ON DIVIDING AN AMOUNT ACCORDING TO
INDIVIDUAL CLAIMS OR LIABILITIES*†

H. PEYTON YOUNG

University of Maryland

A basic principle of distributive justice states that if an allocation among a group of
individuals is fair, then it should be perceived as fair when restricted to each subgroup of
individuals. This ‘consistency’ principle applies in particular to methods for allocating taxes
among citizens according to their ability to pay. Every consistent, continuous taxation method
optimizes an additively separable objective function. Optimal allocations may be computed by
a simple Lagrange multiplier technique. Similar results hold for allocating assets among
creditors according to their claims. The techniques are illustrated by constructing an objective
function for a bankruptcy method from the Babylonian Talmud.

1. Introduction. How should a tax be allocated among a group of individuals
according to their ability to pay? This is one of the central issues in the theory of
public finance, about which there is a centuries-old literature [13]. In posing this
question we do not mean to suggest that we have a particularly new or startling answer
to the tax problem. Rather, our aim is to show how traditional solutions to taxation
relate to a much wider class of problems in distributive justice. These problems are
concerned with how to allocate a resource according to individuals’ claims on it, or
how to allocate a cost based on individuals’ liabilities for it. Three other examples of
this type spring to mind. In a bankruptcy case the assets of the bankrupt entity are to
be allocated fairly among the creditors according to their claims. In a business
partnership, the gains and losses are divided among the partners according to their
contributions. In a legislature, seats are apportioned among the constituent states
according to their populations (or among the parties according to their vote totals). All
of these problems have three elements in common: there is a fixed quantity to be
distributed among a specified group of individuals; each individual has a numerical
claim on the resource (or a potential liability for the cost); the focus is on the fairness
or justice of the solution.

The reasons for choosing taxation as the motivating example are, first, familiarity;
second, the richness of solutions. More particularly we are interested in the form
of these solutions. Inevitably, taxes are allocated via tax schedules. How could it be
otherwise? Yet we shall show that this type of solution involves a crucial assumption
that forms the principal focus of our study. Namely, it takes for granted that how much
subgroup of individuals shares a given amount of tax should depend only on their own
taxable incomes. This assumption seems innocuous enough, yet it has far-reaching

*This work was supported by the National Science Foundation under Grant SES-831 9530 at the
University of Maryland.
†Received February 12, 1985; revised February 20, 1986.
OR/MS Index 1978 subject classification. Primary: 231 Games/group decisions Secondary: 509 Organi-
tional studies/motivation/Incentives.
Key words. Resource allocation, fairness, optimality.

398
consequences. Interpreted as a general principle of distributive justice, it says that a fair allocation is fair when restricted to any subgroup of the population. To put it another way, in order to know whether a subgroup of individuals is treated fairly relative to each other, it is enough to know what their claims are and how much they are allotted; not on how much the others are allotted.

This property of an allocation method will be called consistency. It says that the criterion of equity on which a method rests should apply locally; equity can be judged on the basis of pairwise (and n-wise) comparisons. This idea reaches well beyond taxation. For example, it figures in the apportionment of representation in legislatures [2]. If a state with a population of \( p \) is given \( a \) seats, and another with population \( p' \) is given \( a' \) seats, then this should represent a fair apportionment of \( a + a' \) seats among these two states irrespective of how the other states share their seats. This property was implicit in E. V. Huntington's early 'pairwise comparisons' approach to apportionment [7]. It was formulated as a general axiom of fair division by this author and M. L. Balinski [2, p. 141]. In the case of apportionment it characterizes an historically important family of methods (known as rank-index methods) that are widely used in practice [2, Appendix, Section 8]. For example, the current method for apportioning the U.S. House of Representatives is of this type.

The same principle lies behind an ancient Talmudic method for adjudicating bankruptcy cases. In this context consistency says that the way any two (or any \( n \)) creditors divide their portion of the assets should depend only on their own claims against the assets [1]. This principle also lies behind some well-known bargaining solutions. Harsanyi pointed out that the Nash solution is consistent [6], and recently T. Lensberg has given a beautiful characterization of Nash's and other bargaining solutions based on this property [8], [9], [10]. Other solution concepts that have been axiomatized via consistency are the nucleolus of cooperative games (Sobolev [14]), and egalitarian profit-sharing methods (Moulin [11]).

In this paper we shall show that every consistent (continuous and symmetric) allocation method must be based on something akin to a tax schedule. Secondly, we shall show that any method based on such a tax schedule must, in effect, maximize an additive measure of 'welfare'; equivalently, it must minimize some additive measure of 'loss' or 'sacrifice'. The meaning of this result is that a criterion of distributive equity is based on pairwise comparisons (i.e., is in a sense 'decentralized') if and only if it is formally equivalent to maximizing the sum of individuals' gains or minimizing the sum of individuals' losses. Equity and optimality are thus revealed to be intimately related concepts.

The plan of the paper is as follows. In §2 we give a precise definition of consistency. §3 defines the concept of a parametric tax schedule, which is the form of many historical tax proposals. §4 shows that a continuous, symmetric method is consistent if and only if it is representable by a parametric tax schedule. §5 shows the equivalence between consistency and minimizing an additive loss function in the presence of continuity and symmetry. We also discuss the connections between these results and Lensberg's work in bargaining theory. §6 shows how the methods of proof developed in the preceding sections can be used to construct an objective function for the Talmudic bankruptcy method. §7 describes how the theory can be adapted to profit-sharing methods.

2. Definition of terms. An allocation problem is a pair \((x; T)\), where \(x \in \mathbb{R}_+^I\) is a vector of positive real numbers (the claims), which are indexed by some finite nonempty subset \(I\) of positive integers (the individuals), and \(T\) is the total to be allocated, \(0 \leq T \leq \Sigma x_i\). A solution is a vector \(t \in \mathbb{R}_+^I\) such that \(\Sigma t_i = T\) and \(0 \leq t_i \leq x_i\).
for all \( i \in I \). An allocation method is a function \( F \) that assigns to every problem a unique solution \( t = F(x; T) \).

We shall only consider methods that are symmetric in the sense that allocations are determined by individuals' claims, not by who they are; that is, if \( t = F(x; T) \) and \( x_i = x_j \), then \( t_i = t_j \). Further we shall require that \( F \) be continuous on every domain \( \{(x, T) \in \mathbb{R}^I \times \mathbb{R} : x > 0, 0 \leq T \leq \Sigma x_i\} \) for every nonempty, finite subset \( I \) of positive integers. For the sake of concreteness we shall think of \( F \) as a method of taxation. In this case \( x_i \) represents \( i \)'s taxable income, \( t_i \) is the tax assessed on \( i \), and \( T \) is the total amount of tax to be levied.

In general, a method \( F \) is consistent if for any finite nonempty subset \( I \) of individuals

\[
\forall x \in \mathbb{R}_{++}^I, \forall J \subseteq I \ (J \neq \emptyset) \ [t = F(x; T) \implies t_J = F(x_J; \Sigma_J t_i)].
\]

A weaker condition is pairwise consistency, which requires (1) only for all pairs of individuals, i.e., whenever \( |J| = 2 \). In the presence of continuity and symmetry, pairwise consistency implies consistency, as we shall subsequently see.

