - A \) is dominant diagonal if and only \( 1 - a - b > 0 \). Therefore the system \( IU.2 \) is normally benevolent if and only if \( 1 - a - b > 0 \).

With this special structure we can actually solve for \((I - A)^{-1}\). We know from Lemma DDDM that \( I - A = I - aJ^{-1} - bJ \) will have one and only one bounded, non-negative inverse matrix if and only if \( 1 - a - b > 0 \). A good way to find this inverse is to factor \( I - aJ^{-1} - bJ \) into the product of two matrices with easily computed inverses. We show in the appendix how this is done. The result has a very clean, simple form. In particular, future utilities which are \( k \) generations removed are discounted at a rate \( \beta^k < 1 \) and utilities

\(^2\) In his paper, Kimball solves a system of difference equations to find the equivalent of a row of our matrix \((I - A)^{-1}\). He finds two roots for his system of difference equations and therefore represents the typical element of the matrix as a linear combination of power series in the two roots. As it turns out, and as Kimball observes, one of Kimball's roots exceeds one in absolute value. Therefore the only bounded matrix that satisfies Kimball's difference equations is the one which is a power series in the smaller root.
of past generations which are $k$ generations removed are discounted at a rate $\alpha^k < 1$ where $\alpha/\beta = a/b$.

**Proposition 5.** Suppose that there is a denumerable infinity of consumers with interrelated utility functions of the form:

$$U_t = u_t(c_t) + aU_{t-1} + bU_{t+1}$$

for every generation $t$ where $a \geq 0$ and $b \geq 0$. This system is normally benevolent if and only if $a + b < 1$. If $a + b < 1$, then the preferences of the $t$th generation over allocations can be represented by a utility function of the form:

$$U_t = \sum_{j=1}^{\infty} \alpha^j u_t(c_{t-j}) + u_t(c_t) + \sum_{j=1}^{\infty} \beta^j u_t(c_{t+j})$$

where $0 < \alpha < 1$, $0 < \beta < 1$ and $\frac{\alpha}{\beta} = \frac{a}{b}$. More specifically, $\alpha = \frac{1-\sqrt{1-4ab}}{2b}$ and $\beta = \frac{1-\sqrt{1-4ab}}{2a}$. The preferences over allocations represented by these utility functions are the only regular preferences that are induced by the original system of interrelated utility functions.

**Benefactors who want to see their gifts consumed**

Suppose that each generation is concerned about the happiness of its predecessors and of its successors but is more concerned about consumption that it can observe than about consumption that it cannot observe. An additively separable model that has this character is the following. For each generation, $U_t = \sum_{i=-n}^{m} k_t i u_i(c_t) + \sum_{i=-\infty}^{\infty} a_t U_i$ where the consumer is specially interested in the consumption of the last $n$ generations and of the next $m$ generations. This system of utility functions is represented by the matrix equation $U = Ku + AU$ or equivalently, $(I - A)U = Ku$. If $I - A$ is dominant diagonal, we can solve this system for the independent system of utilities defined over allocations. We then have $U = (I - A)^{-1} Ku$ where $(I - A)^{-1}$ is a nonnegative matrix. If $K$ is also a nonnegative matrix, then $(I - A)^{-1} K$ will also be a nonnegative matrix and so the system of utility functions will be normally benevolent.

A special case of this type is where for each generation $t$,

$$U_t = k_1 u(c_{t-1}) + u(c_t) + k_2 u(c_t) + aU_{t-1} + bU_{t+1}.$$ 

This system can be written as the matrix equation $U = k_1 J^{-1} u + u + k_2 J u + AU$ where $A = aJ^{-1} + bJ$. Then $U = (I - A)^{-1} (k_1 J^{-1} + I + k_2 J) u$. Recalling our solution for $(I - A)^{-1}$ in the case of two-sided altruism, we can now compute the vector $(I - A)^{-1} (k_1 J^{-1} + I + k_2 J) u$. The $t$th element of this vector is seen to be

$$\sum_{j=1}^{\infty} \alpha^j (k_1 \alpha^{-1} + 1 + k_2 \alpha) u(c_{t-j}) + (k_1 \alpha + 1 + k_2 \beta) u(c_t) + \sum_{j=1}^{\infty} \beta^j (k_1 \beta + 1 + k_2 \beta^{-1}) u(c_{t+j})$$

