

NBER WORKING PAPER SERIES

RANDOMIZATION IN OPTIMAL INCOME TAX SCHEDULES

Dagobert L. Brito

Jonathan H. Hamilton

Steven M. Slutsky

Joseph E. Stiglitz

Working Paper No. 3289

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 1990

We thank David Sappington for valuable comments and James Mirrlees and Margaret Meyer for suggesting the example in Section IV in a different context. Joseph Stiglitz thanks the Hoover Institution, the John M. Olin Foundation, and the National Science Foundation for financial support. This paper is part of NBER's research program in Taxation. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

RANDOMIZATION IN OPTIMAL INCOME TAX SCHEDULES

ABSTRACT

The optimal income tax problem, since it requires self-selection constraints which define nonconvex feasible sets, is one of the many problems in economics for which randomization in the solution may be desirable. For a two-class economy, we characterize the optimal random tax schedules and we present necessary and sufficient conditions for the desirability of local randomization. The standard single-crossing restriction on preferences is not required for these results. We also show that randomization can be beneficial without violating (ex post as well as ex ante) horizontal equity. Lastly, we give an example to demonstrate that the gains from randomization may be large.

Dagobert L. Brito
Department of Economics
Rice University
Houston, TX 77251

Jonathan H. Hamilton
Department of Economics
University of Florida
Gainesville, FL 32611

Steven M. Slutsky
Department of Economics
University of Florida
Gainesville, FL 32611

Joseph E. Stiglitz
Department of Economics
Stanford University
Stanford, CA 94305

I. INTRODUCTION

Asymmetric information is an important factor in many areas of economics, including optimal income taxation, regulation, and labor contracts. The presence of asymmetric information adds self-selection or incentive compatibility constraints to an uninformed principal's maximization problem. The self-selection constraints produce feasible sets which are generally not convex. Given the nonconvexity, the uninformed may find it desirable to offer random schedules of opportunities to the informed. Weiss [1976], Stiglitz [1982], Fellingham, Kwon, and Newman [1984], Maskin and Riley [1984], and Arnott and Stiglitz [1988] have demonstrated benefits from randomization in a variety of contexts.

Several unresolved problems leave unclear the significance and relevance of these results. First, those papers utilize special assumptions on preferences, such as separability over commodities and "single-crossing" across types of individuals, without exploring how readily these can be relaxed. Second, only sufficient conditions for randomization to improve locally upon the deterministic solution have been found, and these sufficient conditions are not easy to interpret. It is therefore difficult to assess whether randomization is desirable only when preferences are in some way anomalous or whether it yields benefits in many circumstances. Third, it is unclear whether randomization is only a theoretical curiosity or whether it could be implemented in practice. Important practical considerations are the complexity and administrative feasibility of the randomization scheme. In addition, there are questions of ethical acceptability, in particular concerning horizontal equity--equal treatment of equals. While individuals of the same type are treated identically ex ante, randomization induces ex post utility differences within a class. This may be especially objectionable in

optimal taxation. It may also be an obstacle in principal-agent problems. Often, agents supply effort after resolution of the randomness. If they care about their relative status in a way that affects productivity, then inequitable random solutions may fail to yield gains over nonstochastic ones. Finally, previous work has only shown that deterministic solutions can be improved upon without deriving the optimal random solution. Thus, the magnitude of the gain from randomization is unknown. If the gains are small, randomization would not be worth the costs of its implementation.

In this paper, we treat each of these problems for an optimal income tax model with two classes of individuals and two commodities. The results directly translate to other two-class selection models. Unlike previous literature, we make no special assumptions about preferences other than concavity. We derive necessary and sufficient conditions for some local randomization to improve on the nonrandom solution. Essentially, these duplicate the sufficient conditions, such as those found by Stiglitz [1982], for a 50-50 randomization over two bundles. The set of preferences for which some randomization can be beneficial is the same as that for which a 50-50 randomization is desirable. Nevertheless, the generalization allows us to understand better the gains from randomizing. It also allows us to consider other questions such as whether randomization is desirable if the two classes have preferences which differ only slightly.

We also derive a sufficient condition for the desirability of nonlocal randomization. This condition and its proof add much insight into the value of randomization. Randomization can be desirable with any pair of ordinal preferences which differ at the nonrandom solution, provided that the utility function of the class whose self-selection constraint binds is sufficiently concave. The proof constructs a randomization over bundles which are

indifferent for the class for which they are intended. While this is not necessarily the optimal randomization, it shows that random solutions need not violate ex post horizontal equity. If horizontal equity were a political constraint on tax schedules, the gains from randomization might be reduced, but not eliminated.

Finally, we give an example in which we calculate the first-best, the nonrandom second-best, and the random second-best Pareto optima. Randomization has a major effect on the outcome. In this example, randomization allows the first-best Pareto frontier to be attained even when one group's self-selection constraint is strictly binding, despite the nonrandom second best being quite far from the unconstrained frontier! In our numerical example, at some points on the frontier, society would be willing to give up over 12% of its resources, rather than give up the ability to use random solutions.

Therefore, random solutions can be significant in selection models. They can arise for any set of ordinal preferences and the gains can be quite large. It may thus be a mistake to disregard the possibility of randomization in such models.

II. THE BASIC MODEL AND A CHARACTERIZATION THEOREM

A. Structural Assumptions

Consider a society composed of two different classes of individuals, denoted A and B. Individuals within each class are identical, but individuals in different classes differ either in tastes or abilities. The government does not know to which class any individual belongs, but does know the numbers of individuals in each class, denoted N_i , $i = a, b$. Individuals consume a single good, C, and earn income, Y. People in each class have a utility

function over these bundles $V^i(C, Y)$, $i = a, b$, with $\partial V^i / \partial C \equiv V_C^i > 0$ and $\partial V^i / \partial Y \equiv V_Y^i < 0$. The maximum income that individuals in each class can earn is bounded from above by K^i , so that $Y^i \leq K^i$, $i = a, b$. The marginal rate of substitution for an individual is denoted $MRS^i(C, Y) \equiv -V_Y^i / V_C^i > 0$, $i = a, b$.

The following assumptions about $V^i(C, Y)$ are made throughout:

- (A1) $V^i(C, Y)$, $i = a, b$, is twice continuously differentiable in C and Y ;
- (A2) $V^i(C, Y)$, $i = a, b$, is strictly concave in C and Y ;
- (A3) $MRS^a(C, Y)$ and $MRS^b(C, Y)$ differ at almost every (C, Y) bundle and those points where the MRS^i 's are equal do not lie on the no-tax budget line.

Assumption (A1) is made for convenience in exposition but can be relaxed without difficulty. Assumption (A2) of concavity instead of quasiconcavity insures that expected utility functions describe convex preferences. Assumption (A3) is crucial to guarantee that the groups actually have different preferences since, if their indifference curves coincide, redistribution between groups is impossible. Assumption (A3) allows multiple crossings of the indifference curves of the two classes. With multiple crossings, there will be bundles at which the indifference curves of the two groups are tangent. Such tangencies are not ruled out as long as they form only a discrete set of curves in the (C, Y) plane.¹

A special case satisfying these assumptions is that considered in Stiglitz [1982] in which the utility functions $V^i(C, Y)$ arise from common underlying preferences over consumption and labor with the classes A and B having different abilities (and wages). Let L^i be hours worked and w_i the wage rate of group i . Then $L^i = Y/w_i$ and $V^i(C, Y) = U(C, Y/w_i)$ where U is the common utility function over C and L . If A is the more able group ($w_a > w_b$) then $MRS^a(C, Y) < MRS^b(C, Y)$ at each (C, Y) .²

Because the government lacks information about abilities, it must offer a common set of opportunities to all individuals. With randomization, the opportunities may be lotteries over tax schedules. For any tax schedule, each type of individual will have a unique best consumption-income bundle on that schedule. For a tax schedule T , let $(C^i(T), Y^i(T))$ be the utility-maximizing bundle for group i . We will take the standard approach in selection problems of solving directly for the consumption-income bundles subject to selection constraints. We can then derive a tax schedule which supports these choices.³ Our approach here is to study the properties of Pareto efficient tax schedules, rather than those which maximize social welfare. (The two approaches are clearly closely related since, by varying the welfare weights for the groups, we can effectively map out the entire constrained Pareto frontier.)

B. Timing of the Randomization

The random action by the government is assumed to take place after individuals reveal their types, but before they decide their levels of effort. The government constructs two lotteries of tax schedules, one intended for individuals in group A and one for those in group B. Each individual must decide which lottery to participate in. Then, at random, a tax schedule is assigned to the individual. The individual, given his tax schedule, decides on the amount of income to produce. Since the individual has already declared his type, every schedule can be constructed to allow only one (C, Y) bundle to be chosen, even if the individual has misrevealed his type. Thus, each lottery can be viewed as a lottery over different (C, Y) bundles -- one for each possible tax schedule that may be drawn. We call this timing post-revelation pre-effort randomization.

Two other timing sequences are also possible. One of these resolves the uncertainty before agents reveal types to the tax authority. The government randomly assigns tax schedules to individuals without knowing their types. Each tax schedule in the lottery is equivalent to a pair of (C, Y) bundles, one intended for each type. Considering the government's problem as a choice of (C, Y) bundle, each pair of bundles must satisfy self-selection constraints. We will call this pre-revelation randomization.

The other alternative to our randomization scheme is to resolve the uncertainty after effort choices are made. In all outcomes, agents of the same type produce the same pre-tax income, but consumption differs depending upon the amount of tax. When individuals reveal their types (by choosing effort levels), these lotteries can be viewed as ones in which only consumption differs. This situation arises when tax laws are randomly enforced. This set of lotteries is feasible whether agents choose effort before or after resolution of the uncertainty. We denote this sequence post-revelation post-effort randomization. Figure 1 shows the timing of all three schemes.

We consider post-revelation pre-effort randomization since its outcome dominates, at least weakly, the other two sequences.

Theorem I: Post-revelation pre-effort randomization dominates either post-revelation post-effort randomization or pre-revelation pre-effort randomization.

All proofs are deferred to the Appendix.

If pre-effort randomization is not feasible, post-effort randomization might still be desirable for some sets of preferences. Similarly, if post-revelation pre-effort randomization is not possible, pre-revelation

randomization can be desirable if there are nonconvexities in the utility possibility frontier (see Stiglitz [1982]). This possibility exists because pre-revelation randomization can relax the government's budget constraint. There exists similar alternative timings of randomization in situations where both moral hazard and adverse selection are present. For insurance markets, the same three possibilities would arise if individuals of different risk classes (known to them) chose the level of care. Arnott and Stiglitz [1988] consider randomization in insurance markets with either adverse selection or moral hazard, but not both together.

