POVERTY, INEQUALITY AND REDISTRIBUTION UNDER LEXICOGRAPHIC SOCIAL WELFARE

Unlike aversion to *inequality*, aversion to *poverty* resists formalization in welfare economics. One way to assign normative significance to the poverty line is to allow the welfare measure to exhibit a discrete loss from poverty (DLP) at z. However, the resulting redistribution scheme prioritizes headcount-reducing transfers to the borderline poor over transfers to the very poorest, rendering the DLP measure unattractive.

The paper remedies this by transcending the conventional real valued welfare measure. It proposes a lexicographic L^*-ordering, where the first rank criterion corresponds to an inequality-based evaluation function, while the second rank criterion corresponds to an evaluation function that exhibits DLP. The redistribution scheme entails transfers to the poorest until the first rank criterion is satisficed; only then may transfers be allocated to the borderline poor. The model's parameters can represent varying degrees of concern for the poorest, highlighting its flexibility as a framework for welfare evaluation.

1. INTRODUCTION

The redistribution of incomes is motivated by society's aversion to inequality and poverty. The formal derivation of an optimal redistribution scheme however reveals some difficulties associated with combining these two concepts. Inequality is typically measured in terms of a strictly convex function of incomes of each person in society. The method for measuring poverty is strikingly different: however the poverty indicator is finally computed, the first step involves classification of persons as poor or nonpoor, based on an income threshold z. The threshold identifies a fixed standard of living that just permits the satisfaction of basic needs. Hence, while inequality is essentially a relative concept, poverty involves a notion of being deprived in some absolute sense.

The measurement of inequality fits snugly within conventional welfare economics, in which social well-being is represented by a real-valued, continuous function of individual utilities. Such functions are informative regarding ordinal rankings of income distributions; no special meaning is assigned to a particular income threshold. Within this framework, a distinct concern for the poor is difficult to formalize.

A promising approach proposed in the literature is to allow a discontinuity at z in the social welfare function. That is, welfare may drop discretely as y_i, the income of person i, falls to just below z. Such a discrete drop does introduce a special meaning to the poverty line. Unfortunately, a welfare function with discrete loss from poverty (DLP) yields an unpleasant implication for the optimal redistribution scheme. If the least poor person is just a small amount ε short of the poverty line, then a transfer of ε to this person is preferred to a transfer of ε to the poorest person, even if the latter's income is close to zero. This clashes with widely held intuitions about the treatment of inequality and deprivation. As Ravallion (1994) observes: “If one starts instead from the value judgment that (subject to information and incentive constraints) the *poorest* in terms of the agreed welfare measure (y_i / z) should always get the highest priority, then jumps are ruled out (p. 1330).”

Thus the dilemma tackled by this paper: DLP welfare is a useful and perhaps indispensable device for understanding poverty; however introducing a welfare jump at z within the standard real-valued framework leads to prioritization of the least poor over the poorest in the transfer scheme.

To resolve this dilemma, this paper proposes a lexicographic model of choice. This framework contrasts with standard welfare theory, which orders income vectors according to a single criterion of social well-being. The model developed here orders income distributions according to a succession of criteria, implying a vector-valued welfare function. The flexibility of the approach allows us to formulate an ordering that incorporates a discrete loss from poverty, yet mandates that transfers be first allocated to the abject poorest.

The lexicographic approach has been suggested before in the poverty literature. In this paper it is employed for the first time to address the dilemma of modeling a discrete loss from poverty while prioritizing the most severely deprived. In so doing, the issues of poverty and inequality are linked in a coherent and fundamental way.

The discussion is structured as follows: Section 2 reviews previous studies on optimal redistribution given the dual concerns of poverty and inequality. Section 3 presents the lexicographic model and the resulting income redistribution scheme. Section 4 concludes.

2. REVIEW OF RELATED WORK

Since Atkinson (1970), economists have been sensitive to the welfare assumptions embedded in an inequality measure. One of the most important assumptions is the Dalton principle (more commonly referred to as the Pigou-Dalton transfer principle), which requires that inequality measure fall when a transfer is made from an upper to a lower part of a distribution. Sen (1976) ushered in a parallel scrutiny of poverty measures, likewise driven by the Dalton principle. The literature on inequality and poverty has however seldom converged completely on a common set of assumptions. As pointed out in the Introduction, the missing link is an adequate rationale for the poverty threshold.

2.1. Explaining the poverty line

Lewis and Ulph (1988) have drawn attention to this problem. They point out that standard welfare economics and normative public finance adopts a social welfare function whose arguments are the indirect utilities of persons. The welfare function is monotone increasing, symmetric, and strictly quasiconcave. The indirect utilities are determined by identical, continuous, differentiable, and concave functions. "Within such a theory there is *nothing* in the consumption behavior of consumers nor in the construction of individual and social welfare functions that gives one particular level of income the characteristics and significance the poverty line has in the poverty literature" (p. 119, italics supplied.)