Consistency states that an allocation should be invariant when viewed by any subgroup of individuals. In the context of taxation, it says that the way a given amount of tax is shared among a subgroup should depend only on the abilities to pay of those individuals and the amount that they have to divide. Suppose for example that three individuals have taxable incomes of $10,000, $20,000 and $30,000 respectively, and that they pay $2,000, $7,000, and $11,000 in taxes. It seems reasonable that the same method, when applied to allocating a tax of $9,000 among two individuals with incomes of $10,000 and $20,000, should allocate it as $2,000 and $7,000 and not in some other way. Also, if $18,000 is to be allocated among two individuals with incomes of $20,000 and $30,000, it should be divided as $7,000 and $11,000, and so forth. In fine, whatever distributive principle is embodied in a method \( F \), it should be a principle that applies uniformly and consistently to any pair of individuals, to any threesome, indeed to any size subgroup.

3. Parametric methods. The method of proportional division is clearly consistent. So is the method of equal division (the head tax) in which everyone is taxed the same amount, subject to ability to pay. The mirror image of the head tax is to tax everyone so that their after-tax incomes are equal, subject to no one being subsidized. These methods all have simple analytic descriptions. Let \( \lambda \) be a parameter that depends implicitly on the total tax \( T \) to be levied. The proportional method (or flat tax) is given by the formula

\[
t_i = \lambda x_i, \quad 0 \leq \lambda \leq 1,
\]

where \( \lambda \) is set so that \( \Sigma t_i = T \). The equal division method (or head tax) is given by the formula

\[
t_i = \min\{\lambda, x_i\}, \quad 0 \leq \lambda \leq \infty,
\]

where \( \lambda \) is determined so that \( \Sigma t_i = T \). The levelling tax has the representation

\[
t_i = \max\{x_i - 1/\lambda, 0\}, \quad 0 \leq \lambda \leq \infty,
\]

where \( \lambda \) is chosen so that \( \Sigma t_i = T \).

In general, let \( f(x, \lambda) \) be any real-valued function of two scalar variables \( x \) and \( \lambda \), where \( x > 0 \), and \( \lambda \) ranges over some closed interval \([a, b]\) of the extended reals. For
each \(x, f\) is assumed to be weakly monotone increasing and continuous in \(\lambda\) with \(f(x, a) = 0\) and \(f(x, b) = x\). All of the above parametric descriptions have these properties. (The requirement that \(f\) be monotone increasing in \(\lambda\) is the reason for \(1/\lambda\) instead of \(\lambda\) in (4).)

Given such an \(f\), define an allocation method \(F\) as follows. For every problem \((x, T)\)

\[
t = F(x, T) \quad \text{iff} \quad \exists \lambda \forall i \quad [t_i = f(x_i, \lambda) \quad \text{and} \quad \sum t_i = T].
\] (5)

Such a method \(F\) is said to be parametric with representation \(f\). Many classical tax formulas are of the parametric type. For example, in 1889 A. J. Cohen Stuart [4], and later, Edgeworth [5], suggested a family of progressive tax schedules that include the following:

\[
t_i = \max \left\{ 0, x_i - x_i^\lambda \right\}, \quad 0 \leq \lambda \leq \infty.
\]

In 1901, G. Cassel [3] proposed the following family of schedules:

\[
t_i = x_i^2/(x_i + 1/\lambda), \quad 0 \leq \lambda \leq \infty.
\]

Another interesting class of parametric tax functions are those that cause all taxpayers to 'sacrifice equally' in loss of utility. One postulates a continuous, strictly increasing utility function \(u(x), u(x) - u(x - t) = \lambda\), where \(\lambda\) is the constant level of sacrifice. Inverting \(u\) yields the tax function \(t = x - u^{-1}[u(x) - \lambda]\), which is parametric in the level of sacrifice \(\lambda\). The class of 'equal sacrifice' methods is treated axiomatically in [15] and [16]. We cite these examples in order to underline the great variety of parametric tax methods, and to suggest the relevance of methods other than the garden-variety types such as equal or proportional division.

4. A representation theorem for consistent methods. A method \(F(x, T)\) is continuous if it is jointly continuous in both of its arguments. Every parametric method is symmetric, and consistent. We shall show that, in the presence of continuity, the converse also holds. In fact, it is enough to assume only pairwise consistency, together with symmetry and continuity, to explicitly construct a parametric representation of a method.

Theorem 1. A continuous taxation method \(F\) is symmetric and pairwise consistent if and only if it is representable by a continuous parametric function.

Lemma 1. If \(F\) is symmetric, continuous, and pairwise consistent, then \(F\) is monotone increasing in \(T\), that is

\[
0 \leq T \leq T' \leq \sum x_i \Rightarrow F(x; T) \leq F(x; T').
\]

Proof of Lemma 1. The proof is by contradiction: suppose that \(F\) satisfies the hypotheses of Lemma 1 but is not monotone increasing in \(T\). By pairwise consistency there exists a two-agent problem that is not monotonic. That is (by symmetry) there exists some pair \((x_1, x_2) > 0\) such that

\[
(t_1, t_2) = F(x_1, x_2; t_1 + t_2), \quad (t'_1, t'_2) = F(x_1, x_2; t'_1 + t'_2),
\]

where \(t_1 + t_2 < t'_1 + t'_2\), \(t_2 > t'_2\), and \(t_1 < t'_1\).

Choose an integer \(n\) sufficiently large that

\[
t_1 + nt_2 > t'_1 + nt'_2
\] (6)
Consider the \((n + 1)\)-agent problem \(x = (x_1, x_2, x_3, \ldots, x_2)\) in which \(x_2\) is replicated \(n\) times. For all \(T\) such that \(0 \leq T \leq x_1 + nx_2\), define

\[
\alpha(T) = F_1(x_1, x_2, x_3, \ldots, x_2; T) + F_2(x_1, x_2, x_3, \ldots, x_2; T).
\]

Thus \(\alpha(T)\) is the combined amount that agents 1 and 2 get when the total is \(T\) in the \((n + 1)\)-agent problem. \(\alpha\) is a continuous function of \(T\) and \(\alpha(0) = 0\), \(\alpha(x_1 + nx_2) = x_1 + x_2\). By continuity there exists \(T'\), \(0 \leq T' \leq x_1 + nx_2\), such that \(\alpha(T') = t'_1 + t'_2\).