C. The Optimal Tax Problem

To specify formally the government's optimization with the possibility of random taxation, consider the lotteries to be offered to the two groups A and B. Each group will be offered a set of bundles and a probability for each bundle in the set. The lotteries thus are $L^i((C^{ih}, Y^{ih}), \pi_{ih})$, $i = a, b$, and $h = 1, \dots, k(i)$ where $k(i)$ is some finite number. The government can choose both the bundles and the probabilities in each lottery. The maximization problem with randomization is:

$$\text{Max } \alpha N_a \sum_{h=1}^{k(a)} \pi_{ah} V^a(C^{ah}, Y^{ah}) + (1 - \alpha) N_b \sum_{h=1}^{k(b)} \pi_{bh} V^b(C^{bh}, Y^{bh}) \quad (1)$$

$$\text{s. t. } \sum_{h=1}^{k(i)} \pi_{ih} V^i(C^{ih}, Y^{ih}) \geq \sum_{h=1}^{k(j)} \pi_{jh} V^j(C^{jh}, Y^{jh}) \quad i = a, b, j \neq i : \lambda_i$$

$$N_a \sum_{h=1}^{k(a)} \pi_{ah} (C^{ah} - Y^{ah}) + N_b \sum_{h=1}^{k(b)} \pi_{bh} (C^{bh} - Y^{bh}) \leq 0 \quad : \mu$$

$$\sum_{h=1}^{k(i)} \pi_{ih} = 1, \quad i = a, b,$$

$$\pi_{ih} \geq 0, \quad i = a, b \text{ and } h = 1, \dots, k(i)$$

$$0 \leq Y^{ih} \leq K^i, \quad i = a, b, \quad h = 1, \dots, k(i)$$

$$0 \leq C^{ih} \leq M^i, \quad i = a, b, \quad h = 1, \dots, k(i)$$

α is a weight which can be varied to find the entire Pareto frontier, λ_i and μ are the multipliers on the self-selection and resource constraints. The upper bounds M^i and K^i bound the feasible set.⁴

The objective function and constraints incorporate ex ante expected values. For the self-selection constraints and the government's objective function, this is reasonable. The government designs the tax scheme and individuals decide which type to declare before realization. It is therefore of no consequence if, ex post, someone prefers the bundle obtained by someone of the other type. However, production occurs after the realization of the random process. Having total consumption exceed incomes in some realizations is not feasible even if expected total consumption and income are equal.⁵ Despite this, an ex ante resource constraint can be appropriate under some randomization procedures. Assume the government places one tax function for each individual of a particular type into an urn and the individuals who reveal that type draw without replacement. Then the ex ante expected net tax paid by an individual is exactly paid per person ex post by the group. Other procedures could also justify an ex ante constraint.⁶ We assume such a procedure is used.

D. Existence and Characterization of the Optimal Taxes

Under our assumptions, a solution to the optimal tax problem always exists.

Theorem II: For every α , there exists a pair of lotteries

$L^a((C^{ah}, Y^{ah}), \pi_{ah}), h=1, \dots, k(a)$ and $L^b((C^{bh}, Y^{bh}), \pi_{bh}), h=1, \dots, k(b)$

which solves the government's maximization problem.

To characterize the optimal lotteries, let $V^{i0}, i = a, b$, be the optimal utilities in the no-tax situation, let $((C^{ih}(\alpha), Y^{ih}(\alpha)), \pi_{ih}(\alpha)), i = a, b, h=1, \dots, k(i)$, be the optimal lotteries as a function of α , and let

$EV^i(\alpha) \equiv \sum_{h=1}^{k(i)} \pi_{ih}(\alpha) V^i(C^{ih}(\alpha), Y^{ih}(\alpha)), i = a, b$, be the optimal expected utilities achieved as functions of α . The deterministic problem is identical to this with the imposition of constraints that $k(a) = k(b) = 1$.

Theorem III: The solution to (1) satisfies the following properties:

(i) The self-selection constraint does not bind for any group whose expected utility is greater than in the no-tax situation. That is,

if $EV^i(\alpha) \geq V^{i0}$ then, $\sum_{h=1}^{k(j)} \pi_{jh}(\alpha) V^i(C^{jh}(\alpha), Y^{jh}(\alpha)) < EV^i(\alpha), i = a, b, j \neq i$.

(ii) A self-selection constraint can bind for at most one group. That is,

if $EV^i(\alpha) = \sum_{h=1}^{k(j)} \pi_{jh}(\alpha) V^i(C^{jh}(\alpha), Y^{jh}(\alpha))$ then $EV^j(\alpha) >$

$\sum_{h=1}^{k(i)} \pi_{ih}(\alpha) V^j(C^{ih}(\alpha), Y^{ih}(\alpha)), i = a, b, \text{ and } j \neq i$.

(iii) Production efficiency is satisfied in that total consumption equals total income. That is, $\mu > 0$.

(iv) If the self-selection constraint of one group is not binding, then it is not desirable to randomize the other group's bundle which is

undistorted. That is, if $\lambda_i = 0$, then $k(j) = 1$ and

$MRS^j(C^{j1}(\alpha), Y^{j1}(\alpha)) = 1, i = a, b, \text{ and } j \neq i$.

(v) There is no benefit to randomizing over more than three bundles. At each random bundle, the MRS of the group consuming there must lie between the MRS of the other group and the MRT (which equals 1). That is, for $i = a, b$, and $j \neq i$, if $\lambda_j > 0$, then an optimal solution exists with $k(j) \leq 3$ and at each $h = 1, \dots, k(j)$, either

$$1 < MRS^j(C^{jh}(\alpha), Y^{jh}(\alpha)) < MRS^i(C^{jh}(\alpha), Y^{jh}(\alpha)),$$

$$MRS^i(C^{jh}(\alpha), Y^{jh}(\alpha)) < MRS^j(C^{jh}(\alpha), Y^{jh}(\alpha)) < 1, \text{ or} \quad (2)$$

$$MRS^i(C^{jh}(\alpha), Y^{jh}(\alpha)) = MRS^j(C^{jh}(\alpha), Y^{jh}(\alpha)) = 1.$$

Conditions (i)-(iv) are essentially the same as in the deterministic optimal tax problem with utilities replaced by expected utilities. (See Brito, et al. [1990].) Note that condition (ii) holds at the optimum even if multiple crossings of the indifference makes it feasible to have both self-selection constraints binding. Condition (v) differs somewhat from the deterministic case since the possibility of $MRS^i = MRS^j = 1$ cannot be ruled out at one of the random bundles. Note that, at one of the random bundles, both MRS's may exceed 1, while at another they may both be smaller. If a stronger single crossing assumption were made, then both MRS's would have to be strictly greater than 1 at all of the random bundles or strictly less at all of the random bundles.

At the optimum, at most one class could face a nondegenerate lottery and that lottery never requires more than three bundles. Theorem III does not tell us whether either group actually faces a nondegenerate lottery. First order conditions are of little use for showing that a nondegenerate lottery exists or for finding the random solution, since the solution when randomization is prohibited also satisfies the first order conditions when

randomization is allowed. To see this, denote the nonrandom solution as a lottery over $k(i)$ bundles where each bundle is identical to the optimal nonrandom bundle. First order conditions with respect to π_{ih} are trivially satisfied and those with respect to C^{ih} and Y^{ih} reduce to those in the deterministic problem. Hence, a nondegenerate random solution exists only when multiple solutions to the first order conditions exist, with the extra solutions asymmetric (and not readily apparent). However, since the self-selection constraints involve the difference in utilities at different bundles and therefore do not define convex sets, multiple solutions are possible. In the next section, we consider when randomization can be shown to be desirable.

III. NECESSARY AND SUFFICIENT CONDITIONS FOR RANDOMIZATION

Solving for the optimal random solution, except for special examples, can be quite difficult. Instead, in this section, we consider conditions under which the optimal deterministic solution can be improved upon by introducing randomness in the solution. Whether randomization is desirable depends on the concavity of the utility representations and on the shapes of the indifference curves. In Theorem IV, we focus on the utility representations and show that for any pair of indifference curves which differ at the optimal deterministic bundle, randomizing is beneficial provided the utility representation is sufficiently concave.⁷

Theorem IV: Consider any V^i and V^j satisfying assumption (A1)-(A3) and any α such that the self-selection constraint of class i is strictly binding ($\lambda_i > 0$) in the deterministic problem. Then there exists a concave increasing transformation of V^i such that $k(j) > 1$ holds at the optimum, whether or not V^j is similarly transformed.

To establish this result, we randomize over a set of bundles between any of which j is indifferent, but i is not. See Figure 2 for an illustration of the construction. The transformation of utility that makes i 's utility function more concave reduces the desirability of j 's lottery for individuals in group i . This weakens i 's self-selection constraint and allows for adjustments which raise j 's expected utility. Having the initial randomization over two bundles which are chosen to be indifferent to the optimum nonrandom bundle given to j is a convenient way to demonstrate that an improvement over the deterministic solution exists. This particular randomization has a special property--it does not violate horizontal equity (defined as identical agents having equal utilities even if they have different budget sets). All type j 's obtain equal utility, although with different consumption bundles. The randomization causes utilities to be stochastic only for type i 's who would mimic type j 's. In equilibrium, of course, no imitation occurs. While this may not be the optimal randomization, it remains feasible even when ex post horizontal equity is required.

The argument underlying Theorem IV is not limited to models with just two classes of individuals but generalizes directly to models with any finite set of classes of individuals. Consider the solution to the optimal tax problem when randomization is prohibited. Consider any bundle (C^0, Y^0) which is assigned to only one class and which some other classes view as indifferent to the bundles assigned to them. Replacing (C^0, Y^0) with some lottery over bundles on the same indifference curve for the class to whom it is assigned will weaken the self-selection constraints of the other classes who are indifferent to taking (C^0, Y^0) , provided those classes have sufficiently concave utility representations. Only when a pooling equilibrium exists will it not necessarily be possible to replace the deterministic bundle with a

random one over bundles which yield the same utility to all groups assigned the original bundle.

In Theorem IV, only the utility function of individuals in group i is transformed. However, making the utility function of individuals in group j more concave as well does not change the result since the randomization we use in establishing the result involves bundles over which j is indifferent. Hence, the result also applies to situations in which the two groups have the same utility functions but differ in ability.

Theorem IV shows that the desirability of randomization cannot be ruled out by looking just at indifference maps, since randomization can arise for essentially any indifference maps under some utility representation. On the other hand, the shape of the indifference map does have some effect on whether randomization is desirable. To investigate this, we restrict attention to local randomization. A local randomization is improving if, for any positive ϵ , there exists a lottery each of whose outcomes is within ϵ of the deterministic solution and which is Pareto preferred to it. The next theorem gives a necessary and sufficient condition for local randomization which depends in part on the curvature of indifference curves. Assume $\lambda_a > 0$ in the deterministic solution to the problem so that the marginal tax rate is not zero at (C^b, Y^b) . Let $H^i(C^b, Y^b)$ be the Hessian of the utility function of type i at (C^b, Y^b) .

Theorem V: Assume $\lambda_a > 0$ in the nonrandom solution and consider any probability triple (π_1, π_2, π_3) , at least two of which are positive and $\pi_1 + \pi_2 + \pi_3 = 1$. Let $\theta = \{V_C^b(C^b, Y^b) | 1 - MRS^b(C^b, Y^b) | \} / \{V_C^a(C^b, Y^b) | 1 - MRS^a(C^b, Y^b) | \}$. There exists some local randomization with these probabilities around the nonrandom solution which improves on that solution iff the matrix $H^b(C^b, Y^b) - \theta H^a(C^b, Y^b)$ is not negative semidefinite.

The condition in Theorem V depends upon properties of the indifference map directly through the term θ and indirectly through the Hessians H^a and H^b which are affected by both the curvature of indifference curves and the degree of concavity of the utility function. Both Hessians H^a and H^b are negative definite. Essentially, by taking the difference between them, the condition is satisfied or violated depending upon which of them is more concave at B's bundle. If V^b is more concave, then H^b will dominate H^a and the difference between them will also be negative definite and local randomization will tend not to be desirable. If V^a is more concave, then the difference between the two Hessians is less likely to be negative definite and local randomization will be beneficial. This is complicated by the fact that H^a is multiplied by θ . From Theorem III(V), $|1 - MRS^b(C^b, Y^b)| < |1 - MRS^a(C^a, Y^a)|$. Hence, although $V_c^a(C^b, Y^b)$ and $V_c^b(C^b, Y^b)$ are hard to compare in general, there is some tendency for θ to be less than 1. This in turn increases the extent to which the concavity of V^a must exceed that of V^b for local randomization to be desirable. To summarize, local randomization is desirable if and only if the utility function of the group actually consuming the bundle is significantly less concave than that of the group whose self-selection constraint is binding at that bundle.