A more elaborate model could take account of the fact that one only observes consumption during the parts of the lifetime of adjacent generations when one is, oneself alive.
where, as in the previous discussion, \(a = \frac{\alpha}{\alpha + \beta}\) and \(b = \frac{\beta}{\alpha + \beta}\). One can divide this expression through by the positive number \(k_1\alpha + 1 + k_2\beta\) to obtain an equivalent and more neatly expressed representation of the same preferences. This is

\[
U_t = k_1^* \sum_{j=1}^{\infty} \alpha^j u(c_{t-j}) + u(c_t) + k_2^* \sum_{j=1}^{\infty} \beta^j u(c_{t+j})
\]

where \(k_1^* = \frac{k_1\alpha^{-1} + 1 + k_2\beta}{k_1\alpha + 1 + k_2\beta}\) and \(k_2^* = \frac{k_1\alpha + 1 + k_2\beta}{k_1\alpha + 1 + k_2\beta}\). Each generation will agree with its successor about the part of the allocation having to do with the third generation and later, but will not necessarily agree about how to allocate resources between the second and third generation. It is also worth noticing that the weights on the period utilities are nonnegative in this model so long as \(a + b < 1\) and the \(k_i\)'s are nonnegative even if one or both of the \(k_i\)'s are bigger than unity.

**On intergenerational consistency**

An important question in the theory of intergenerational preferences is "When is it true that each generation agrees with its successor about how to allocate resources among all future generations?" This question was posed by Strotz (1956). The implications of intertemporal inconsistency for intergenerational planning were discussed by Pollak (1968) and Blackorby, Nissen, Primont and Russell (1973) among others. We say that there is forward-looking intergenerationally consistency if every generation \(t\) agrees with its successor about how to allocate consumption among all generations starting from \(t + 1\). Most of the earlier discussion in the literature deals with the case of "one-sided altruism". For a stationary economy with one-sided altruism, there is forward-looking intergenerational consistency if preferences of generation \(t\) are representable by a utility function of the form

\[U_t = U(u_t(c_t), U_{t-1}, U_{t+1})\]

where \(U(\cdot, \cdot)\) is an increasing function of both of its arguments. It is immediate from inspection of this utility function that any change in intergenerational consumption streams that leaves the consumption of generation \(t\) unchanged will increase \(U_t\) if and only if it increases \(U_{t+1}\).

The question of forward-looking intergenerational consistency in the case of two-sided altruism was raised in an interesting debate between Burbridge (1983), (1984), and Buiter and Carmichael (1984) and was further clarified by Kimball (1986). In the case of two-sided altruism, we have \(U_t = U(u_t(c_t), U_{t-1}, U_{t+1})\) where \(U(\cdot, \cdot\cdot)\) is an increasing function of each of its arguments. Since \(U_t\) depends on \(U_{t-1}\) as well as on \(u_t(c_t)\) and \(U_{t+1}\), and since generation \(t - 1\) is concerned about future consumption bundles, it is not immediately clear that generation \(t\) will agree with its successor about consumption allocations among generations after \(t\). All that is immediate from the monotonicity of \(U_t\) as a function of \(U_{t+1}\) is that generation \(t\) agrees with its successor about how to allocate resources among future generations so long as the allocations being compared are of equal utility to generation \(t - 1\).