By pairwise consistency and symmetry,

\[
F(x_1, x_2, x_3, \ldots, x_2; T') = (t'_1, t'_2, t'_2, \ldots, t'_2).
\]

Therefore \(T' = t'_1 + nt'_2\). Since \(\alpha(0) = 0 \leq t_1 + t_2 < \alpha(T')\), the continuity of \(\alpha\) implies that there exists \(T\) such that

\[
0 \leq T \leq T' \quad \text{and} \quad \alpha(T) = t_1 + t_2.
\]

By pairwise consistency and symmetry

\[
F(x_1, x_2, x_3, \ldots, x_2; T) = (t_1, t_2, t_2, \ldots, t_2).
\]

Therefore \(T = t_1 + nt_2\). By choice of \(n\), \(T = t_1 + nt_2 > t'_1 + nt'_2 = T'\). But by (8), \(T \leq T'\), which is a contradiction. \(\blacksquare\)

Suppose that \(F\) were strictly monotone increasing in \(T\). Then the proof of Theorem 1 would be quite simple. Consider the solutions for the two-agent problem \((x, 1; T)\) as \(T\) varies from 0 to \(x + 1\). For every \(\lambda \in [0, 1]\) define \(t = f(x, \lambda)\) if and only if \((t, \lambda) = F(x, 1; t + \lambda)\). The continuity and strict monotonicity of \(F\) imply that \(f\) is well-defined, continuous, and strictly monotonic in \(\lambda\). We claim that \(f\) is a parametric representation of \(F\). Fix \(t^* = F(x; T^*)\) where \(x = (x_1, x_2, \ldots, x_n) > 0\) and \(0 \leq T^* \leq \sum x_i\). Consider the solutions \((t, \lambda) = F(x, 1; T)\) as \(T\) varies from 0 to 1 + \(\sum x_i\). By continuity there exists a value of \(T\), say \(T'\), such that at \(T'\) agents 1 and 2 get a combined total of \(t^*_1 + t^*_2\). By pairwise consistency they must therefore divide it as \((t^*_1, t^*_2)\). Likewise there is a value \(T = T''\) such that agents 1 and 3 get a combined total of \(t^*_3 + t^*_5\). By pairwise consistency they must divide it as \((t^*_3, t^*_5)\). Since \(F\) is strictly monotonic and agent 1 gets \(t^*_1\) at both \(T'\) and \(T''\), it follows that \(T' = T''\). A continuation of this line of argument shows that there is a value \(T = T^*\) and a value \(\lambda = \lambda^*\) such that \(F(x_1, x_2, \ldots, x_n, 1; T^*) = (t^*_1, t^*_2, \ldots, t^*_n, \lambda^*)\). Pairwise consistency implies then that \((t^*_i, \lambda^*) = F((x_i, 1); t^*_i + \lambda^*)\), so by definition of \(f\), \(t^*_i = f(x_i, \lambda^*)\) for all \(i\).

Conversely, suppose that \(f(x_i, \lambda) = t_i\) for some \(\lambda\) and all \(i\). Let \(T = \sum t_i\). The above argument implies that there exists a \(\lambda'\) such that \(f(x_i, \lambda') = t_i\) and \(\sum t'_i = \sum t_i = T\). Since \(f\) is monotonic in \(\lambda\), it follows that \(t'_i = t_i\) for all \(i\). Therefore, \(f\) is a parametric representation of \(F\).

The preceding argument is quite general, and will be appealed to several times below. Namely, to show that \(f(x, \lambda)\) parametrically represents some method \(F\), it is enough to know that \(f\) is monotonic in \(\lambda\) (not necessarily strictly so), and that \(T = F(x; T)\) implies that \(t_i = f(x_i, \lambda)\) for some \(\lambda\) and all \(i\).

Unfortunately, many interesting allocation methods are not strictly monotonic. For example, neither the head tax nor the levelling tax is strictly monotonic, because of the boundary constraints. Nor is the Talmudic method that will make its appearance in \S 6. To handle these cases a variant of the above method of proof is required.
Let $F$ satisfy the hypotheses of Theorem 1. By Lemma 1, $F$ is monotone increasing in $T$. For every $x, y > 0$ and $0 \leq t \leq x$, define

$$g(y, x, t) = \max\{F_1(y, x; T): F_2(y, x; T) = t, 0 \leq T \leq y + x\}. \quad (9)$$

Thus $g(y, x, t)$ is the maximum amount that player 1 pays (his ability to pay being $y$), when paired with player 2 who pays $t$ (and whose ability to pay is $x$). The function $g(y, x, t)$ is well-defined and $0 \leq g(y, x, t) \leq y$ for all $y > 0$.

**Lemma 2.** For each fixed $x > 0$ and $0 \leq t \leq x$, $g(y, x, t)$ is continuous in $y$.

**Proof.** Fix $x > 0$ and $0 \leq t \leq x$, and let $\langle y_k \rangle \rightarrow y$, where $y_k > 0$ for all $k$. Fix $k$. For the three-agent problem $(y_k, y, x; T)$ there exists a maximum value of $T$ such that $F_3(y_k, y, x; T) = t$. Call this value $T^*$, and let

$$F(y_k, y, x; T^*) = (u_k^*, u^*, t). \quad (10)$$

By pairwise consistency and symmetry,

$$(u^*, t) = F(y, x; u^* + t), \quad (u_k^*, t) = F(y_k, x; u_k^* + t),$$

Let $g(y, x, t) = u$ and $g(y_k, x, t) = u_k$. By the definition of $g$, $u^* \leq u$ and $u_k^* \leq u_k$. Suppose that $u^* < u$. Since $F$ is continuous and monotone increasing in $T$, there is some $T > T^*$ such that

$$F_2(y_k, y, x; T) + F_3(y_k, y, x; T) = u + t > u^* + t. \quad (11)$$

By the definition of $g$, $(u, t) = F(y, x; u + t)$. By (11) and pairwise consistency it follows that $F_2(y_k, y, x; T) = u$ and $F_3(y_k, y, x; T) = t$. But then $T^*$ was not maximal, a contradiction. Hence $u^* = u$. By a similar argument deduce that $u_k^* = u_k$. Therefore (10) implies that

$$(u_k, u, t) = F(y_k, y, x; u_k + u + t). \quad (12)$$

This holds for all $k$. Let $u'$ be any limit point of the sequence $\langle u_k \rangle$. By (12) and the continuity of $F$, $(u', u, t) = F(y, x, u; u' + u + t)$. By symmetry, $u' = u$. Thus $\lim_{k \rightarrow \infty} g(y_k, x, t) = u = g(y, x, t)$, so $g$ is continuous in $y$. $\blacksquare$

Let $\phi(y)$ be some continuous probability density function on $\mathbb{R}_+$ that is positive for all $y > 0$ and has finite mean. We may think of $\phi$ as representing the distribution of incomes in some hypothetical society. For specificity we shall take the exponential distribution $\phi(y) = e^{-y}$, which has mean 1, although any convenient distribution could be used in the following construction. For all $x > 0$ and $0 \leq t \leq x$, define

$$h(x, t) = \int_0^\infty g(y, x, t) e^{-y} dy. \quad (13)$$

$h(x, t)$ is the expected amount that a randomly selected person would pay when paired with an individual who pays $t$ and whose ability to pay is $x$. Since $0 \leq g(y, x, t) \leq y$ for all $y$,

$$0 \leq h(x, t) \leq \int_0^\infty ye^{-y} dy = 1. \quad (14)$$

**Lemma 3.** For each $x > 0$, $h(x, t)$ is strictly monotone increasing in $t$. 

Proof. Fix \( x > 0 \) and suppose that \( x \geq t' > t \geq 0 \). Since \( g \) is monotone increasing in \( t \),

\[
h(x, t') = \int_0^\infty g(y, x, t') e^{-y} \, dy \geq \int_0^\infty g(y, x, t) e^{-y} \, dy = h(x, t).
\]