When the indifference curves of the two groups have very different slopes, then θ will tend to be smaller than 1 and local randomization is less likely to be desirable. It does not follow that if the two groups have very similar preferences the possibilities for local randomization are increased since both will not only be near each other but will also be close to 1. In fact, as shown in the next corollary, if the preferences of A's and B's are similar, local randomization is not desirable. Assume the utility functions belong to a family parameterized by a scalar P , so that they become

$v^a(C, Y, P_a)$ and $v^b(C, Y, P_b)$.

Corollary I: Assume that preferences are related such that at all $P = P_a = P_b$, $v^a(C, Y, P) = F(v^b(C, Y, P))$. Then, for P_a near P_b , no local randomization is desirable unless F is sufficiently concave.

A special case satisfying the assumptions needed for this result is the standard case of two classes with identical preferences but different abilities. If the abilities are close to each other, no local randomization is desirable regardless of which group the redistribution favors.⁸ Both Theorem IV and Corollary I show that concavifying utility makes randomization desirable. Theorem IV concerns the degree of concavity of each utility function separately, while Theorem V and Corollary I involve the concavity of one utility function relative to the other. For P_a near P_b , if both functions undergo the same transformation, no local randomization will be desirable. Theorem IV demonstrates that, if ordinal preferences differ between the groups, randomization will be desirable if the common transformation is extreme enough. Thus, this randomization must be a nonlocal randomization of bundles offered a group in a lottery.

A crucial point about Theorem V is that the condition for local randomization given any probability triple is independent of the probabilities. If local randomization can improve on the deterministic solution given one probability triple then it will be beneficial for any probability triple. Any restrictions on the possible probabilities which prevent the optimal randomization from being achieved will not prevent the deterministic solution from being dominated by some other suboptimal randomization.

In order to see that the condition of Theorem V can be satisfied, consider the special case where all individuals have identical additive utility functions over consumption and labor with the groups differing only in ability.

Corollary II: Consider the utility functions $V^i(C, Y) = \psi(C) - \gamma(Y/w_i)$, $i = a, b$, where the w_i are ability parameters with $0 < w_b < w_a$. To satisfy concavity of V^i , ψ is concave and γ convex. Let $L^i = Y/w_i$.

- (i) No randomization, local or nonlocal, is desirable either if $\lambda_a > 0$ and $L\gamma'''(L)/\gamma'(L) \geq -2$ at all L or if $\lambda_b > 0$ and $L\gamma'''(L)/\gamma'(L) \leq -2$ at all L .
- (ii) Consider any probabilities (π_1, π_2, π_3) with at least two positive and with $\pi_1 + \pi_2 + \pi_3 = 1$. There exists some local randomization with these probabilities around the nonrandom optimum which improves on the nonrandom solution iff:

$$\lambda_a > 0 \text{ and } \frac{\gamma''(Y^b/w_a)}{w_a^2(1 - MRS^a(C^b, Y^b))} > \frac{\gamma''(Y^b/w_b)}{w_b^2(1 - MRS^b(C^b, Y^b))} \quad (3)$$

$$\text{or } \lambda_b > 0 \text{ and } \frac{\gamma''(Y^a/w_b)}{w_b^2(MRS^b(C^a, Y^a)-1)} > \frac{\gamma''(Y^a/w_a)}{w_a^2(MRS^a(C^a, Y^a)-1)} \quad (4)$$

Consider condition (3). Since $w_a > w_b$ and $1 - MRS^a(C^b, Y^b) > 1 - MRS^b(C^b, Y^b) > 0$, the denominator of the LHS exceeds that of the RHS in (3). Hence, $\gamma''(Y^b/w_a)$ must exceed $\gamma''(Y^b/w_b)$ by a sufficient amount for this condition to be satisfied. Since $Y^b/w_a < Y^b/w_b$, this can be satisfied by a sufficiently negative value of γ''' . A sufficient condition for (3) to be

satisfied is that the derivative of $\bar{y}''(Y/w)/[w^2(1 - MRS)]$ with respect to w be positive. Taking this derivative yields the following sufficient condition for the desirability of local randomization:

$$L\bar{y}'''(L)/\bar{y}''(L) < -2 - [\bar{y}'(L) + \bar{y}''(L)]/[w\psi'(C) - \bar{y}'(L)] \quad (5)$$

where $w\psi'(C) - \bar{y}'(L) = w\psi'(C)(1 - MRS(C^b, Y^b)) > 0$. This is stricter than the necessary condition for randomization in part (i) since it requires $L\bar{y}'''(L)/\bar{y}''(L)$ to be strictly less than -2 by a finite amount.

These necessary conditions for randomization with $\lambda_a > 0$ and $\lambda_b > 0$ are both violated unless $L\bar{y}'''(L)/\bar{y}''(L) = -2$. Thus, randomization is desirable nowhere on the utility possibility frontier (UPF) if $L\bar{y}'''(L)/\bar{y}''(L) = -2$.⁹ Furthermore, randomization cannot be desirable both with $\lambda_a > 0$ and $\lambda_b > 0$ unless $L\bar{y}'''(L)/\bar{y}''(L) + 2$ changes sign along the UPF as labor supply changes. Since sufficient conditions for local randomization with $\lambda_a > 0$ and $\lambda_b > 0$ are opposite in sign, randomization will be desirable somewhere on the UPF for many allowable $\bar{y}(\cdot)$ functions.¹⁰

IV. BENEFITS OF RANDOMIZATION: AN EXAMPLE

The potential desirability of randomization in self-selection problems has been shown in several contexts. In this paper, we have given a more intuitive explanation of why randomization can yield improvements and have sought to determine the scope of randomization by deriving a necessary and sufficient condition for some local randomization to be preferred to the nonrandom solution. These results do not specify the optimal randomization and, therefore, do not give any insight into the magnitude of the gain from using randomized solutions. In the following example, the problem is explicitly solved and the optimal random and nonrandom solutions explicitly

compared. Although the example has special features, it shows that the gains from randomization can be quite large. Hence, in general, randomization cannot be disregarded as only a minor modification of the nonrandom solution.

Assume an infinite population with two types of individuals which have equal measure. Each consumes a commodity, C , and leisure, $R \equiv 1 - Y$. The two types have equal earning ability but different tastes. We use a slightly different notation here to reflect this difference. There is no need to transform the problem to get a common budget constraint. Their von Neumann-Morgenstern utility functions are $U^a(C^a, R^a) = \log C^a + \log R^a$ and $U^b(C^b, R^b) = C^b + \log R^b$. C and R are constrained by an aggregate production possibility frontier $C + R = K$. For convenience in analyzing the solution, we consider only the case of $K > 4$. The first best Pareto problem can then be written as:

$$\begin{aligned}
 U^b(U^a, K) &\equiv \text{Max } C^b + \log R^b & (6) \\
 \text{s.t. } &\log C^a + \log R^a \geq 2 \log U^a \\
 &C^a + C^b + R^a + R^b = K \\
 &C^a, C^b, R^a, R^b \geq 0
 \end{aligned}$$

The maximized value function $U^b(\hat{U}^a, K)$ is a representation of the utility possibility frontier. Explicitly solving this problem yields the solution:

$$\begin{aligned}
 C^a = R^a = \hat{U}^a, C^b = K - 1 - 2\hat{U}^a, \text{ and } R^b = 1, & \text{ if } \hat{U}^a \leq (K-1)/2 \\
 C^a = R^a = \hat{U}^a, C^b = 0, \text{ and } R^b = K - 2\hat{U}^a, & \text{ if } \hat{U}^a \geq (K-1)/2
 \end{aligned}$$

Substituting these solutions into the utility functions yields the first best Pareto frontier.

$$\begin{aligned}
 U^b &= K - 1 - 2\hat{U}^a & \text{if } \hat{U}^a \leq (K-1)/2 \\
 U^b &= \log(K - 2\hat{U}^a) & \text{if } (K-1)/2 \leq \hat{U}^a < K/2
 \end{aligned}$$

For any values of \hat{U}^a and K , the first best solution can be checked to see if either type would prefer the bundle of the other type of individual. It then follows that for $0 \leq \hat{U}^a < K^{\frac{1}{2}} - 1$, type A individuals prefer the bundle given to type B's. For $K^{\frac{1}{2}} - 1 \leq \hat{U}^a \leq U^*$, neither self-selection constraint binds where U^* is implicitly defined by $3U^* + \ln U^* = K - 1$ for which $K^{\frac{1}{2}} - 1 < U^* < K - 1$ holds. Finally, for $U^* < \hat{U}^a < K/2$, the self-selection constraint of type B individuals is binding. See Figure 3.

Consider the nonrandom second best problem with self-selection:

$$\begin{aligned} \hat{U}^b(\hat{U}^a, K) &\equiv \text{Max } C^b + \log R^b && (7) \\ \text{s.t. } \log C^a + \log R^a &\geq 2 \log \hat{U}^a \\ C^a + R^a + C^b + R^b &= K \\ \log C^a + \log R^a &\geq \log C^b + \log R^b \\ C^b + \log R^b &\geq C^a + \log R^a \\ C^a, R^a, C^b, R^b &\geq 0 \end{aligned}$$

From the first best results, for $0 \leq \hat{U}^a < K^{\frac{1}{2}} - 1$, the constraint $\log C^a + \log R^a \geq \log C^b + \log R^b$ holds with equality and the constraint $C^b + \log R^b \geq C^a + \log R^a$ holds with strict inequality. For $K^{\frac{1}{2}} - 1 \leq \hat{U}^a \leq U^*$, neither self-selection binds and the second best solution is also the first best one. For $U^* < \hat{U}^a < K/2$, $\log C^a + \log R^a > \log C^b + \log R^b$ and $C^b + \log R^b = C^a + \log R^a$ must hold at the second best solution.

Randomization is never desirable in the region when neither self-selection constraint binds. In this problem, for \hat{U}^a large enough that the self-selection constraint for type B individuals binds, randomization is also not desirable. In that circumstance, only the type A bundle could be random. But, if type A consumption or leisure were random, its certainty equivalent level would be less than its expected value. Replacing the random

distribution by its certainty equivalent would leave the self-selection and Pareto constraints still satisfied. Extra resources would also become available. This would allow an increase in both types' utilities. Thus, since we are interested in the benefits of randomization, the second best problem is solved only in the region with $0 \leq \hat{U}^a < K^{\frac{1}{2}} - 1$ and the type A self-selection constraint binding.

For low levels of K and \hat{U}^a , C^b may equal zero in the second best solution. Note that the self-selection constraint for type A's is always satisfied if $C^b = 0$ and thus no randomization is necessary. For $\hat{U}^a \geq 1$ and $K > 4$, $C^b > 0$ holds in all second best optima. For $1 \leq \hat{U}^a < K^{\frac{1}{2}} - 1$, the self-selection constraint and the minimum utility constraint for type A's bind. The solution is $C^a = R^a = \hat{U}^a$, $C^b = [K - 2\hat{U}^a + (K^2 - 4\hat{U}^a K)^{\frac{1}{2}}]/2$, and $R^b = [K - 2\hat{U}^a - (K^2 - 4\hat{U}^a K)^{\frac{1}{2}}]/2$. In this region, $U^b = \log [(K - 2\hat{U}^a - (K^2 - 4\hat{U}^a K)^{\frac{1}{2}})/2] + [K - 2\hat{U}^a + (K^2 - 4\hat{U}^a K)^{\frac{1}{2}}]/2$. The first and second best Pareto frontiers are shown in Figure 4.