To fill this gap, Lewis and Ulph lay a microeconomic basis for the poverty line. They hypothesize the existence of a discrete good called *inclusion*, acquired by purchasing a basket of commodities that costs a minimum of z. For example, to enjoy inclusion, a person may be required to observe dietary norms, where compliance entails some minimum expense. Consumption with inclusion always yields higher utility than consumption without inclusion. Along with a few other assumptions, they obtain an indirect utility function that falls continuously as income falls up to z, where it drops discretely downwards by a constant δ. This represents a discrete loss in utility due to a failure to consume inclusion.

From the discontinuous indirect utility function, a DLP welfare function can be derived with the usual assumptions; the welfare function can be shown to take the form $W(\bar{y}) - I - \delta H$, where W is a strictly convex, additively separable function of individual incomes, $\bar{y}$ is mean income, I is an inequality measure and H is the headcount ratio or poverty incidence. The δH term introduces discontinuity in the social welfare function. Hence, in evaluating optimal redistribution, transfers to the poorest (given fixed $\bar{y}$) must be traded off with the potential gains from reducing the poverty incidence.

2.2. The dilemma for optimal antipoverty transfers

Bourguignon and Fields (1997) derive optimal transfers to the poor when a poverty measure (which captures society's loss from poverty) exhibits the DLP feature. Let P be such a poverty measure, which is additively decomposable, nonincreasing in the incomes of the poor, and falls discretely by an amount δ whenever an erstwhile poor reaches the poverty line. They show that P can be decomposed into two parts as follows:

(1) $P = P_C + \delta H$

where P_C is a poverty measure sharing the properties of *P* except that it is continuous at *z*.

The resulting transfer scheme exhibits the dilemma confronting DLP measures. Let person 1 and person *D* with incomes y_1 and y_D respectively denote the poorest and the least poor person. Suppose y_D can be made arbitrarily close to *z* while y_1 is some fixed amount close to zero. Consider a transfer of the amount $z - y_D$; without loss of generality, let $y_2 - y_1 > z - y_D$. Then there is a y_D such that a transfer of the amount $z - y_D$ should go to person *D* rather than to person 1, however severe the latter's deprivation. The reason is that as y_D approaches *z*, the benefit from transferring $z - y_D$ to the poorest person approaches zero, whereas the benefit from making the same transfers to *D* remains fixed at δ. The DLP poverty measure is therefore rendered unattractive by this implication for optimal transfers.

2.3. Relative versus absolute poverty lines

The basis for the foregoing argument is the existence of a positive δ. One may object that the inclusion argument for δ applies only when the poverty line is deemed absolute, whereas in practice the calculation of poverty lines may be based on a measure of central tendency of an income distribution, and are therefore *relative*. We avoid wading here into a relative-versus-absolute lines debate; Madden (2001) provides an good overview of the issues and advocates a generalized poverty line that includes both relative and absolute components. Instead, we note that the problem of prioritizing the least poor remains even if *z* includes a relative component. For example, the fact that *z* rises with a rise in mean income may be irrelevant to a model that assumes a mean-preserving redistribution.

2.4. A lexicographic approach

The generalized representation of choice is the lexicographic model, which encompasses both continuous and discontinuous preferences (Chipman, 1960). A lexicographic approach to poverty reduction is however seldom encountered in the literature. The most prominent mention is in Atkinson (1987). He proposes two evaluation functions: the first involves the poverty measure *P*, while the second involves a welfare measure $\bar{y} - I$. The lexicographic scheme involves minimizing *P*; and, where P ranks two distributions equally, maximizing $\bar{y} - I$. This suggestion is taken up by Creedy (1997), who examines the case in which incentive problems complicate the derivation of an optimal redistribution scheme.

Consider however the benchmark case in which redistribution involves no deadweight burden, and the transfer budget is limited. Suppose (as usual) *P* and *I* are additively separable measures, with *I* subject to the Dalton principle, though *P* need not be. In fact, Atkinson asserts that *P* may as well be *H*, under a reasoning that is similar to the inclusion argument of Lewis and Ulph. If so then the optimal redistribution scheme is simple: the first criterion requires all antipoverty transfers to be devoted solely to lifting the most number of the least poor up to *z*. The second criterion is then relevant for allocating the remainder of the antipoverty budget, as well as in levying contributions from the nonpoor. Atkinson’s lexicographic model therefore advocates assisting the least poor ahead of the poorest.

In modeling lexicographic choice for poverty reduction, we need not however be confined to the Atkinson proposal. In the next section we provide an alternative lexicographic model that is consistent with a broad set of poverty and inequality norms; in particular, it incorporates a discrete loss from poverty, while allowing the most severely deprived to enjoy first priority in the redistribution scheme.