We wish to conclude that \( h(x, t') > h(x, t) \). Suppose not. Then \( g(y, x, t') = g(y, x, t) \) for almost all positive \( y \). In particular, every neighborhood of \( x \) contains such a \( y \). Therefore, we can choose a positive sequence \( \langle y_k \rangle \) converging to \( x \) such that \( g(y_k, x, t) = g(y_k, x, t') \) for all \( k \). Let \( u_k = g(y_k, x, t) = g(y_k, x, t') \). By definition of \( g \),

\[
(u_k, t') = F(y_k, x; u_k + t') \quad \text{and} \quad (u_k, t) = F(y_k, x; u_k + t) \quad \text{for all } k.
\]

Let \( u' \) be any limit point of \( \langle u_k \rangle \). From the preceding and the continuity of \( F \) conclude that

\[
(u', t') = F(x, x; u' + t') \quad \text{and} \quad (u', t) = F(x, x; u' + t).
\]

By symmetry, \( u' = t' \) and also \( u' = t \), so \( t = t' \), contrary to assumption. ■

Proof of Theorem 1. For every \( x > 0 \) and \( 0 \leq t \leq x \), define

\[
\begin{align*}
\hbar^-(x, t) &= \lim_{t' \to t^-} h(x, t'), \quad \hbar^+(x, t) = \lim_{t' \to t^+} h(x, t'), \\
h^-(x, 0) &= 0, \quad h^+(x, x) = 1.
\end{align*}
\]

(15)

For every \( \lambda \in [0, 1] \) and every \( x > 0 \), define

\[
t = f(x; \lambda) \quad \text{iff} \quad \hbar^-(x, t) \leq \lambda \leq \hbar^+(x, t).
\]

(16)

For every \( \lambda \in [0, 1] \) and every \( x > 0 \), there is at least one \( t \) satisfying (16). There is exactly one such \( t \) because whenever \( t < t' \),

\[
\hbar^-(x, t) \leq \hbar^+(x, t) < \hbar^-(x, t') \leq \hbar^+(x, t').
\]

Since \( h(x, t) \) is strictly increasing in \( t \), it follows that \( f(x, \lambda) \) is continuous and (weakly) monotone increasing in \( \lambda \) for each \( x > 0 \).

We claim that \( f \) represents the method \( F \). As we remarked before the proof of Lemma 2, it is enough to show that if \( t^* = F(x, T^*) \), then \( t_i^* = f(x_i, \lambda^*) \) for some \( \lambda^* \) and all \( i \). Let \( t^* = F(x, T^*) \). Given \( y > 0 \), consider the solutions \( (u, t_1, t_2, \ldots, t_n) = F(y, x_1, x_2, \ldots, x_n; T) \) as \( T \) varies from \( 0 \) to \( y + \sum x_i \). By the continuity of \( F \) there exists, for every \( i \) and \( j \), at least one value \( T = T_{ij} \) such that agents \( i \) and \( j \) get a combined total of \( t_i^* + t_j^* \). By pairwise consistency they must split it as \( (t_i^*, t_j^*) \). For each of the chosen values \( T_{ij} \), agent \( i \) gets the same amount, namely \( t_i^* \). Because \( F \) is monotone in \( T \), agent \( i \) gets \( t_i^* \) in the closed real interval \( \mathcal{S}_i \) spanned by the values \( \{T_{ij}; 1 < j < n\} \). The set of \( n \) intervals \( \mathcal{S}_i \) overlap pairwise, since \( \mathcal{S}_j \) and \( \mathcal{S}_k \) contain the common element \( T_{jk} \). Therefore all of the intervals contain a common element \( T \). The solution at \( T \) has the form \( (\bar{u}, t_1^*, t_2^*, \ldots, t_n^*) = F(y, x_1, x_2, \ldots, x_n; T) \). By pairwise consistency it follows that

\[
(\bar{u}, t_i^*) = F(y, x_i; \bar{u} + t_i^*) \quad \text{for all } i.
\]

(17)
By definition of $g$, $\bar{u} \leq g(y, x_i, t_i^*)$ for all $i$. Let $u^*$ be the maximum value of $\bar{u}$ for which (17) holds. Then $u^* = \min_i \{ g(y, x_i, t_i^*) \}$, which is a continuous function of $y$. Define $\gamma(y) = \min_i \{ g(y, x_i, t_i^*) \}.$

Fix $i$, $y > 0$, and $t_i < t_i^*$. Let $u = g(y, x_i, t_i)$ and $u^* = \gamma(y)$. By definition of $g$, $(u, t_i) = F(y, x_i; u + t_i)$. Since $F$ is monotone increasing, the fact that $t_i < t_i^*$ and $(u^*, t_i^*) = F(y, x_i; u^* + t_i^*)$ implies that $u + t_i < u^* + t_i^*$ and $u \leq u^*$. Thus for all $i$ and all $y > 0$,

$$g(y, x_i, t_i) \leq \gamma(y) \leq g(y, x_i, t_i^*) \quad \text{whenever } t_i < t_i^*.$$ 

Letting $\lambda = \int_0^\infty \gamma(y) e^{-y} \, dy$ it follows that

$$h(x_i, t_i) \leq \lambda \leq h(x_i, t_i^*) \quad \text{whenever } t_i < t_i^*.$$ 

Hence

$$h^-(x_i, t_i^*) \leq \lambda \leq h^+(x_i, t_i^*) \quad \text{for all } i,$$

so

$$t_i^* = f(x_i, \lambda) \quad \text{for all } i.$$ 

Q.E.D.

It remains only to show that $f$ is a jointly continuous function of $x$ and $\lambda$. Let $\langle(x_k, \lambda_k)\rangle \to (x, \lambda)$. Further, let $t' = f(x_k, \lambda_k)$ and $t'' = f(x, \lambda_k)$. Then

$$(t', t'') = F(x_k, x; t' + t'').$$

We already know that $f$ is continuous in $\lambda$ for each fixed $x$, hence $\langle t'' \rangle \to t = f(x, \lambda)$. Let $t'$ be any limit point of $\langle t'_k \rangle$. By the continuity of $F$, $(t', t) = F(x, x; t' + t)$. Symmetry implies that $t' = t$. Therefore, $f(x_k, \lambda_k)$ converges to $f(x, \lambda)$, so $f$ is continuous.

5. **Consistency and optimality.** A common approach to allocating a cost or tax is to minimize some collective measure of “loss”, or to maximize some collective measure of “welfare.” In the simplest case, the same measure is applied to all individuals and the results are added up. The difficulty with this approach lies in choosing the appropriate measure. The advantages are its intuitive appeal and analytical convenience. In this section we shall show that every symmetric, continuous, and consistent method minimizes an additively separable, strictly convex loss function. Equivalently, the method maximizes an additively separable, strictly concave welfare function. We shall show explicitly how to construct such a function based on a parametric representation of the method.

Consider first some simple examples of objective functions. Let $t$ be an allocation of taxes associated with some vector of taxable incomes $x$. A plausible measure of loss is the total “sacrifice” suffered by all individuals. Let $u(x)$ be a differentiable, strictly concave function that represents the ‘utility’ of income level $x > 0$. The same function is assumed to apply to all individuals. For example, take $u(x) = \ln x$. The total sacrifice imposed by $t$ is $S(x, t) = \Sigma [\ln x_i - \ln(x_i - t_i)]$. Minimizing $S(x, t)$ subject to $\Sigma t_i = T$ is a standard exercise in constrained optimization. Form the Lagrangian

$$L(t) = \Sigma [\ln x_i - \ln(x_i - t_i)] + \lambda (T - \Sigma t_i).$$

Setting $\partial L/\partial t_i = 0$ yields the solution $1/(x_i - t_i) = \lambda$ for all $i$. That is, $t_i = x_i - 1/\lambda$, which is the levelling tax. If the constraint $0 \leq t \leq x$ is also imposed, then the solution is $t_i = \max\{0, x_i - 1/\lambda\}$. 