Consider the possibility of randomization when $1 \leq \hat{U}^a < K^{\frac{1}{2}} - 1$. If C^b is random, the self-selection constraint for type A individuals is now $\log C^a + \log R^a \geq E(\log C^b) + \log R^b$. Consider any random solution with $E(C^b) = C^b$ where C^b is the first best solution. Then the resource constraint continues to be satisfied. The minimum utility constraint is unaffected. The utility of type B's is unchanged. Now consider the self-selection constraint. If, in the randomization, $C^b = 0$ arises with any probability $\pi > 0$, then $E(\log C^b) = -\infty$ and the self-selection constraint is satisfied for C^a and R^a at their first-best levels.

Of course, this solution violates ex post horizontal equity. A solution almost as good which satisfies horizontal equity can be found. Let $C_1^b = 0$, $C_2^b = [K - 1 - 2U^a]/(1 - \pi)$, $R_2^b = [1 - \pi R_1^b]/(1 - \pi)$ and R_1^b determined to

satisfy $\log R_1^b = C_2^b + \log [(1 - \pi R_1^b)/(1 - \pi)]$. As π goes to zero, R_1^b converges to e^{K-1-2U^a} . The bundles $(0, R_1^b)$ and $(C_2^b, (1 - \pi R_1^b)/(1 - \pi))$ lie on the same indifference curve for type B individuals. As π goes to zero, the (C_2^b, R_2^b) bundle arises almost all the time and is arbitrarily close to the first best bundle. The bundle $(0, R_1^b)$ arises with small probability. Thus, $E(U^b)$ is approximately equal to the first best level.

Next, we show that significant gains can result by achieving the first rather than the second best. Consider $U^a = 1$. Consider the first best and assume society discards a fraction δ of the available resources K such that the first best and nonrandom second best yield the same outcome. Since there is only a single resource constraint, $1 - \delta$ is equivalent to Debreu's [1951] coefficient of resource utilization. That is, δ solves $(1 - \delta)K - 3 = [K - 2 + (K^2 - 4K)^{\frac{1}{2}}]/2 + \log ([K - 2 - (K^2 - 4K)^{\frac{1}{2}}]/2)$. Solving for $1 - \delta$ yields:

$$1 - \delta = 1/2 + [4 + (K^2 - 4K)^{\frac{1}{2}}]/2K + [\log ([K - 2 - (K^2 - 4K)^{\frac{1}{2}}]/2)]/K$$

At $K = 4$, $1 - \delta = 1$ and no resources can be discarded. The maximum value of δ occurs approximately at $K = 9.5$ where δ is approximately equal to 12.7%.

Of course, this example has two significant special features. U^b is linear in R^b so type B's are risk neutral with respect to this commodity while U^a is so highly risk averse with respect to C^a that as C^a goes to zero, U^a goes to negative infinity. Both of these assumptions can be relaxed and randomization will remain desirable, but the benefits will decline.

V. CONCLUSIONS

The use of a random solution creates ex post asymmetries among individuals within the same class. Given strictly concave utility functions,

the individuals would like to reduce the asymmetries. This could be done in several ways. Either the individuals in a class could directly pool risks with each other or they could seek to buy insurance against the uncertainty induced by the randomization in the tax schedules. Either of these would frustrate the government's attempt to relax the self-selection constraints through the use of randomization. These problems may not be severe since individuals might find trade in leisure hours difficult. To use randomization successfully, the government must inhibit the access of individuals to other markets. To rule out risk pooling, the government would have to be able to monitor individuals to be certain that income reported by one individual was actually earned by that individual and not by someone else. In ruling out insurance against government randomization, the government would have to be careful not to prevent socially desirable insurance against exogenous uncertainty.

This type of problem is not unique to randomization. If the government is restricted to nonstochastic tax schedules, similar problems arise. The consumption-income opportunity sets defined by the optimal tax schedules are nonconvex. Individuals would seek to frustrate the government's exploitation of nonconvex schedules by engaging in trades or by gambling to convexify their opportunity sets. If the government cannot prevent these types of behavior, further constraints on the feasible set of non-random tax structures need to be imposed. It is unclear whether the administrative problems in preventing these types of responses to the government policies are greater with randomization or without. One simplifying feature of a practical randomization scheme is that the probabilities need not be chosen optimally for randomization to be desirable.

In summary, we have offered a general characterization of the random

solution to the optimal income tax problem and have presented conditions under which randomization is desirable in order to gain insight into the question of whether randomized solutions are merely a curiosum or whether they should be significant in the design of government policies. Randomization loosens self-selection constraints. We have argued that random solutions need not violate horizontal equity, can arise for any pair of ordinal preferences for the two classes which differ provided one or both classes have significantly concave utility representations, and can yield significant gains over nonrandom solutions.

Footnotes

1. Brito, et al. [1990], characterize the deterministic optimal income tax problem under these weak assumptions for the case of n classes. Guesnerie and Seade [1982] also establish properties of the solution without assuming single crossing.
2. This requires an additional assumption on $U(C,L)$. Note that $MRS^i(C,Y) = - [U_L(C,L^i)/w_i U_C(C,L^i)]$. If $w_a > w_b$ then $L^a < L^b$. The result holds if the direct effect of the higher wage is not countered by the effects of a lower L on the MRS. Differentiating $-[U_L(C,Y/w)/w U_C(C,Y/w)]$ with respect to w yields $dMRS(C,Y)/dw = (U_L/w U_C) - (L/w^2)d(-U_L/U_C)/dL$. A sufficient condition for $dMRS(C,Y)/dw < 0$ is $d(-U_L/U_C)/dL \geq 0$ which holds if C is not inferior. See Sadka [1976].
3. The schedule $T(Y)$ need not be differentiable. In fact, it will generally be nondifferentiable at the incomes chosen by the two groups.
4. In the deterministic problem, the bounds on Y^i along with resource balance automatically bound C^i . Here, bounds on C^{ih} do not follow from the expected resource constraint since that constraint bounds the products $\pi_{ih} C^{ih}$ only. As some π_{ih} goes to zero, the corresponding C^{ih} could be made arbitrarily large. M^i could be chosen sufficiently large to bound the feasible set without affecting the solution.
5. If some C^{ih} exceeds $N_a K^a + N_b K^b$ (maximum total production), even if it arises with very low probability, the outcome is not feasible when this bundle arises. Such ex post violations of feasibility can clearly occur even if the M^i are set to restrict each C^{ih} to be less than $N_a K^a + N_b K^b$.
6. An alternative justification of the constraint is that each individual of a revealed type draws a tax schedule at random from the same distribution. As the number of individuals gets large, the distribution of ex post resource balance will always have values violating the constraint. However, the law of large numbers implies that the per capita violation of the constraint goes to zero. Thus if the

consumption vectors are reduced by ϵ , the probability that the constraint is satisfied goes to 1 as the population grows. Hence, the solution can be viewed as an ϵ -equilibrium. Obtaining only an ϵ -equilibrium with randomization is not a real restriction since that is all that a deterministic solution can achieve. At least one self-selection constraint holds with equality at the optimum. Some individuals then are indifferent over two bundles only one of which is it desirable for them to choose. Guaranteeing that these individuals select the appropriate bundle can only be done by deviating slightly from the optimal bundles.

7. Maskin and Riley [1980] have a similar result that a sufficient concavification of the utility function can make randomization desirable. Their result is for a continuum of individuals with identical preferences but different abilities. They require the additional assumption that preferences ordinally are additive in consumption and labor and are linear in consumption.
8. This result holds not just because little redistribution is desired when individuals have similar abilities so that neither self-selection constraint binds. In the deterministic problem, if $\alpha = 0$ and the A's are more able, redistribution is carried out until A's self-selection constraint binds. Randomization will not improve on this solution.
9. If $\bar{x} = K_1 \ln L + K_2 L$, where $K_1 < 0$ and $K_2 > 0$, then $L\bar{x}''/\bar{x}''' = -2$ so that randomization will never occur. ($\bar{x}' > 0$ holds for this function at small values of L.) If there is quadratic disutility of labor with $\bar{x} = KL^2$, then (i) of Corollary II is satisfied so that randomization does not occur when $\lambda_a > 0$. Substituting into (4) shows that local randomization cannot occur when $\lambda_b > 0$. Whether nonlocal randomization can occur is unclear.
10. Stiglitz [1982] derived the conditions in Theorem V and Corollary II as sufficient for local randomization over two bundles arising with equal probability. The results here are stronger: these conditions are both necessary and sufficient for the existence of a local randomization that

improves on the nonrandom solution. The same condition applies for any arbitrarily given probabilities over three bundles. Due to an error in calculation, Stiglitz [1982] asserts a result opposite that of Corollary I.

References

- Arnott, R.J. and J.E. Stiglitz, "Randomization with Asymmetric Information: A Simplified Exposition," Rand Journal of Economics, 19: 344-362, (1988).
- Brito, D., J. Hamilton, S. Slutsky and J. Stiglitz, "Pareto Efficient Tax Structures," Oxford Economic Papers, 1990 forthcoming.
- Debreu, G., "The Coefficient of Resource Utilization," Econometrica 19: 273-292, (1951).
- Fellingham, J.C., Y.K. Kwon and D.P. Newman, "Ex Ante Randomization in Agency Models," Rand Journal of Economics 15: 290-301 (1984).
- Guesnerie, R. and J. Seade, "Nonlinear Pricing in a "Finite Economy," Journal of Public Economics 17: 157-179 (1982).
- Maskin, E. and J. Riley, "Income v. Leisure Taxes," Discussion paper no. 174, UCLA Department of Economics, (1980).
- _____, "Optimal Auctions with Risk-Averse Buyers," Econometrica 52: 1473-1518 (1984).
- Mirrlees, J.A., "An Exploration in the Theory of Optimum Income Taxation," Review of Economic Studies 38: 175-208 (1971).
- Sadka, E., "On Income Distribution, Incentive Effects and Optimal Income Taxation," Review of Economic Studies 43: 261-268, (1976).
- Stiglitz, J.E., "Self-Selection and Pareto Efficient Taxation," Journal of Public Economics 17: 213-240 (1982).
- Weiss, L.W., "The Desirability of Cheating Incentives and Randomness in the Optimal Income Tax," Journal of Political Economy 89: 1343-1352 (1976).

APPENDIX

PROOFS OF THEOREMS

PROOF OF THEOREM I: Pre-effort randomization has a lottery over (C,Y) bundles that differ in both arguments. In post-effort randomization, all bundles in the lottery have the same Y. The extra flexibility available to the government makes pre-effort randomization at least weakly superior.

In pre-revelation randomization, the government randomly assigns a tax schedule to each individual. Corresponding to the tax schedules are a pair of bundles, one for each type. Every such pair must satisfy self-selection constraints. Post-revelation pre-effort randomization imposes self-selection constraints only on expected utilities over the lotteries. This latter constraint is weaker and thus the government's objective function may attain a higher value. Q.E.D.

PROOF OF THEOREM II: The constraint set is nonempty since the deterministic solution yields a set of degenerate lotteries which are feasible in the problem with randomization. The constraint set is closed and all variables $(\pi_{ih}, C^{ih}, Y^{ih})$ are bounded. Hence, the constraint set is compact. The continuous objective function then attains a maximum on the nonempty compact constraint set. Q.E.D.