Surprisingly, in the case of additively separable two-sided altruism, preferences turn out to have forward-looking intergenerational consistency after all. From Proposition 4 of the previous section it follows that if \(U_t = aU_{t-1} + u_t(c_t) + bU_{t+1}\) for all \(t\), then one can represent preferences of generation \(t\) over future allocations, holding consumption of generations up until \(t\) constant, by a utility function of the form \(U_t = C + \sum_{i=1}^{\infty} \beta^i u_t(c_t)\). Preferences of generation \(t + 1\), holding consumption of generations up until \(t\) constant,
can be represented by $U_{t+1} = C' + \sum_{i=1}^{\infty} \beta^{i-1} u_i(c_i)$. Therefore it is clear that generation $t$'s preferences over allocations among generations starting with $t + 1$ are represented by a utility function that is an increasing affine transformation of the function that represents generation $t + 1$'s preferences over the same set of allocations. It follows that if all of the consumptions before generation $t$ are already determined, then generation $t$ and generation $t + 1$ will agree about everything else. Therefore there is forward-looking intergenerational consistency.

This conclusion appears to contradict the view expressed by Burbridge who argues that "consistent behavior" requires that utility functions over allocations must take the form $U_t = \sum_{i=-N}^{\infty} \beta^i u_{t+i}$. Notice that Burbridge's proposed utility function is strikingly different from the utility function corresponding to two-sided altruism as stated in Proposition 4. For Burbridge, more distant historical periods have geometrically increasing weights. For the case of two-sided altruism, more distant historical periods have geometrically decreasing weights.

**Proposition 6.** If every generation has a utility function of the form $U_t = aU_{t-1} + u_t(c_t) + bU_{t-1}$, then there is forward-looking intergenerational consistency of preferences.

There is a symmetric result concerning backward-looking consistency. While it is not possible, given current technology, to go back and rearrange the past, it might interest historians to notice that with two-sided altruism of this type, all generations will agree with their predecessors about what they would have liked their history to have been. If this were the case, there would be little work for historical revisionists.

4. Conclusion

I have studied systems of utility functions in which the utility functions of two or more people simultaneously depend on each others' utilities. In the cases that I study, these interdependent utility systems induce preferences of each individual over allocations of goods among the interrelated population. For the purposes of demand theory, it is these latter preferences that are of the greatest interest, since it is simpler to study how one buys commodities for others than to study how one buy utility for others. On the other hand, introspection and casual observation often seem to offer stronger and more interesting hypotheses about the nature of interdependent utilities than they do about preferences over allocations. Hypotheses about the nature of utility interdependence imply special structure for the preference over allocations. The purpose of this paper has been to explain the way in which this happens. I have resisted the temptation to expand our discussion from preferences on to demand theory. That is the logical next step. For example the special utility structure that we find in the case of two-sided altruism leads to some interesting implications for the economics of social security and inheritance. But that is a subject for another paper.

There is another application of this theory that I think is worth pursuing. That is the theory of memory and anticipation for a single individual. The formalism of two-sided altruism, for example, could be interpreted as follows. Suppose that my current happiness, $U_t$, depends on my current consumption, $c_t$, on how happy I expect to be next period, $U_{t+1}$, and how pleasant my memory of past happiness is, $U_{t-1}$. In a simple example, then I might have $U_t = u_t(c_t) + aU_{t-1} + bU_{t+1}$ for all time periods of my life. This hypothesis about memory and anticipation induces a special structure on intertemporal preferences
just as hypotheses about interrelated utility functions induce preferences on allocations. One important difference for the model of individuals is that finiteness of life is now an essential part of the story. Everybody's life has a beginning and an end. Young people are different from old people both because they have a shorter past, and a longer future. Explicit models of memory and anticipation of the kind we have proposed give a useful way of taking this into account. It would be interesting to study more sophisticated models of this type.
5. Appendix

Proof of Lemma DDDM

Let $A$ be a nonnegative denumerable matrix and let $(I - A)D$ have positive row sums for some bounded diagonal matrix $D$. This proof follows the proof for the finite-dimensional case which was presented by David Gale (1960). The only extra ingredient that we need that is not required for the finite dimensional case is that we must use the boundedness of the matrix $D$ when we demonstrate that $\lim_{n \to \infty} (I + A + A^2 + \ldots + A^n)$ exists.