In fact, the levelling method minimizes total sacrifice for any increasing, differentiable, strictly concave utility function \( u(x) \). This is because every optimal solution \( t \) satisfies \( u'(x_i - t_i) = \lambda \) for some \( \lambda \) and all \( i \). Since \( u \) is strictly concave, its derivative \( u' \) is strictly decreasing, hence invertible, so \( t_i = x_i - [u']^{-1}(\lambda) \). Taking account of the constraint \( 0 \leq t_i \leq x_i \), it follows that the minimum occurs when \( t_i = \max\{0, x_i - \lambda'\} \) for some \( \lambda' \geq 0 \).

Other types of loss functions are also of interest. For example, minimizing the simple sum of squares \( \sum t_i^2 \) subject to \( \sum t_i = T \) and \( 0 \leq t_i \leq x_i \) yields the head tax: \( t_i = \min\{x_i, \lambda/2\} \). Minimizing the weighted sum of squares \( \sum t_i^2/x_i \) yields the flat tax: \( t_i = (\lambda/2)x_i \).

In general, an objective function is symmetric and additively separable if it takes the form \( \sum H(x, t_i) \) for some function \( H(x, t) \) of scalar \( x > 0 \) and scalar \( t, 0 \leq t \leq x \). If, in addition, \( H(x, t) \) is a strictly convex function of \( t \) for each \( x \), then for every \( x > 0 \) and \( 0 \leq T \leq \Sigma x_i \), the following optimization problem has a unique solution \( t^* \):

\[
\min \sum H(x, t_i) \text{ subject to } \sum t_i = T \text{ and } 0 \leq t_i \leq x_i \text{ for all } i. \tag{18}
\]

We now show that every symmetric, continuous, and pairwise consistent allocation method can be cast in this form.

**Theorem 2.** A symmetric, continuous allocation method is pairwise consistent if and only if its solutions minimize a symmetric, continuous, additively separable, strictly convex objective function.

We shall show how to explicitly construct such a function from any method with the given properties. First we illustrate the idea of the proof in the special case when \( F \) is symmetric, continuous, pairwise consistent and strictly monotonic in \( T \).

For all \( \lambda \in [0, 1] \) define the function

\[
t = f(x, \lambda) \iff (\lambda, t) = F(1, x; \lambda + t).
\]

As shown before the proof of Lemma 2, \( f \) represents \( F \). It is also continuous and strictly monotonic in \( \lambda \) for each \( x \). Define \( f^{-1}(x, t) = \lambda \) if and only if \( f(x, \lambda) = t \). \( f^{-1} \) is continuous and strictly monotone increasing in \( t \) for each \( x \). Therefore,

\[
H(x, t) = \int_0^t f^{-1}(x, t') \, dt'
\]

is continuous, strictly convex in \( t \), and has continuous first partial derivative \( \partial H(x, t)/\partial t = f^{-1}(x, t) \). (See [12, Theorem 24.2].) We claim that, with \( H \) defined in this way, \( F(x, T) \) is an optimal solution to (18) for any feasible problem \( (x, T) \). Given a feasible problem \( (x, T) \), form the Lagrangian

\[
L(t) = \sum H(x, t_i) + \lambda(T - \sum t_i).
\]

Suppose now that \( t^* \) minimizes \( \Sigma H(x, t_i) \) subject to \( \Sigma t_i = T \) and \( 0 \leq t_i \leq x_i \) for all \( i \). Suppose further that the latter constraints are not binding at the optimum, so that \( 0 < t^*_i < x_i \) for all \( i \). Then \( \partial L(t^*)/\partial t_i = 0 \) for all \( i \), which means that \( \partial H(x, t^*_i)/\partial t_i = \lambda \) for all \( i \). Since \( \partial H(x, t^*_i)/\partial t_i = f^{-1}(x_i, t^*_i) \), it follows that \( f^{-1}(x_i, t^*_i) = \lambda \) and \( t^*_i = f(x_i, \lambda) \) for all \( i \). Because \( f \) represents \( F \), it follows that \( t^* = F(x^*, T^*) \). We have therefore shown that if \( t^* \) is an optimal interior solution to (18), then \( t^* \) is an \( F \)-allocation. Conversely, if \( t^* \) is an \( F \)-allocation and \( 0 < t^* < x \), then the reverse line of reasoning shows that \( t^* \) optimizes (18). Indeed, an extension of
this argument shows that $F$-allocations also optimize (18) even when some of the constraints $0 \leq t_i \leq x_i$ are binding. This establishes Theorem 2 when $F$ is strictly monotone increasing in $T$.

In case $F$ is not strictly monotonic in $T$, the following more detailed argument is required.

**Proof of Theorem 2.** If $F(x; T)$ minimizes a symmetric, continuous, additively separable and strictly convex objective function, then it is straightforward to check that $F$ is symmetric, continuous, and pairwise consistent. Conversely, suppose that $F$ is symmetric, continuous, and pairwise consistent. By Theorem 1, $F$ has a parametric representation $t = f(x, \lambda)$, where $\lambda \in [0, 1]$. Let $\lambda = h(x, t)$ be any pseudo-inverse of $f$: a strictly monotone increasing function of $t$ defined for all $x > 0$ and $0 \leq t \leq x$, such that

$$t = f(x, \lambda) \iff h^-(x, t) \leq \lambda \leq h^+(x, t),$$

where $h^-$ and $h^+$ are defined as in (15). The function $H(x, t) = \int_0^t h(x, t') dt'$ is a closed, convex function of $t$ for each $x$ (see [12, Theorem 24.2]). Moreover, since $h$ is strictly increasing in $t$, $H$ is strictly convex in $t$ for each $x$.

Fix $x_1, x_2, \ldots, x_n > 0$ and $0 \leq T \leq \sum x_i$. We shall show that $t^*$ minimizes $\sum H(x_i, t_i)$ subject to the constraints if and only if $t_i^* = f(x_i, \lambda)$ for some $\lambda$ and all $i$. Define $f_0(t) = \sum H(x_i, t_i)$ and $f_1(t) = T - \sum t_i$. Extend $f_0(t)$ to all of $\mathbb{R}^n$ by letting $H(x_i, t_i) = +\infty$ whenever $t_i < 0$ or $t_i > x_i$. Then the optimization problem

$$\min \sum H(x_i, t_i) \quad \text{subject to} \quad \sum x_i = T \text{ and } 0 \leq t_i \leq x_i$$

(20)

takes the form

$$\min f_0(t) \quad \text{subject to } f_1(t) = 0 \text{ and } t \in \mathbb{R}^n.$$  