PROOF OF THEOREM III: (i) Let $\bar{C}^i(\alpha) \equiv \sum_{h=1}^{k(i)} \pi_{ih}(\alpha) C^{ih}(\alpha)$ and $\bar{Y}^i(\alpha) \equiv \sum_{h=1}^{k(i)} \pi_{ih}(\alpha) Y^{ih}(\alpha)$, $i = a, b$. Then, since $V^i(C^i, Y^i)$ is strictly concave $V^i(\bar{C}^i(\alpha), \bar{Y}^i(\alpha)) \geq EV^i(\alpha)$, with strict inequality if the type i lottery is

nondegenerate. If, contrary to the statement of the theorem, $\sum_{h=1}^{k(j)} \pi_{jh}(\alpha) V^i(C^{jh}(\alpha), Y^{jh}(\alpha)) = EV^i(\alpha)$, then $V^i(\bar{C}^j(\alpha), Y^j(\alpha)) \geq EV^i(\alpha)$, with strict inequality if the type j lottery is nondegenerate, also follows from strict concavity. By assumption, $EV^i(\alpha) \geq V^{i0}$ so that $V^i(\bar{C}^i(\alpha), \bar{Y}^i(\alpha)) \geq V^{i0}$ and $V^i(\bar{C}^j(\alpha), \bar{Y}^j(\alpha)) \geq V^{i0}$ must hold. Since the indifference curve through V^{i0} is tangent to the line $C = Y$, $\bar{C}^i(\alpha) > \bar{Y}^i(\alpha)$ and $\bar{C}^j(\alpha) > \bar{Y}^j(\alpha)$ must hold. Hence, $N_i(\bar{C}^i(\alpha) - \bar{Y}^i(\alpha)) + N_j(\bar{C}^j(\alpha) - \bar{Y}^j(\alpha)) > 0$ must hold. But $N_i \sum_{h=1}^{k(i)} \pi_{ih}(\alpha)(C^{ih}(\alpha) - Y^{ih}(\alpha)) + N_j \sum_{h=1}^{k(j)} \pi_{jh}(\alpha)(C^{jh}(\alpha) - Y^{jh}(\alpha)) = N_i(\bar{C}^i(\alpha) - \bar{Y}^i(\alpha)) + N_j(\bar{C}^j(\alpha) - \bar{Y}^j(\alpha))$, so resource balance would be violated if the self-selection constraint for i holds with equality.

(ii) Even with randomization, the point (V^{a0}, V^{b0}) must lie on the constrained utility possibility frontier. There exists α_0 such that $V^i(\alpha_0) = V^{i0}$, $i = a, b$. For any $\alpha \neq \alpha_0$, either $V^a(\alpha) > V^{a0}$ or $V^b(\alpha) > V^{b0}$. From (i), the group better off than in the no-tax case cannot have a binding self-selection constraint. Thus, both self-selection constraints cannot bind.

(iii) Any increased resources can always be given to the group whose self-selection constraint is not binding without violating any other constraint. Since this will raise the value of the objective function, the multiplier on the resource constraint is strictly positive.

(iv) Consider the first order conditions with respect to each C^{jh} and Y^{jh} . In each condition, π^{jh} multiplies every term, so it drops out leaving conditions identical in form to those in the deterministic problem. Since $\lambda_i = 0$, it is immediate that at each h , $MRS^j(C^{jh}, Y^{jh}) = 1$.

To see that $k(j) > 1$ is not needed at an optimum, consider the case where the self-selection constraint for group i holds with strict inequality (that is, $\sum_{h=1}^{k(i)} \pi_{ih} V^i(C^{ih}, Y^{ih}) > \sum_{h=1}^{k(j)} \pi_{jh} V^i(C^{jh}, Y^{jh})$). Having $k(j) > 1$ in a

nondegenerate lottery (some π_{jh} and π_{jk} positive and $(C^{jh}, y^{jh}) \neq (C^{jk}, y^{jk})$) cannot be optimal. With the probabilities fixed, consider a change in the (C^{jh}, y^{jh}) which moves each closer to the mean bundle without changing the mean. Such a shift will continue to satisfy both the bounds on the C^{jh} and y^{jh} and the resource balance constraint. The probabilities do not change, satisfying the constraints that they sum to 1. From concavity, $\sum_{h=1}^{k(j)} \pi_{jh} v^i(C^{jh}, y^{jh})$ will rise but since the self-selection constraint for i holds with strict inequality, if the shift is not too large, this constraint will still hold. Since $\sum_{h=1}^{k(j)} \pi_{jh} v^j(C^{jh}, y^{jh})$ rises, the self-selection constraint for group j must still hold. Thus, the bundles after the shift yield a feasible lottery. Since $\sum_{h=1}^{k(j)} \pi_{jh} v^j(C^{jh}, y^{jh})$ rises, the objective function rises showing that the prior bundle was not optimal.

(v) The first order conditions with respect to C^{jh} and y^{jh} are the same as those in the deterministic problem since the π_{jh} cancel from the conditions. To show the condition on the MRS's, it is enough to consider only degenerate lotteries. Assume $\lambda_a > 0$ so that from (ii), $\lambda_b = 0$ must hold. The first order conditions with respect to C^b and y^b are:

$$(1 - \alpha)N_b v_c^b(C^b, y^b) - \lambda_a v_c^a(C^b, y^b) - \mu N_b = 0$$

$$(1 - \alpha)N_b v_y^b(C^b, y^b) - \lambda_a v_y^a(C^b, y^b) + \mu N_b = 0$$

Solving for μN_b and combining terms yields:

$$(1 - \alpha)N_b [v_c^b(C^b, y^b) + v_y^b(C^b, y^b)] = \lambda_a [v_c^a(C^b, y^b) + v_y^a(C^b, y^b)]$$

or: $(1 - \alpha)N_b v_c^b(C^b, y^b) [1 - MRS^b(C^b, y^b)] = \lambda_a v_c^a(C^b, y^b) [1 - MRS^a(C^b, y^b)]$

Adding $\lambda_a v_c^a(C^b, y^b) MRS^b(C^b, y^b)$ to both sides, combining terms, and substituting the first order condition for C^b yields:

$$\mu N_b [1 - MRS^b(c^b, Y^b)] = \lambda_a V_c^a(c^b, Y^b) [MRS^b(c^b, Y^b) - MRS^a(c^b, Y^b)]$$

Since λ_a (by assumption) and μ (from part (iii)) are positive, the result follows. Unlike the solution when random solutions are prohibited, both MRS's equal to unity can not be ruled out.

If $\lambda_b > 0$, similar manipulations of the first order conditions for C^a and Y^a show the required results.

To see that $k(j) > 3$ is not required, consider the optimal lotteries. Fix the quantities (c^{ih}, y^{ih}) , $i = a, b$, and $h = 1, \dots, k(i)$, making (1) a linear programming problem in the π_{ih} . From part (iv), one of the lotteries is degenerate with $\pi_{ji} = 1$. From part (ii), at most one self-selection constraint can bind. Thus, of the five original constraints, at most three (a self-selection constraint, the resource balance constraint, and the constraint that the probabilities in the nondegenerate lottery sum to 1) are binding. A solution exists with the number of nonzero variables no greater than the number of binding constraints. Q.E.D.

PROOF OF THEOREM IV: The proof requires the following lemma.

Lemma I: Consider any utility function $U(X)$, any bundles X^1 , X^2 , and X^3 with $U(X^1) > U(X^2)$ and $U(X^3) > U(X^2)$, and any π , $0 < \pi < 1$. There exists some concave increasing transformation $V(X) = F(U(X))$ such that $\pi V(X^1) + (1 - \pi)V(X^2) < V(X^3)$.

Proof: It is sufficient to provide an example of a transformation under which the result holds. Let $U^2 \equiv U(X^2)$ and consider $F_n(U) = (U - U^2)^{1/n}$. Then $V_n(X^2) = (U(X^2) - U^2)^{1/n} = 0$ for all n . The result then reduces to there existing an n^* with $\pi V_{n^*}(X^1) < V_{n^*}(X^3)$. This in turn follows if there exists an n^* for which $V_{n^*}(X^3)/V_{n^*}(X^1) = [(U(X^3) - U^2)/(U(X^1) - U^2)]^{1/n} > \pi$. Since $\lim_{n \rightarrow \infty} [(U(X^3) - U^2)/(U(X^1) - U^2)]^{1/n} = 1$ and since $\pi < 1$, the existence of the

required n^* follows.

Q.E.D.

Proof of Theorem: Given α , let (C^i, Y^i) and (C^j, Y^j) be the solutions to the deterministic problem. Since $\lambda_j > 0$ and $MRS^j(C^j(\alpha), Y^j(\alpha)) \neq 1$, there must exist two bundles (C^{j1}, Y^{j1}) and (C^{j2}, Y^{j2}) with $v^j(C^{j1}, Y^{j1}) = v^j(C^{j2}, Y^{j2}) = v^j(C^j(\alpha), Y^j(\alpha))$ and such that $C^{j1} - Y^{j1} > C^j(\alpha) - Y^j(\alpha) > C^{j2} - Y^{j2}$. Therefore, a $\hat{\pi}$ ($0 < \hat{\pi} < 1$), must exist with $\hat{\pi}(C^{j1} - Y^{j1}) + (1 - \hat{\pi})(C^{j2} - Y^{j2}) = C^j(\alpha) - Y^j(\alpha)$. $v^i(C^{j2}, Y^{j2}) > v^i(C^j(\alpha), Y^j(\alpha)) > v^i(C^{j1}, Y^{j1})$ since j 's indifference curve through $(C^j(\alpha), Y^j(\alpha))$ lies between i 's indifference curve and the 45° line through that bundle. (For the nonrandom solution, the possibility that $MRS^j(C^j(\alpha), Y^j(\alpha)) = MRS^i(C^j(\alpha), Y^j(\alpha)) = 1$ can be ruled out, in contrast to Theorem II(v); see Brito et al. [1990].) See Figure 5. From Lemma I, there exists a transformation of v^i such that $\hat{\pi}F(v^i(C^{j1}, Y^{j1})) + F(v^i(C^j(\alpha), Y^j(\alpha))) = F(v^i(C^i(\alpha), Y^i(\alpha)))$. Since $v^j(C^{j1}, Y^{j1}) = v^j(C^{j2}, Y^{j2}) = v^j(C^j(\alpha), Y^j(\alpha))$, $\hat{\pi}v^j(C^{j1}, Y^{j1}) + (1 - \hat{\pi})v^j(C^{j2}, Y^{j2}) > v^j(C^i(\alpha), Y^i(\alpha))$ so the self-selection constraint for j continues to be satisfied. The self-selection constraint for i is now satisfied with strict inequality. Hence, all constraints hold. If $1 > MRS^j(C^j(\alpha), Y^j(\alpha)) > MRS^i(C^j(\alpha), Y^j(\alpha))$, the bundles (C^{j1}, Y^{j1}) and (C^{j2}, Y^{j2}) can be chosen close enough to $(C^j(\alpha), Y^j(\alpha))$ so that $1 > MRS^j(C^{j1}, Y^{j1})$ and $1 > MRS^j(C^{j2}, Y^{j2})$. Then consider the bundles $(C^{j1} + \delta_1, Y^{j1} + \delta_1)$ and $(C^{j2} + \delta_2, Y^{j2} + \delta_2)$. δ_1 and δ_2 are chosen so that $v^j(C^{j1} + \delta_1, Y^{j1} + \delta_1) = v^j(C^{j2} + \delta_2, Y^{j2} + \delta_2)$. These continue to satisfy resource balance. For small enough $\delta_1, \delta_2 > 0$, it will still follow that $\hat{\pi}F(v^i(C^{j1} + \delta_1, Y^{j1} + \delta_1)) + (1 - \hat{\pi})F(v^i(C^{j2} + \delta_2, Y^{j2} + \delta_2)) < F(v^i(C^i(\alpha), Y^i(\alpha)))$ so i 's self-selection constraint still holds. Since $1 > MRS^j$ at both bundles, $v^j(C^j + \delta_1, Y^j + \delta_1) > v^j(C^{j1}, Y^{j1})$ and $v^j(C^j + \delta_2, Y^j + \delta_2) > v^j(C^{j2}, Y^{j2})$. Hence, the lottery $((C^{j1} + \delta_1, Y^{j1} + \delta_1), (C^{j2} + \delta_2, Y^{j2} + \delta_2), \hat{\pi})$ is

feasible, raises j 's welfare and thus improves upon the nonrandom solution $(C^j(\alpha), Y^j(\alpha))$. If $1 < MRS^j(C^j(\alpha), Y^j(\alpha)) < MRS^i(C^j(\alpha), Y^j(\alpha))$ a similar proof follows with $(C^{j1} - \delta_1, Y^{j1} - \delta_1)$ and $(C^{j2} - \delta_2, Y^{j2} - \delta_2)$ forming the lottery improving on $(C^j(\alpha), Y^j(\alpha))$. Q.E.D.