If $I - A$ is dominant diagonal, there exists a bounded vector $x^* \geq 0$ such that $Ax^* \ll x^*$. (The vector $x^*$ whose components are the diagonal elements of the bounded diagonal matrix $D$ such that $(I - A)D$ has positive row sums has this property.) Therefore there exists a real number $\lambda$ where $0 < \lambda < 1$ and $Ax^* \ll \lambda x^*$. By induction it follows that $A^n x^* \ll \lambda^n x^*$.

For any integer $n$, define the matrix $B(n) = (I + A + A^2 + \ldots + A^n)$. Since the matrix $A$ is non-negative it must be that for all $i$ and $j$, $B_{ij}(n + 1) \geq B_{ij}(n)$. Furthermore since $Ax^* \ll \lambda x^*$, it follows that for all $n$,

$$B(n)x^* \leq (1 + \lambda + \lambda^2 + \ldots + \lambda^n)x^* < \frac{1}{1 - \lambda}x^*.$$

But since $x^* \gg 0$ is a bounded vector and since $B(n) \geq 0$ for all $n$, it follows that the sequence $B_{ij}(n)$ is bounded. Since a bounded monotonic sequence must converge to a limit, it follows than $\lim_{n \to \infty} B_{ij}(n) = B_{ij}$ exists and hence the matrix $B = \lim_{n \to \infty} B(n)$ is well defined. It is straightforward to verify that $B \geq 0$ and $B(I - A) = (I - A)B = I$.

We also need to show that $B$ is the only bounded inverse for $(I - A)$. If $B$ and $B'$ are both bounded inverses, then $B(I - A) = B'(I - A) = (I - A)B = (I - A)B' = I$. Then $(I - A)(B - B') = 0$. Since $B$ and $B'$ are, by assumption, bounded matrices, the matrix $B - B'$ will also be bounded. Therefore every column of $B - B'$ is a bounded vector. Suppose that $(I - A)x = 0$ for a bounded vector, $x$. Where $B(n) = (I + A + A^2 + \ldots + A^n)$, $0 = B(n)(I - A)x = (I - A^n)x$. From our argument of the previous paragraph it follows that $\lim_{n \to \infty} A^n = 0$. Therefore if $x$ is a bounded vector it must be that $x = 0$. From this it follows that every column of $B - B'$ is a zero vector and hence that $B = B'$.

Proof of Lemma DDDMC

Suppose $B = (I - A)^{-1} \geq 0$ and $A \geq 0$. Choose an arbitrary vector $\bar{x} \gg 0$. Let $x^* = B\bar{x}$. Then $(I - A)x^* = (I - A)B\bar{x} = \bar{x} \gg 0$. Set $D$ be the diagonal matrix whose diagonal entries are the elements of $x^*$. The row sums of $(I - A)D$ must all be positive. Therefore $I - A$ is dominant diagonal. It follows from Lemma DDDM that $(I - A)^{-1} = \sum_{t=0}^{\infty} A^t$. 

14
Details in the proof of Proposition 5.

When \(a > 0\) and \(b > 0\) and \(a + b < 1\), we can find a unique pair of parameters \(\alpha\), and \(\beta\) such that \(0 < \alpha < 1\) and \(0 < \beta < 1\), and such that \(a = \frac{\alpha}{1 + \alpha\beta}\) and \(b = \frac{\beta}{1 + \alpha\beta}\). This claim requires proof, which is supplied in the form of the Lemma BIJ, below.

A simple calculation verifies that

\[
I - \frac{\alpha}{1 + \alpha\beta}J^{-1} - \frac{\beta}{1 + \alpha\beta}J = \frac{1}{1 + \alpha\beta}(I - \alpha J^{-1})(I - \beta J).
\]

If \(\beta < 1\) and \(\alpha < 1\), then \((I - \beta J)\) and \((I - \alpha J^{-1})\) are dominant diagonal matrices. It follows from Lemma DDDM that \((I - \beta J)^{-1} = I + \sum_{t=1}^{\infty} \beta^t J^t\) and \((I - \alpha J)^{-1} = I + \sum_{t=1}^{\infty} \alpha^t J^{-t}\).