Assume first that $0 < T < \sum x_i$. Then there exists $t$ in the relative interior of the domain where $f_0(t)$ is finite such that $f_1(t) = 0$. Hence [12, Corollary 28.2.2] there exists a Kuhn-Tucker coefficient for (21), that is, a coefficient $\lambda$ such that if $t^*$ minimizes (20), then $t^*$ minimizes $f_0(t) + \lambda f_1(t)$ over all $t$ in $\mathbb{R}^n$. Now

$$\min_{t \in \mathbb{R}^n} \{ f_0(t) + \lambda f_1(t) \} = \lambda T + \sum_{0 \leq t_i \leq x_i} \min \{ H(x_i, t_i) - \lambda t_i \}. \quad (22)$$

Hence if $t^*$ minimizes (20), then for all $i$, $\lambda$ is a subgradient of $H$ at $(x_i, t_i^*)$, that is,

$$H(x_i, t_i^*) - \lambda t_i \geq H(x_i, t_i^*) - \lambda t_i^*.$$ 

By [12, Theorem 24.2] and the construction of $H$, the subgradient set of $H$ at $(x_i, t_i^*)$ is the interval

$$\partial H(x_i, t_i^*) = [h^-(x_i, t_i^*), h^+(x_i, t_i^*)].$$

Therefore $h^-(x_i, t_i^*) \leq \lambda \leq h^+(x_i, t_i^*)$ for all $i$, so by the choice of $h$ it follows that

$$t_i^* = f(x_i, \lambda) \text{ for all } i.$$

If $T = 0$, the unique optimum of (20) is $t^* = 0$, and by definition of $f$, $t_i^* = f(x_i, 0) = 0$ for all $i$. If $T = \sum x_i$, the unique optimum of (20) is $t^* = x$ and by definition of $f$, $t_i^* = f(x_i, 1) = x_i$ for all $i$. Thus in all cases, if $t^*$ optimizes (20), then $t_i^* = f(x_i, \lambda)$ for some $\lambda$ and all $i$. Hence $t^* = F(x; T)$, because $f$ represents $F$.

Conversely, suppose that $t' = F(x; T)$, that is, $t_i' = f(x_i, \lambda')$ for some $\lambda'$ and all $i$. Let $t^*$ optimize (20) for this $x$ and $T$. By the above, $t_i^* = f(x_i, \lambda)$ for some $\lambda$ and all $i$. Since $\sum t_i^* = \sum t_i' = T$ and $f$ is monotone in $\lambda$, it follows that $t^* = t$. 


It remains only to check that \( H(x, t) \) is continuous. This would be immediate if \( h \) were continuous, but \( h \) need not be. Observe first that for every \( x > 0 \) and \( 0 \leq t', t'' \leq x \),

\[
|H(x, t') - H(x, t'')| \leq |t' - t''|,
\]

because \( 0 \leq h(x, t) \leq 1 \) for all \( 0 \leq t \leq x \).

Let \( \langle (x_k, t_k) \rangle \to (x^*, t^*) \), where \( 0 \leq t^* \leq x^* \) and \( 0 \leq t_k \leq x_k \) for all \( k \). We wish to show that \( \langle H(x_k, t_k) \rangle \to H(x^*, t^*) \). If \( t^* < x^* \), then for all sufficiently large \( k \), \( t_k \leq x^* \), so \( H(x_k, t^*) \) is defined and

\[
|H(x_k, t_k) - H(x^*, t^*)| \leq |H(x_k, t_k) - H(x_k, t^*)| + |H(x_k, t^*) - H(x^*, t^*)| \\
\leq |t_k - t^*| + |H(x_k, t^*) - H(x^*, t^*)|.
\]

On the other hand, if \( t^* = x^* > 0 \), then for every small \( \epsilon > 0 \) and all sufficiently large \( k \), \( H(x_k, t^* - \epsilon) \) is defined and

\[
|H(x_k, t_k) - H(x^*, t^*)| \leq |H(x_k, t_k) - H(x_k, t^* - \epsilon)| \\
+ |H(x_k, t^* - \epsilon) - H(x^*, t^* - \epsilon)| \\
+ |H(x^*, t^* - \epsilon) - H(x^*, t^*)| \\
\leq |t_k - t^*| + \epsilon + |H(x_k, t^* - \epsilon) - H(x^*, t^* - \epsilon)| + \epsilon.
\]

In either case, therefore, to show the continuity of \( H \), it suffices to show that \( \langle H(x_k, t^*) \rangle \to H(x^*, t^*) \) whenever \( \langle x_k \rangle \to x^* \) and \( t^* < x^* \). By definition of \( h \), \( f(x_k, h(x_k, t)) = t \) for all \( k \). For any limit point \( \lambda \) of the sequence \( \langle h(x_k, t) \rangle \), the continuity of \( f \) implies that \( f(x^*, \lambda) = t \). By definition of \( h \), \( h^- (x^*, t) \leq \lambda \leq h^+ (x^*, t) \). Therefore, if \( h \) is continuous at \( (x^*, t) \) (that is, if \( h^- (x^*, t) = h^+ (x^*, t) \)), then \( \langle h(x_k, t) \rangle \to h(x^*, t) = \lambda \). For fixed \( x^* > 0 \), \( h(x^*, t) \) is monotone increasing in \( t \). Hence the set of points \( t \) such that \( h \) is discontinuous at \( (x^*, t) \) is countable. Thus \( \langle h(x_k, t) \rangle \to h(x^*, t) \) for almost all \( t \). Further, the integral of \( h(x_k, t) \) is uniformly bounded for all \( k \):

\[
0 \leq \int_0^t h(x_k, t) \, dt \leq \int_0^* dt \leq t^*.
\]

Therefore the Lebesgue convergence theorem implies that

\[
H(x_k, t^*) = \int_0^t h(x_k, t) \, dt \to \int_0^* h(x^*, t) \, dt = H(x^*, t^*),
\]

so \( H \) is continuous.

There is an interesting connection between Theorem 2 and a recent result of T. Lensberg on the bargaining problem [8], [9], [10]. For every finite nonempty set \( S \) of positive integers, a bargaining problem for \( S \) consists of a compact, convex subset \( C \) of \( \mathbb{R}^S_{+} \) that specifies all possible vectors \( u \) of utilities ("payoffs") that the agents can achieve by cooperating. It is assumed that \( C \) contains some strictly interior point of \( \mathbb{R}^S_{+} \), and that for every \( u \in C \) and \( 0 \leq u' \leq u, u' \in C \). The players 'bargain' over which vector in \( C \) will be the outcome. A bargaining method is a single-valued function \( F(C) \subseteq C \) defined for all bargaining problems.
A taxation problem can be framed as a bargaining problem by viewing the distribution of after-tax income as the bone of contention. If \( x \in \mathbb{R}^n_+ \) is the distribution of taxable incomes and \( T \) the tax total, then the total amount of after-tax income is

\[ U = \sum x_i - T. \]

The bargaining set in this case may be defined as \( C = \{ (u) \in \mathbb{R}^n_+ : 0 \leq u \leq x \text{ and } \sum u_i \leq U \} \). Symmetry, continuity, and consistency all have natural interpretations in the bargaining framework (for detailed definitions, see [9], [10]). Another reasonable property is Pareto optimality: if \( u \geq F(C) \) and \( u \in C \), then \( u = F(C) \). In the case of taxation, this corresponds to the requirement that \( \sum t_i = T \).