3. A LEXICOGRAPHIC MODEL OF WELFARE

The formal setting of the model is as follows: society's alternatives are found in Euclidean space R_+^N, whose points are denoted by a vector *y*. Each element of *y* denotes the income of person *h*, where $h \in^i [1, N]$ (alternative notation for h = 1, 2, …, *N*). In this section, we first present the preference structure in R_+^N, followed by the constraints facing the redistribution scheme. We then derive the optimal redistribution scheme and discuss its implications.

3.1. Lexicographic choice

Consider the following additively separable inequality measure *I*:

(2) $$I(y) = I(y_1, y_2, ..., y_N) = \frac{1}{N}\sum_{h=1}^{N} g(y_h),$$

where $g : R_+^1 \to R_{++}^1$. The function *g* is strictly decreasing in R_+^1, and is continuous, twice differentiable, and strictly convex in R_+^1. That is, *I* rises as income becomes more unequal given the same mean. Consider also a poverty measure *P* that takes the form:

(3) $$P(y) = P(y_1, y_2, ..., y_N) = \frac{1}{N}\sum_{h=1}^{N} L(y_h) + \delta H .$$

Following Bourguignon and Fields, *P* is decomposed into two parts as in equation (1). *L* has the same properties as *g*, while $\delta > 0$ is the discrete loss associated with poverty.

Symmetry of *I* and *P* allows the elements of *y* to be rearranged without disturbing the ordering in R_+^N, hence we focus only on *distributions*. Henceforth *y* denotes vectors whose elements are arranged in nondescending order; furthermore, *y* represents a mixed group of the poor and the nonpoor, i.e. there is a number *D* > 0 such that $y_h < z$, $h \in^i [1, D]$, and $y_h > z$, $h \in^i [D+1, M]$

To see the discontinuity at *z*, consider the function $\pi(y_h) = L(y_h) + d$, where *d* is defined as follows:

$$d = \begin{cases} \delta & \text{if } y_h < z \\ 0 & \text{otherwise} \end{cases}.$$

Then $\lim_{y_D \to z} \pi(y_h) = \delta$, but $\lim_{y_{D+1} \to z} \pi(y_h) = 0$, i.e. the left side and right side limits are unequal at *z*. Since $P = \frac{1}{N}\sum_{h=1}^{N} \pi(y_h)$, discontinuity of π carries over to *P*, making it a DLP measure.

The evaluation functions are as follows: Let $I^* \geq 0$ be some constant. For the first rank criterion, the corresponding evaluation function is $\tilde{I}$,where:

$$\tilde{I} = \begin{cases} -I, & I \geq I^* \\ -I^* & \text{otherwise} \end{cases}.$$

For the second rank criterion, the evaluation function is simply –*P*. The vector-valued objective function is denoted *V*, where $V(y)' = [\tilde{I}(y) \quad -P(y)]$. We say that y^1 is preferred to y^0 if and only if:

$$\tilde{I}(y^1) > \tilde{I}(y^0)\text{; or } \tilde{I}(y^1) = \tilde{I}(y^0) \text{ and } -P(y^1) > -P(y^0).$$

Within a constraint set *Y*, optimal *y* is the most preferred *y* based on a lexicographic comparison; with obvious notation, it represents a solution to the problem

(5) $$\text{Lex}\max_{y} V(y) \text{ s.t. } y \in Y.$$

The model proposed here follows the L^*-ordering suggested by Encarnación (1964).[1] That is, the first-ranked evaluation function $\tilde{I}$ reaches a maximum value $-I^*$, or its "satisficing" level, after which no further redistribution can further increase $\tilde{I}$. Poverty reduction then becomes the relevant choice criterion.

Our model clearly adopts some of the features of the Atkinson proposal, in that there are two criteria, one being inequality-based, and the other being poverty-based. Our model is distinguished by the following:

1. The inequality-based criterion is assigned the first rank, whereas in Atkinson's proposal it is assigned the second rank.
2. The inequality-based criterion is subject to a satisficing level I^*. The magnitude of I^* calibrates the intensity of society's aversion to "pure" inequality. At one extreme is $I^* = 0$, i.e. society is concerned only with inequality and not with poverty. As I^* rises, society increasingly values the alleviation of poverty.
3. *P* falls monotonically for incomes beyond *z*, whereas Atkinson's *P* (and poverty measures in general) are unaffected by the incomes of the nonpoor. (Strict monotonicity of *P* is not however essential to our results.)

[1] Encarnación and others impose continuity on the criterion-specific evaluation functions, contrary to the specification of *P*. Here we are simply following the example of Atkinson.