Multiplying these two matrices together, we find that \((I - \beta J)^{-1}(I - \alpha J^{-1}) = I + \frac{1}{1 - \alpha\beta}B\) where \(B\) is the matrix such that \(B_{ij} = \beta^{j-i}\) for \(j \geq i\) and \(B_{ij} = \alpha^{i-j}\) for \(i \geq j\). Therefore \(U = \frac{1 + \alpha\beta}{1 - \alpha\beta}Bu\). Since \(\frac{1 + \alpha\beta}{1 - \alpha\beta} > 0\), the same preferences can be represented more simply by \(U = Bu\). Carrying out this matrix product for a single row, we find the formula stated in Proposition 5 for a representative generation’s utility over allocations.\(^4\)

Lemma BIJ. Let \(X\) be the set \(\{(\alpha, \beta) \mid 0 < \alpha \leq 1, 0 < \beta \leq 1\}\). Let \(Y\) be the set \(\{(a, b) \mid 0 < a + b \leq 1\}\). Let \(F\) be the function from \(X\) to \(Y\) such that \(F(\alpha, \beta) = (\frac{\alpha}{\alpha + \beta}, \frac{\beta}{\alpha + \beta})\). The function \(F\) is a bijection (i.e. one-to-one and onto \(Y\)). The unique choice of \((\alpha, \beta)\) for which \(F(\alpha, \beta) = (a, b)\) is \((1 - \frac{\beta}{2b}, \frac{1 - \beta}{2a})\).

**Proof**

First we show that \(F(\alpha, \beta) \in Y\) for all \((\alpha, \beta)\) in \(X\). If \(a = \frac{\alpha}{1 + \alpha\beta}\) and \(b = \frac{\beta}{1 + \alpha\beta}\), then \(a \geq 0, b \geq 0\), and \(a + b = \frac{\alpha + \beta}{1 + \alpha\beta}\). But \(1 + \alpha\beta - (\alpha + \beta) = (1 - \alpha)(1 - \beta) \geq 0\). Therefore \(a + b \leq 1\). It follows that \((a, b) \in Y\).

Now we show that if \((a, b) \in Y\) then there is exactly one \((\alpha, \beta)\) in \(X\) such that \((a, b) = F(\alpha, \beta)\). Let \(x = 1 + \alpha\beta\). If \((a, b) = F(\alpha, \beta)\), then \(\alpha = ax\) and \(\beta = bx\). To solve for \(x\), we see that \(\alpha\beta = abx^2\) or equivalently, \(x - 1 = abx^2\). Solving this quadratic equation for \(x\) yields two roots. These are \(x = \frac{1 \pm \sqrt{1 - 4ab}}{2}\). Therefore there are two choices of \((\alpha, \beta)\) such that \((a, b) = F(\alpha, \beta)\). But only the root \(x = \frac{1 - \sqrt{1 - 4ab}}{2}\) gives \((\alpha, \beta)\) in \(X\). We check this last assertion as follows. If \((\alpha, \beta) \in X\), then \(1 > \alpha\beta = abx^2\). Carrying out the multiplication for \(x^2\), and rearranging terms, one finds that for \(x = \frac{1 - \sqrt{1 - 4ab}}{2ab}\), \(abx^2 < 1\) if and only if \(1 - 4ab < \sqrt{1 - 4ab}\) and for \(x = \frac{1 + \sqrt{1 - 4ab}}{2ab}\), \(abx^2 > 1\) if and only if \(1 - 4ab < \sqrt{1 - 4ab}\). Whenever \(a\) and \(b\) are nonnegative and \(a + b < 1\), it must be that \(0 < 1 - 4ab < 1\) and hence \(1 - 4ab < \sqrt{1 - 4ab}\). Therefore we conclude that only the root \(x = \frac{1 - \sqrt{1 - 4ab}}{2ab}\) gives