In [9], Lensberg showed that a bargaining method that is continuous, pairwise consistent, and Pareto optimal must maximize a separable objective function of the form \( \sum \eta_i(u_i) \), subject to \( u \in C \), where each of the extended real-valued functions \( \eta_i \) is continuous and strictly monotone increasing, and every finite partial sum \( \sum \eta_i(u_i) \) is strictly quasi-concave.

Lensberg’s result does not imply Theorem 2 however. First, it is derived from the class of all bargaining problems, instead of the special class of linear bargaining problems treated here. Second, in the present framework, individuals are differentiated by a continuous parameter \( x \) that measures their ability to pay and constrains the allocation that they may receive. The value of the objective function \( \sum H(x_i, t_i) \) depends on both the value \( x_i \) and the amount \( t_i \) that \( i \) is allocated, but not on \( i \) itself. In Lensberg’s model, by contrast, the individuals have different objective functions that may or may not depend on their ability to pay.

6. An example from the Babylonian Talmud. We illustrate the application of Theorems 1 and 2 with a two-thousand year old claims problem from the Babylonian Talmud (Kethubot 93a) that has been beautifully analyzed by Aumann and Maschler [1].

A man had three wives, each of whom has a claim on the deceased’s estate. The Talmud gives solutions for three possible cases.

<table>
<thead>
<tr>
<th>Size of Estate</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
</tr>
<tr>
<td>Amount of Claim</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>200</td>
</tr>
<tr>
<td>300</td>
</tr>
</tbody>
</table>

In general, let \( x \) denote a set of \( n \) claims on an estate having total value \( T \). A bankruptcy problem is a pair \( (x; T) \) such that \( 0 \leq T \leq \sum x_i \). A solution is a vector \( t \in \mathbb{R}^n \) such that \( \sum t_i = T \) and \( 0 \leq t_i \leq x_i \) for all \( i \). It is clear that methods for solving bankruptcy problems are formally equivalent to methods of taxation.

Given a bankruptcy problem \( (x; T) \), where \( x \in \mathbb{R}^n_+ \), define the following cooperative game: for each set \( J \subseteq I \) of claimants, let \( v(J) \) denote the amount left over (if any) after all of the claims of those not in \( J \) have been met in full. Thus \( v(J) = \max\{0, T - \sum_{i \notin J} x_i\} \). Aumann and Maschler show that each of the Mishna solutions is the nucleolus of this cooperative game. They show further that the nucleolus can be calculated as follows. Think of \( T \) as increasing incrementally from zero:

- (a) For \( T = 0 \), let \( t_i = 0 \) for all \( i \);
- (b) For \( 0 < T < \sum x_i/2 \) give equal increments to all \( i \) who currently have \( t_i < x_i/2 \);
- (c) For \( T = \sum x_i/2 \), let \( t_i = x_i/2 \) for all \( i \);
- (d) For \( \sum x_i/2 < T \leq \sum x_i \), give equal increments to all \( i \) for whom \( x_i - t_i \) is a maximum.
From this description of the Aumann-Maschler method it is evidently consistent, symmetric, and continuous. It is not strictly monotonic, however, because when some claimant gets half of her claim, she ceases to get any more until all of the other claimants have gotten at least half of theirs.

The proof of Theorem 1 can be applied to construct a parametric representation of the Aumann-Maschler method, and the proof of Theorem 2 to construct an objective function that the method optimizes. The first step is to define the function $g(y, x, t)$. As in (9), $g(y, x, t)$ is the maximum amount that an individual with a claim of $y$ can get when an individual with a claim of $x$ gets the amount $t$. The function $h(x, t)$ is found by integrating $\int_{0}^{x} g(y, x, t) e^{-y} dy$. There are three cases:

<table>
<thead>
<tr>
<th>Case</th>
<th>$g(y, x, t)$</th>
<th>$h(x, t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t &lt; x/2$</td>
<td>$\min(t, y/2)$</td>
<td>$(1 - e^{-2t})/2$</td>
</tr>
<tr>
<td>$t = x/2$</td>
<td>$\max(y - x/2, y/2)$</td>
<td>$(1 + e^{-x})/2$</td>
</tr>
<tr>
<td>$t &gt; x/2$</td>
<td>$\max(y - x + t, y/2)$</td>
<td>$(1 + e^{-2(x-t)})/2$</td>
</tr>
</tbody>
</table>

Invert $\lambda = h(x, t)$ to obtain the parametric representation

$$t = f(x, \lambda) = -(1/2)\ln(1 - 2\lambda), \quad 0 \leq \lambda \leq (1 - e^{-x})/2,$$

$$t = f(x, \lambda) = x/2, \quad (1 - e^{-x})/2 \leq \lambda \leq (1 + e^{-x})/2,$$

$$t = f(x, \lambda) = x + (1/2)\ln(2\lambda - 1), \quad (1 + e^{-x})/2 \leq \lambda \leq 1.$$

The above representation can be simplified by making a monotone increasing, continuous transformation from $\lambda$ to a new parameter $\mu$ as follows:

$$\mu = 2/\ln(1 - 2\lambda), \quad 0 \leq \lambda \leq (1 - e^{-x})/2,$$

$$\mu = (4\lambda - 2)e^{x}/x, \quad (1 - e^{-x})/2 \leq \lambda \leq (1 + e^{-x})/2,$$

$$\mu = -2/\ln(2\lambda - 1), \quad (1 + e^{-x})/2 \leq \lambda \leq 1.$$

This yields the equivalent parametric representation

$$t = f^{*}(x, \mu) = \begin{cases}  
-1/\mu, & -\infty \leq \mu \leq -2/x, \\
(x/2, & -2/x \leq \mu \leq 2/x, \\
x - 1/\mu, & 2/x \leq \mu \leq \infty. 
\end{cases} \quad (24)$$

To construct an objective function, consider the following pseudo-inverse of $f^{*}$:

$$\mu = h^{*}(x, t) = -1/t, \quad 0 \leq t \leq x/2,$$

$$\mu = h^{*}(x, t) = 1/(x - t), \quad x/2 \leq t \leq x.$$

Integrate $h^{*}(x, t)$ to obtain $H(x, t)$:

$$H(x, t) = -\ln t \quad 0 \leq t \leq x/2,$$

$$H(x, t) = -\ln(x - t) \quad x/2 \leq t \leq x.$$

The functions $f(x, \mu)$ and $H(x, t)$ are illustrated for $x = 2$ in Figures 1 and 2.
Figure 1. $f(2, \mu)$.

Figure 2. $H(2, t)$.

Figure 3. $f(x, \mu)$, $\mu < 0$. 
If the Talmudic method is applied to the construction of tax schedules, the results are rather peculiar. Fix some value of the parameter \( \mu \) in (24) and consider the resulting function of \( x \). If \( \mu < 0 \), then taxation is at the 50% rate for all incomes in the range \( 0 < x \leq 2/|\mu| \), and taxation is constant for all incomes above \( 2/|\mu| \) (see Figure 3). If \( \mu = 0 \) taxation is at the flat rate of 50% for all incomes. If \( \mu > 0 \), then taxation is at the 50% rate for all incomes in the range \( 0 < x \leq 2/\mu \), and after-tax incomes are constant for all incomes above \( 2/\mu \) (see Figure 4).