PROOF OF THEOREM V: Since this proof is complicated (and tedious), we begin by presenting a map of the basic steps in it. The proof is developed in five parts. Part (1) develops conditions that are necessary and sufficient for an arbitrary lottery to improve on the deterministic solution. The lottery must satisfy self-selection, not violate resource balance, and raise B 's expected utility. These conditions are given in (A1), (A4) and (A3). Part (2) specifies the nature of an improving local randomization by specifying three differentiable paths through the deterministic solution each parametrized by the same arbitrary variable t . These paths are then differentiated twice with respect to t to derive necessary and sufficient conditions for a local randomization to be beneficial, that is, to satisfy the requirements in (A1), (A4), and (A3). The local version of those conditions is given in (A13), (A14), and (A15). In part (3), these conditions are algebraically combined to determine a single inequality (A20) which is necessary and sufficient for the existence of an improving local randomization. A crucial step in this derivation is the recognition that the second derivatives of the bundles along the path are not directly restricted. This condition (A20) still involves first derivatives of the three paths with respect to t . In part (4), necessity of the condition in Theorem V is proven by showing that it is implied by (A20). Finally, in part (5), sufficiency of the condition in Theorem V is proven by showing that condition implies (A20).

(1) Conditions for Beneficial Randomization:

From Theorem III, $\lambda_a > 0$ implies $\lambda_b = 0$ and so it is not desirable to randomize A's tax schedule. Furthermore, the optimal randomization to offer B need involve at most three tax schedules. This result carries over to finding lotteries which improve on the nonrandom solution even if the optimal lottery is not found. That is, if a lottery over $k > 3$ bundles is better than the nonrandom solution, a lottery over just 3 bundles must exist which also does better. Hence, we can restrict the analysis to lotteries (π_h, C^{bh}, Y^{bh}) , $h = 1, 2, 3$, where $\sum \pi_h = 1$. Let (C^a, Y^a, C^b, Y^b) denote the nonrandom solution.

A lottery exists which improves upon (C^a, Y^a, C^b, Y^b) if, holding (C^a, Y^a) fixed:

$$V^a(C^b, Y^b) \geq \sum \pi_h V^a(C^{bh}, Y^{bh}) \quad (A1)$$

$$N_a(C^a - Y^a) + N_b \sum \pi_h (C^{bh} - Y^{bh}) \leq 0 \quad (A2)$$

$$\sum \pi_h V^b(C^{bh}, Y^{bh}) \geq V^b(C^b, Y^b) \quad (A3)$$

and at least one of (A1) to (A3) holds with strict inequality. (A1) and (A2) guarantee the lottery is feasible while (A3) says that B's expected utility does not decline. (A1) can be replaced by an equality since, if it held with strict inequality, a mean preserving contraction of the lottery would continue to satisfy (A2) and would raise $\sum \pi_h V^b(C^{bh}, Y^{bh})$ leaving (A3) satisfied. Since (A1) is assumed to hold with strict inequality, a small rise in $\sum \pi_h V^a(C^{bh}, Y^{bh})$ would not violate it. Of course, such a shrinkage would cause (A1) to hold with equality before all randomness was eliminated, or else there would exist a nonrandom vector better than (C^a, Y^a, C^b, Y^b) , a contradiction. If (A2) holds with strict inequality, C^a could be raised to get an improvement while, if (A3) holds with strict inequality, the improvement is direct. Since

$N_a(C^a - Y^a) + N_b(C^b - Y^b) = 0$, equation (A2) can be rewritten as

$$\sum \pi_h (C^{bh} - Y^{bh}) - C^b + Y^b \leq 0 \quad (A4)$$

(2) Description and Derivation of Beneficial Local Randomization

Consider any probability vector (π_1, π_2, π_3) . A local randomization is a path $(C^h(t), Y^h(t))$, $h=1,2,3$ with $(C^h(0), Y^h(0)) = (C^b, Y^b)$, all h , and such that for all $t > 0$ at least two bundles with nonzero probabilities differ from each other. From (A1), (A2), and (A4), a local randomization is improving for this probability vector if a path exists such that for all small $t > 0$:

$$\sum \pi_h V^a(C^h(t), Y^h(t)) = V^a(C^b, Y^b) \quad (A5)$$

$$\sum \pi_h (Y^h(t) - C^h(t)) \geq Y^b - C^b \quad (A6)$$

$$\sum \pi_h V^b(C^h(t), Y^h(t)) \geq V^b(C^b, Y^b) \quad (A7)$$

Differentiating (A5)-(A7) with respect to t around $t = 0$ and recalling that

$MRS^i = -V_y^i/V_c^i$ yields:

$$\sum \pi_h \left[\frac{dC^h(0)}{dt} - MRS^a(C^b, Y^b) \frac{dY^h(0)}{dt} \right] = 0 \quad (A8)$$

$$\sum \pi_h \left[\frac{dY^h(0)}{dt} - \frac{dC^h(0)}{dt} \right] \geq 0 \quad (A9)$$

$$\sum \pi_h \left[\frac{dC^h(0)}{dt} - MRS^b(C^b, Y^b) \frac{dY^h(0)}{dt} \right] \geq 0 \quad (A10)$$

Substituting (A8) into (A9) and (A10) yields:

$$[1 - MRS^a(C^b, Y^b)] \sum \pi_h \frac{dY^h(0)}{dt} \geq 0 \quad (A11)$$

$$[MRS^a(c^b, y^b) - MRS^b(c^b, y^b)] \Sigma \pi_h \frac{dy^h(0)}{dt} \geq 0 \quad (A12)$$

From (2), $1 - MRS^a(c^b, y^b)$ and $MRS^a(c^b, y^b) - MRS^b(c^b, y^b)$ must have

opposite signs so that (A11) and (A12) can both hold if and only if

$\Sigma \pi_h \frac{dy^h(0)}{dt} = 0$ which from (A8) implies that $\Sigma \pi_h \frac{dC^h(0)}{dt} = 0$. Thus, the first

order effects along an improving path must be zero. The gain to randomization

must come from second order effects. If either the second derivatives of net

resources or of B's expected utility are positive, then, since the first

derivatives are zero at $t = 0$, for $t > 0$ the first derivatives, and hence the

functions, will become positive as required. These second derivatives at

$t = 0$ are:

$$\Sigma \pi_h S^{ha} + \Sigma \pi_h [v_c^a(c^b, y^b) \frac{d^2 C^h(0)}{dt^2} + v_y^a(c^b, y^b) \frac{d^2 y^h(0)}{dt^2}] = 0 \quad (A13)$$

$$\Sigma \pi_h \left[\frac{d^2 y^h(0)}{dt^2} - \frac{d^2 C^h(0)}{dt^2} \right] \geq 0 \quad (A14)$$

$$\Sigma \pi_h S^{bh} + \Sigma \pi_h [v_c^b(c^b, y^b) \frac{d^2 C^h(0)}{dt^2} + v_y^b(c^b, y^b) \frac{d^2 y^h(0)}{dt^2}] \geq 0 \quad (A15)$$

where $S^{ih} = \left(\frac{dC^h(0)}{dt} \right)^2 v_{cc}^i(c^b, y^b) + 2 \frac{dy^h(0)}{dt} \frac{dC^h(0)}{dt} v_{cy}^i(c^b, y^b) + \left(\frac{dy^h(0)}{dt} \right)^2 v_{yy}^i(c^b, y^b)$. From the fact that the utility functions are strictly concave,

for any h with $dC^h(0)/dt$ or $dy^h(0)/dt$ not zero, S^{ih} must be negative. For

convenience, the arguments of partials of V^i are deleted since all partials are evaluated at (c^b, y^b) .

(3) A Necessary and Sufficient Condition for Beneficial Local Randomization

Substituting (A13) into (A14) and (A15) yields

$$Z \equiv (1 - MRS^a) \Sigma \pi_h \frac{d^2 y^h(0)}{dt^2} + \Sigma \pi_h S^{ah} / v_c^a \geq 0 \quad (A16)$$

$$[MRS^a - MRS^b] \Sigma \pi_h \frac{d^2 y^h(0)}{dt^2} + \Sigma \pi_h [(S^{bh}/V_c^b) - (S^{ah}/V_c^a)] \geq 0 \quad (A17)$$

Solving (A16) for $\Sigma \pi_h \frac{d^2 y^h(0)}{dt^2}$ and substituting into (A17) gives:

$$\frac{(MRS^b - 1)}{(1 - MRS^a)} (\Sigma \pi_h S^{ah})/V_c^a + (\Sigma \pi_h S^{bh})/V_c^b \geq \frac{(MRS^b - MRS^a)}{(1 - MRS^a)} Z \quad (A18)$$

From (2), $1 - MRS^a$, $1 - MRS^b$, and $MRS^b - MRS^a$ all have the same sign. Hence, (A18) can be rewritten as

$$\frac{\Sigma \pi_h S^{bh}}{V_c^b |1 - MRS^b|} - \frac{\Sigma \pi_h S^{ah}}{V_c^a |1 - MRS^a|} \geq \frac{|MRS^b - MRS^a|}{|1 - MRS^a| |1 - MRS^b|} Z \geq 0 \quad (A19)$$

Given any (π_1, π_2, π_3) , a path $(C^h(t), Y^h(t))$, $h = 1, 2, 3$, yields a local randomization which improves on the nonrandom solution iff $\frac{dC^h(0)}{dt}$, $\frac{dY^h(0)}{dt}$, $\frac{d^2 C^h(0)}{dt^2}$ and $\frac{d^2 Y^h(0)}{dt^2}$ are such that $\Sigma \pi_h \frac{dC^h(0)}{dt} = \Sigma \pi_h \frac{dY^h(0)}{dt} = 0$ but, for some h , $(\frac{dC^h(0)}{dt}, \frac{dY^h(0)}{dt}) \neq 0$, and (A13), (A16), and (A19) hold with a strict inequality in (A16) or (A19). Note that since $(1 - MRS^a) \neq 0$ and $d^2 Y^h(0)/dt^2$ are otherwise unrestricted, for any $(dC^h(0)/dt, dY^h(0)/dt)$, $h = 1, 2, 3$, $d^2 Y^h(0)/dt^2$, $h = 1, 2, 3$, can be chosen to make Z defined in (A16) take any nonnegative value. Given this choice of $d^2 Y^h(0)/dt^2$, $h = 1, 2, 3$, $d^2 C^h(0)/dt^2$, $h = 1, 2, 3$, can always be chosen to make (A13) hold. It therefore follows that local randomization is desirable iff there exists a $(dC^h(0)/dt, dY^h(0)/dt)$, $h = 1, 2, 3$, such that

$$\frac{\Sigma \pi_h S^{bh}}{V_c^b |1 - MRS^b|} > \frac{\Sigma \pi_h S^{ah}}{V_c^a |1 - MRS^a|} \quad (A20)$$

Necessity of (A20) follows directly from (A19). Sufficiency of (A20) holds since, using (A16), $d^2Y^h(0)/dt^2$ can be chosen to make Z arbitrarily small so that (A19) holds.