\(^4\) There are also some unbounded matrices \(B\) such that \(U = Bu\) is a solution to \((I - A)U = u\). To see where these come from, recall that when we parameterized the system with \(a = \frac{\alpha}{1 + \alpha\beta}\) and \(b = \frac{\beta}{1 + \alpha\beta}\), we found that there was exactly one solution for \((\alpha, \beta)\) such that \(0 \leq \alpha < 1\) and \(0 \leq \beta < 1\). But there is also a solution, \((\alpha', \beta')\) where \(0 \leq \alpha' < 1\) and \(0 \leq \beta' < 1\). The matrix \((I - \alpha J^{-1} - bJ)\) can also be factored into the expression \((I - \alpha' J^{-1})(I - \beta' J)\). But when the inverse of this matrix is computed, it turns out to be unbounded.
(\alpha, \beta) in X. Then \((\alpha, \beta) = \left(\frac{1 - \sqrt{1 - 4a\beta}}{2a}, \frac{1 - \sqrt{1 - 4a\beta}}{2a}\right)\) is the one and only point in X such that \(F(\alpha, \beta) = (a, b)\).

Proof of Lemma ET

Suppose that \(U = u + AU\) and \(U^* = u^* + A^*U^*\) are equivalent systems of interdependent utility functions. The system of independent preferences on allocations induced by \(U = u + AU\) is represented by the matrix equation \(U = Bu\) where \(B = (I - A)^{-1}\). A standard result of consumer theory (see e.g. Debreu, 1960) is that if two additively separable functions represent the same preferences, they must be affine transformations of each other. Therefore if \(U^*\) represents the same preferences as \(U\) and is also additively separable in its arguments, it must be that for each \(i\), \(U_i = k_iU_i^* + b_i\) for some constants \(k_i > 0\) and \(b_i\).

Therefore \(U = DU^* + b\) where \(D\) is a diagonal matrix with \(D_i = k_i\). Substituting \(DU^* + b\) for \(U\) in the expression \(U = u + AU\), we have: \(DU^* + b = u + ADU^* + Ab\). Premultiplying by \(D^{-1}\) and rearranging terms, we find that \((I - D^{-1}AD)U^* = u^*\) where \(u^* = D^{-1}u + b^*\) for some vector of constants \(b^*\). Further rearrangement shows that \(U^* = u^* + A^*U^*\) where \(A^* = D^{-1}AD\). This establishes that if the systems \(U = u + AU\) is equivalent to \(U^* = u^* + A^*U^*\) then \(A^* = D^{-1}AD\) for some diagonal matrix \(D\) with positive diagonals and \(u^* = D^{-1}u + b^*\) for some vector of constants \(b^*\). The converse proposition is a matter of straightforward verification.

Proof of Lemma IP

According to Lemma ET, the two systems will be equivalent only if \(A^* = D^{-1}AD\) for some diagonal matrix \(D\) with positive diagonal entries. With this fact in mind, each of the first three claims in Lemma IP follows from elementary results of matrix algebra. If \(I - A\) is dominant diagonal, then it is immediate from the definition that \((I - A)D\) is dominant diagonal where \(D\) is any positive diagonal matrix. But if \((I - A)D\) is dominant diagonal, then \(D^{-1}(I - A)D = I - D^{-1}AD\) must also be dominant diagonal since premultiplying by a positive diagonal matrix simply multiplies every row sum by a positive constant.
Bibliography


Recent CREST Working Papers


87-21: Jeffrey A. Miron and Stephen P. Zeldes, “Production, Sales, and the Change in Inventories: An Identity that Doesn’t Add Up” June 1987.


MichU Bergstrom, Ted
DeptE Systems of Benevo-
CorREST lent Utility Inter-
AUTHOR dependence
TITLE 4

<table>
<thead>
<tr>
<th>DATE DUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/25/87</td>
</tr>
</tbody>
</table>