7. Conclusion: profit-sharing problems. Suppose that \( n \) partners contribute amounts \( y_1, y_2, \ldots, y_n \) to a joint enterprise. Let the total revenue from the enterprise be \( R \). Assuming that \( R \geq \sum y_i \), how should it be shared among the partners? A profit-sharing problem is a pair \((y; R)\) where \( y > 0 \) and \( \sum y_i \leq R < \infty \). A solution of \((y; R)\) is a vector \( r \) such that \( \sum r_i = R \) and \( r_i \geq y_i \) for all \( i \). A profit-sharing method is a function \( F^* \) that gives a unique solution \( r = F^*(y; R) \) for every profit-sharing problem \((y; R)\).

The profit-sharing problem is very similar to the bankruptcy problem, and the results of the preceding sections are easily adapted to it. The definitions of symmetry, continuity, consistency, and monotonicity carry over naturally, and are left to the reader. A parametric profit-sharing function is a real-valued function \( F^*(y, \mu) \) defined for all \( y > 0 \) and all \( \mu \) in some half-open interval \([a, b)\), such that for each fixed \( y \), \( F^* \) is continuous and weakly monotone increasing in \( \mu \), and \( F^*(y, a) = y, \lim_{\mu \to \infty} F^*(y, \mu) = \infty \). A profit-sharing method \( F^* \) is parametric if for some such \( F^* \) and all \((y; R)\),

\[
r = F^*(y; R) \quad \text{iff} \exists \mu \forall i \quad r_i = F^*(y_i, \mu) \quad \text{and} \quad \sum r_i = R.
\]

**Theorem 3.** A continuous profit-sharing method is symmetric and pairwise consistent if and only if it is representable by a continuous parametric function.

**Corollary.** If a profit-sharing method is symmetric, continuous, and pairwise consistent, then it is monotone increasing in \( R \).

The proof of this corollary parallels the proof of Lemma 1 and is left to the reader. Theorem 3 can then be derived from Theorem 1 by a process of inversion. Let \( F^* \) be a profit-sharing method satisfying the three properties, and hence monotonic. Define a taxation method \( F \) as follows. Fix \( k > 0 \). For every \( T \) such that \( 0 < T < \sum x_i \), let

\[
t = F(x; T) \quad \text{if there exists} \quad R \quad \text{such that}
\]

\[
(1/t_1, \ldots, 1/t_n) = F^*(1/x_1, \ldots, 1/x_n; R) \quad \text{and} \quad T = \sum t_i.
\]

(25)
We claim that $F$ is single-valued. Indeed, suppose that $t'$ and $t''$ are two solutions for $F(x; T)$. Then for some $R'$ and $R''$, $(1/t_1', \ldots, 1/t_n') = F*(1/x_1, \ldots, 1/x_n; R')$ and $(1/t_1'', \ldots, 1/t_n'') = F*(1/x_1, \ldots, 1/x_n; R'')$. Since $F*$ is monotonic, either $1/t_i' < 1/t_i''$ for all $i$, or vice versa. From $\sum t'_i = \sum t''_i$, conclude in either case that $t' = t''$.

Expression (25) defines $F$ continuously, symmetrically, monotonically, and consistently. But not necessarily completely, since for some $i$ it may happen that $\lim_{R \to \infty} F^*(1/x_1, \ldots, 1/x_n; R) = c_i < \infty$. The set $J$ of all such indices $i$ is properly contained in the set of all indices $\{1, 2, \ldots, n\}$. Therefore (25) defines $F(x; T)$ only for those $T$ satisfying $\sum 1/c_i < T \leq \sum x_i$. Extend $F$ to $T \leq \sum 1/c_i$ by letting

$$t = F(x; T) \text{ if } t_i = 0 \text{ for all } i \notin J, \sum t_i = T,$$

and there exists $R$ such that

$$[\{1/t_i\}_{i \in J} = F^*((1/x_i); J, R)].$$

Again, there may exist some set $J'$ of indices in $J$ such that for every $i \in J'$, $\lim_{R \to \infty} F^*((1/x_i); J, R) < \infty$. Extend $F$ in the above manner relative to the set $J'$. Since each successive subset of indices is properly contained in the preceding one, this process terminates and defines $F(x; T)$ for all $T, 0 < T \leq \sum x_i$. Finally, let $F(x, 0) = 0$. Then $F$ is symmetric, continuous, and consistent. By Theorem 1, $F$ has a parametric representation $f(x, \lambda)$, where $f$ may be chosen so that the domain of $\lambda$ is $[0, \infty]$. Define $f^*(y, \mu) = 1/f(1/y, 1/\mu)$ for all $\mu \in [0, \infty)$. We claim that $f^*$ represents $F^*$. As remarked earlier, it is enough to show that if $r = F^*(y; R)$ then $r_i = f^*(y, \mu)$ for some $\mu$ and all $i$. By definition of $F$, $r = F^*(y; R)$ implies that $(1/r_1, 1/r_2, \ldots, 1/r_n) = F(1/y_1, \ldots, 1/y_n; T)$ for $T = \sum 1/r_i$. Hence $1/r_i = f(1/y_i, \lambda)$ for some $\lambda$ and all $i$. Therefore $r_i = 1/f(1/y_i, \lambda) = f^*(y, \mu)$ for all $i$, where $\mu = 1/\lambda$. This proves Theorem 3.

Parametric profit-sharing methods may also be derived by optimizing an appropriate objective function. The proof parallels that of Theorem 2. Suppose that $f^*(y, \mu)$ parametrically represents the profit-sharing method $F^*$. The parameter $\mu$ may be chosen (by making an order-preserving, continuous transformation of the parameter) so that its domain of definition is the interval $[0, 1]$. For all $0 < y \leq r \leq \infty$, define $h(y, r) = \sup \{\mu : f^*(y, \mu) = r\}$. $h$ is a strictly monotone increasing function of $r$, and its range is $[0, 1]$. Let

$$h^-(y, r) = \lim_{r' \to r} h(y, r'), \quad h^+(y, r) = \lim_{r' \to r^+} h(y, r'),$$

$$h^-(y, 0) = 0, \quad h^+(y, \infty) = 1.$$

Then $h$ is a pseudo-inverse of $f^*$ in the sense that

$$f^*(y, \mu) = r \iff h^-(y, r) \leq \mu \leq h^+(y, r).$$

As in the proof of Theorem 2, the function $H(y, r) = \int_r^\infty h(y, r') \, dr'$ exists, is continuous, is strictly convex in $r$, and its subgradient set is

$$\partial H(y, r) = [h^-(y, r), h^+(y, r)].$$

An argument similar to the proof of Theorem 2 shows that $r = F^*(y; R)$ iff $r$ minimizes

$$\sum H(y_i, r_i) \text{ subject to } \sum r_i = R \text{ and } r_i \geq y_i \text{ for all } i.$$

**Theorem 4.** A continuous profit-sharing method is symmetric and pairwise consistent if and only if it minimizes a symmetric, continuous, additively separable and strictly convex objective function.
Acknowledgements. The author gratefully acknowledges helpful comments by R. Aumann, D. Foster, M. Maschler, and H. Moulin.

References

SCHOOL OF PUBLIC AFFAIRS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MARYLAND 20742