(4) Necessity of the Condition in Theorem V:

Note that $S^{bh} = qH^b(C^b, Y^b)q^t$ and $S^{ah} = qH^a(C^b, Y^b)q^t$ where $q = (dC^h(0)/dt, dY^h(0)/dt)$ and q^t is the transpose of q . If (A20) holds, then for at least one h , $(qH^bq^t)/(V_c^b|1-MRS^b|) > qH^aq^t/(V_c^a|1-MRS^a|)$ must hold. Cross multiplying by $V_c^b|1-MRS^b|$ and substituting $\theta = V_c^b|1-MRS^b|/(V_c^a|1-MRS^a|)$ gives $qH^bq^t > \theta qH^aq^t$ or $q(H^b - \theta H^a)q^t > 0$. This implies that $H^b - \theta H^a$ cannot be negative semidefinite as asserted.

(5) Sufficiency of the Condition in Theorem V:

If $H^b - \theta H^a$ is not negative semidefinite, then there exists a $\hat{q} \in R^2$ such that $\hat{q}(H^b - \theta H^a)\hat{q}^t > 0$. Multiplying out and substituting for θ yields $(\hat{q}H^b\hat{q}^t)/(V_c^b|1-MRS^b|) > (\hat{q}H^a\hat{q}^t)/(V_c^a|1-MRS^a|)$.

Consider any allowable π_1, π_2 , and π_3 (that is, $\pi_i \geq 0$, all i and $\sum_i \pi_i = 1$). Since at least two of these must be strictly positive, without loss of generality, we can relabel so that $\pi_1 > 0$ and $\pi_2 > 0$. To show sufficiency, we will show there exists at least two different vectors, the sum of whose quadratic forms with $H^b - \theta H^a$ equals that for \hat{q} . Thus (A20) is satisfied. To see this, consider the following five equations in the variables $q^h = (q_1^h, q_2^h)$, $h=1, 2, 3$:

$$\pi_1 q_1^1 + \pi_2 q_1^2 + \pi_3 q_1^3 = 0 \tag{A21}$$

$$\pi_1 q_2^1 + \pi_2 q_2^2 + \pi_3 q_2^3 = 0 \tag{A22}$$

$$\pi_1(q_1^1)^2 + \pi_2(q_1^2)^2 + \pi_3(q_1^3)^2 = (\hat{q}_1)^2 \quad (A23)$$

$$\pi_1(q_2^1)^2 + \pi_2(q_2^2)^2 + \pi_3(q_2^3)^2 = (\hat{q}_2)^2 \quad (A24)$$

$$\pi_1 q_1^1 q_2^1 + \pi_2 q_1^2 q_2^2 + \pi_3 q_1^3 q_2^3 = \hat{q}_1 \hat{q}_2 \quad (A25)$$

First, to show that (A21)-(A25) have at least one solution, consider the following values, $q_1^3 = q_2^3 = 0$, $q_1^2 = -(\pi_1/\pi_2)q_1^1 = [\pi_1/(\pi_2(\pi_1 + \pi_2))]^{\frac{1}{2}}\hat{q}_1$ and $q_2^2 = -(\pi_1/\pi_2)q_2^1 = [\pi_1/(\pi_2(\pi_1 + \pi_2))]^{\frac{1}{2}}\hat{q}_2$. Substituting shows these are in fact a solution.

To conclude, we show that any solution to (A21)-(A25) defines an allowable path that satisfies (A20). For any solution, let $(dC^h(0)/dt) = dy^h(0)/dt = q^h$, $h = 1, 2, 3$. From (A21) and (A22), $\Sigma \pi_h (dC^h(0)/dt) = \Sigma \pi_h (dy^h(0)/dt) = 0$ as required. From (A23) to (A25), $\Sigma \pi_h (q^h H^b (q^h)^t) = \hat{q}^H \hat{q}^t$ and $\Sigma \pi_h (q^h H^a (q^h)^t) = \hat{q}^H \hat{q}^t$. Substituting these into the assumption that $\hat{q}^H \hat{q}^t > \theta \hat{q}^H \hat{q}^t$ yields (A20) as required. Q.E.D.

PROOF OF COROLLARY I:

Let $Q(P_a, P_b) = |1 - MRS^a| \frac{q^H b q^t}{V_c^b} - |1 - MRS^b| \frac{q^H a q^t}{V_c^a}$. From Theorem V, local randomization is desirable at some $P_a \neq P_b$, if $Q(P_a, P_b) > 0$ for some q . At $P_a = P_b$, no redistribution is possible since the indifference maps are identical. Hence, at the deterministic optimum $MRS^a = MRS^b = 1$. Hence $Q(P_a, P_b) = 0$. Differentiating Q with respect to P_a around $P_a = P_b = P$ yields

$$\frac{dQ(P, P)}{dP_a} = \frac{d|1 - MRS^a|}{dP_a} \frac{q^H b q^t}{V_c^b} - \frac{d|1 - MRS^b|}{dP_a} \frac{q^H a q^t}{V_c^a} \quad (A26)$$

Since $v^a = F(v^b)$, $v_c^a = F'v_c^b$, $v_{cc}^a = F'v_{cc}^b + F''(v_c^b)^2$, $v_{yy}^a = F'v_{yy}^b + F''(v_y^b)^2$, and $v_{cy}^a = F'v_{cy}^b + F''v_c^b v_y^b$. When $P_a = P_b$, substituting these into $(qH^a q^t)/v_c^a$ yields:

$$\frac{qH^a q^t}{v_c^a} = \frac{qH^b q^t}{v_c^b} + \frac{v_c^b F''}{F'} (q_1 - q_2)^2 \quad (A27)$$

Substituting (A27) into (A26) yields

$$\begin{aligned} \frac{dQ(P, P)}{dP_a} &= \frac{qH^b q^t}{v_c^b} \left[\frac{d|1 - MRS^a|}{dP_a} - \frac{d|1 - MRS^b|}{dP_a} \right] \\ &\quad - \frac{v_c^b F''}{F'} (q_1 - q_2)^2 \frac{d|1 - MRS^a|}{dP_a} \end{aligned} \quad (A28)$$

For there to be redistribution toward B's with $\lambda_a > 0$, from Theorem III(v), $dMRS^a/dP_a$ and $dMRS^b/dP_a$ must have the same sign and $dMRS^a/dP_a$ must be larger in magnitude. Note that

$dMRS^b/dP_a = (\partial MRS^b/\partial c^b)(dc^b/dP_a) + (\partial MRS^b/\partial y^b)(dy^b/dP_a)$ and since P_a is not an argument of MRS^b , around $P_a = P_b$,

$$\frac{dMRS^a}{dP_a} = \frac{dMRS^b}{dP_a} + \frac{\partial MRS^a}{\partial P_a} \quad (A29)$$

The discussion above showed that $dMRS^a/dP_a$, $dMRS^b/dP_a$, and $\partial MRS^a/\partial P_a$ must all have the same sign. Given this, (A28) becomes:

$$\frac{dQ(P, P)}{dP_a} = \frac{qH^b q^t}{v_c^b} \left| \frac{\partial MRS^a}{\partial P_a} \right| - \frac{v_c^b F''}{F'} (q_1 - q_2)^2 \left| \frac{dMRS^a}{dP_a} \right| \quad (A30)$$

For F'' near zero, $dQ(P, P)/dP_a < 0$ must hold showing the undesirability of

randomization.

Q.E.D.

PROOF OF COROLLARY II: (i) Assume $\lambda_a > 0$. Since both groups share identical preferences over consumption, there is no benefit from any consumption randomization for B. To see this, assume there exists a feasible randomization for B over the three bundles: (C^{b1}, Y^{b1}) , (C^{b2}, Y^{b2}) , and (C^{b3}, Y^{b3}) . Let $\bar{C} = \pi_{b1}C^{b1} + \pi_{b2}C^{b2} + \pi_{b3}C^{b3}$. Consider the lottery $(\bar{C} - \delta, Y^{b1})$, $(\bar{C} - \delta, Y^{b2})$, $(\bar{C} - \delta, Y^{b3})$ where δ solves $\psi(\bar{C} - \delta) = \sum \pi^{bh}\psi(C^{bh})$. Note that $\delta \geq 0$ holds since ψ is concave. This lottery is also feasible since it satisfies A's self-selection constraint. B's expected utility is the same as in the original lottery. The extra resources (δ) can be used to give both A and B higher utilities than in the original lottery. Thus, consumption randomization is suboptimal for these preferences. Let \bar{C}^b be the consumption level in all bundles of the optimal lottery. Then (A1)-(A3) become:

$$\sum \pi_{bh} \bar{x}(Y^{bh}/w_a) = \psi(\bar{C}_b) - \psi(C^b) + \bar{x}(Y^b/w_a) \quad (A31)$$

$$\sum \pi_{bh} Y^{bh} \geq \bar{C}^b - C^b + Y^b \quad (A32)$$

$$\sum \pi_{bh} \bar{x}(Y^{bh}/w_b) < \psi(\bar{C}^b) - \psi(C^b) + \bar{x}(Y^b/w_b) \quad (A33)$$

For (A32) and (A33) to hold, $\bar{C}^b > C^b$ must hold. [To see this, assume $\bar{C}^b \leq C^b$ and let $\bar{Y} = \sum \pi_{bh} Y^{bh}$. From convexity of \bar{x} , $\bar{x}(\bar{Y}/w_b) \leq \sum \pi_{bh} \bar{x}(Y^{bh}/w_b)$. Since $\bar{Y} \geq \bar{C}^b - C^b + Y^b$, then $\psi(C^b) - \bar{x}(Y^b/w_b) \geq \psi(\bar{C}^b) - \bar{x}(\bar{Y}/w_b)$ follows from $1 > MRS^{bh}(C^b, Y^b)$ (which in turn follows from Theorem III(v) and the single crossing property of MRS^a and MRS^b for this utility function). (See Figure 6). These two inequalities violate (A33).]

Consider $M(w) = \sum \pi_{bh} \bar{x}(Y^{bh}/w) - \bar{x}(Y^b/w)$ where from (A31), $M(w_a) = \psi(\bar{C}_b) - \psi(C^b)$. Equation (A33) will hold iff $M(w_b) < M(w_a)$. Since $w_b < w_a$, a sufficient condition for this never to hold is $\partial M/\partial w \leq 0$ for all w

between w_a and w_b . Differentiating yields

$$\partial M/\partial w = -[\Sigma \pi_{bh} y^{bh} \gamma'(y^{bh}/w) - y^b \gamma'(y^b/w)]/w^2 \quad (A34)$$

so that $\partial M/\partial w \leq 0$ follows if

$$\Sigma \pi_{bh} y^{bh} \gamma'(y^{bh}/w) - y^b \gamma'(y^b/w) \geq 0 \quad (A35)$$

Adding and subtracting $\bar{Y} \gamma'(\bar{Y}/w)$ to (A35) yields

$$[\Sigma \pi_{bh} y^{bh} \gamma'(y^{bh}/w) - \bar{Y} \gamma'(\bar{Y}/w)] + [\bar{Y} \gamma'(\bar{Y}/w) - y^b \gamma'(y^b/w)] \geq 0 \quad (A36)$$

Since $\bar{C}^b > C^b$ as shown above, it follows from (A32) that $\bar{Y} > Y^b$. From convexity of γ , $\gamma'' \geq 0$ so that $\bar{Y} \gamma'(\bar{Y}/w) - y^b \gamma'(y^b/w) \geq 0$. Hence, a sufficient condition for (A36) is

$$\Sigma \pi_{bh} y^{bh} \gamma'(y^{bh}/w) - \bar{Y} \gamma'(\bar{Y}/w) \geq 0 \quad (A37)$$

Condition (A37) in turn follows if the function $Y \gamma'(Y/w)$ is convex in Y at all values of Y . Differentiating $Y \gamma'$ twice yields the condition:

$$2\gamma''(T/w)/w + Y \gamma'''(Y/w)/w^2 \geq 0 \quad (A38)$$

Substituting $L = Y/w$ shows that the condition in (i) is sufficient for $M(w_b) \geq M(w_a)$ with (A33) always violated and thus that randomization can never yield an improvement if it holds.

(ii) Since $\Sigma \pi_h(dC^h(0)/dt) = 0$ was shown in the proof of Theorem VI and since there is no consumption randomness here as shown above, $dC/dt = 0$ must hold. Then the condition in Theorem V reduces to

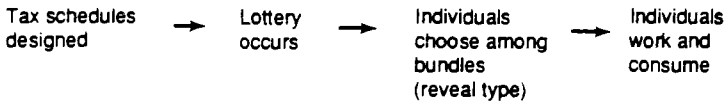
$$\frac{v_{yy}^b}{|1 - MRS^b|} > \frac{v_{yy}^a}{|1 - MRS^a|}$$

Substituting for V_{yy}^b and V_{yy}^a and the fact that $1 > MRS^b > MRS^a$ gives the condition in (3).

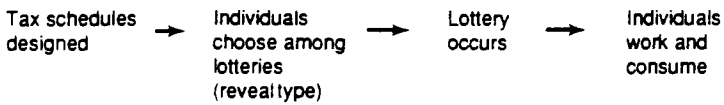
The proofs for $\lambda_b > 0$ follows similarly recognizing that if $\lambda_b > 0$ then (2) implies that $MRS^b(C^a, Y^a) > MRS^b(C^a, \bar{Y}^a) > 1$. To prove (i), it now follows that $\bar{C}^a < C^a$ must hold which then implies $\bar{Y}^a < Y^a$. The result is that $Y^a(Y/w)$ concave means randomization cannot occur. The proof of condition (4) also follows from Theorem V as in (3), reversing the A's and B's in the condition of Theorem V. Q.E.D

ALTERNATIVE TIMINGS OF RANDOMIZATION

PRE-REVELATION PRE-EFFORT RANDOMIZATION



POST-REVELATION PRE-EFFORT RANDOMIZATION



POST-REVELATION POST-EFFORT RANDOMIZATION

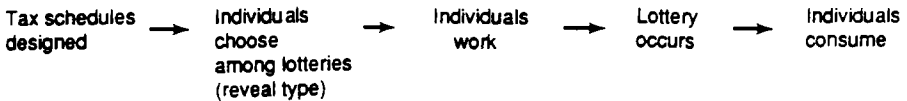


FIGURE 1

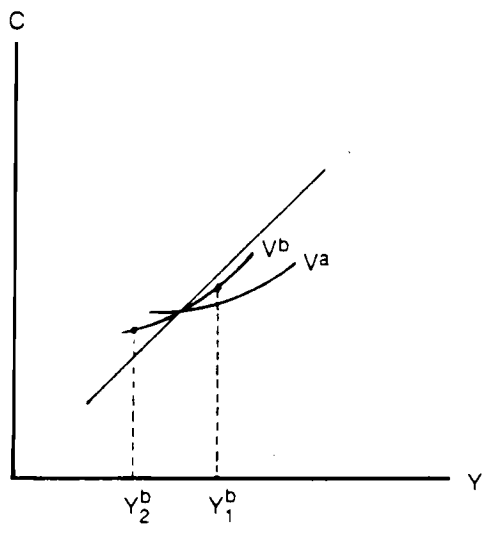


FIGURE 2

This randomization along B's indifference curve maintains B's expected utility, but it lowers A's expected utility from B's bundle when A is very risk averse.

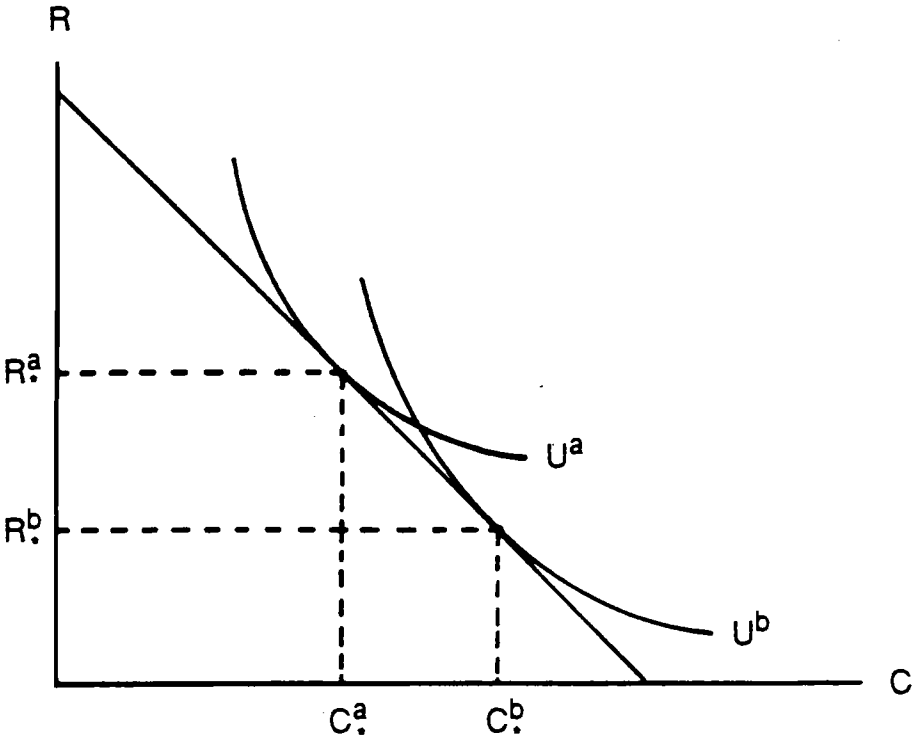


FIGURE 3
THE COMPETITIVE EQUILIBRIUM

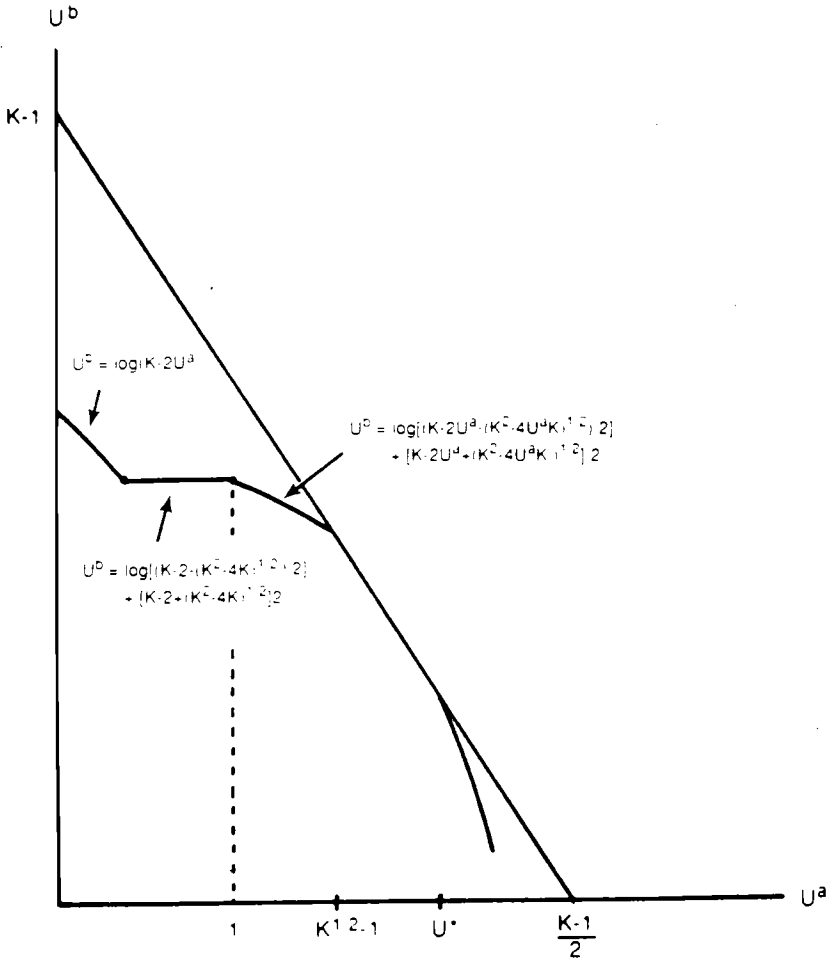


FIGURE 4

THE UTILITY POSSIBILITY FRONTIER

The outer line is the first best utility possibility frontier. The inner lines are segments of the second best utility possibility frontier without randomization for $K > 4$.

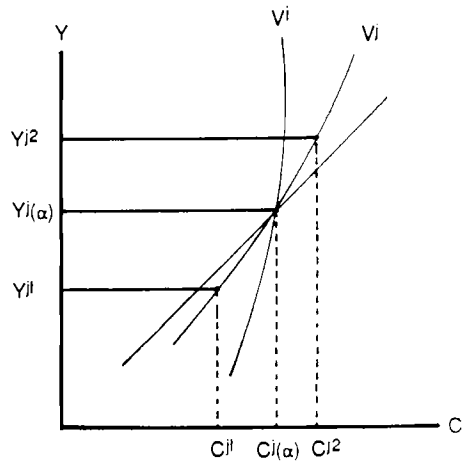
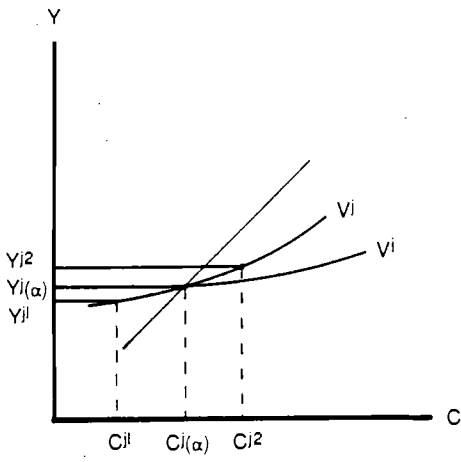


FIGURE 5

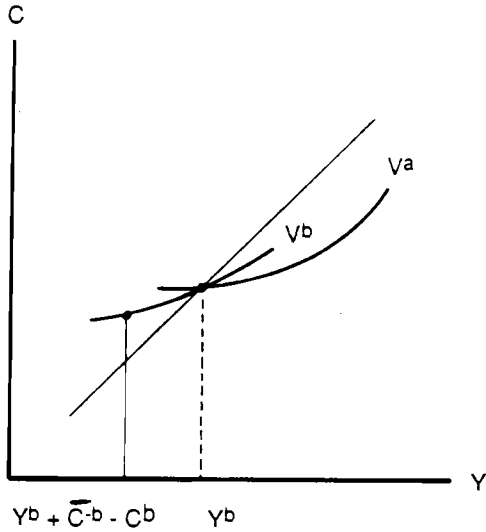


FIGURE 6

If $\bar{C}^b < C^b$, then $1 > MRS^b(C^b, Y^b) > MRS^a(C^b, Y^b)$ implies that $v^b(\bar{C}^b, \bar{Y}^b) < v^b(C^b, Y^b)$ where $\bar{Y}^b > Y^b + \bar{C}^b - C^b$.