3.2. The redistribution scheme

The redistribution scheme corresponds $t \in R_{+}^{N}$, $\tau \in R_{+}^{N}$ such that to a $u't = u'\tau = B$ for the unit vector *u*. Vectors *t* and τ respectively correspond to transfers to recipients and taxes from contributors; *B* is a budget constraint where *B* > 0, and *B* is less than either $\sum_{h=1}^{D}(z - y_h)$ or $\sum_{h=D+1}^{M}(y_h - z)$, i.e. *B* does not eliminate either the poor or the nonpoor. The income vector with redistribution is $y + t - \tau$. Deadweight losses and incentive problems are suppressed in the analysis. We are interested only in net transfers and taxes, which alone can alter *y*; hence $t_h > 0$ implies $\tau_h = 0$, while $\tau_h > 0$ implies $t_h = 0$.

Using the terminology of Bourguignon and Fields, we identify the following special types of transfers and taxes:

DEFINITION. Let *k* and *l* be integers such that $1 \le k < l \le N$. Suppose for $h \in^{i} [k,l]$, $y_h + t_h = y_l + t_l$, and $t_h = 0$ otherwise. Then *t* is an *equalizing transfer* for $[k,l]$. Similarly, for $h \in^{i} [k,l]$, suppose $y_h - \tau_h = y_l - \tau_l$, and $\tau_h = 0$ otherwise. Then τ is an *equalizing tax* for $[k,l]$. If either $t_1 = B$ for $B < y_2 - y_1$, or *t* is an equalizing transfer for $[1, F]$, where $1 < F < D$, then *t* is a *p-transfer*. Similarly, if $t_D = B$ for $B < z - y_D$, or *t* is an equalizing transfer for $h \in^{i} [k, D]$ where *k* >1, then *t* is an *r-transfer*. If $\tau_N = B$ for $B < y_N - y_{N-1}$, or τ is an equalizing tax for $h \in^{i} [k, N]$, where *k* > *D*, then τ is an *r-tax*. We denote the p-transfer, r-transfer, and r-tax as, respectively, $t^{P}(B; y)$, $t^{r}(B; y)$, and $\tau^{r}(B; y)$; if there is no confusion we may suppress the *y* argument. Note for the *p*-transfer that the person *F* with the highest income but still receiving a positive transfer depends on *B*; we may therefore write $F(B)$. Finally, a *mixed transfer of B* (or simply a *mixed transfer*) is a *t* such that $t = t^p(B_1) + t^r(B_2)$ and $B_1 + B_2 = B$.

In short, the p-transfer allocates transfers to the poorest in an attempt to lift them up to a common income. The r-transfer allocates transfers to the leas poor so as to lift as many as possible up to the poverty line. A mixed transfer divides an antipoverty budget into two parts, one allocated as a p-transfer, another as an r-transfer.

3.3. Optimal redistribution

The optimal redistribution scheme can now be derived. We first characterize redistribution as optimized under the evaluation functions *I* and *P*, taken singly; the results are then combined to obtain the solution to the problem in (5) . As a preliminary, we state the following Lemma which merely reiterates Dalton's principle. A proof is reproduced here for completeness.

LEMMA 1. Consider a redistribution scheme *t*, τ such that given *k*, for all $h < k$, $t_k > 0$ implies $\tau_h = 0$ while $\tau_k > 0$ implies $t_h = 0$. Moreover, the elements of $\tilde{y}$ are still arranged in nondecreasing order. Then $I(\tilde{y}) < I(y)$.

PROOF. The Lemma hypothesizes a transfer of *B* from an upper part to a lower part of a distribution. Let $\theta > 0$ be the smallest change in income of any person. Distribution $\tilde{y}$ can be generated by a series of transfers of the amount θ from contributors to recipients yielding a series of vectors $y^1, y^2, ..., y^l, \tilde{y}$ (where the transfer that changes y^l to $\tilde{y}$ involves an amount no greater than θ. The foregoing vectors are denoted $k = 0,1,...l, l+1$ where $y^0 = y$ and $y^{l+1} = \tilde{y}$. To construct these vectors, the following procedure is adopted. For y^k let the transfer θ be from person *j* to person *h*, where *j* is the highest income person in y^{k-1} making a transfer, while *h* is the highest income recipient.[2] Then $h < j$, and $y_h^{k-1} < y_h^k \leq y_j^k < y_j^{k-1}$. For ease of notation, consider the transition from y to y^1. By the mean value theorem, there exists y_m^u, y_m^l such that $y_m^u \in (y_j, y_j - \theta)$, $y_m^l \in (y_h, y_h + \theta)$, $\theta g'(y_m^u) = g(y_j - \theta) - g(y_j)$, and $\theta g'(y_m^l)$ $= g(y_h) - g(y_h + \theta)$. Then $y_m^u > y_m^l$; by strict convexity of *g*, $g'(y_m^u) < g'(y_m^l)$. Hence upon rearrangement, $g(y_j - \theta) + g(y_h + \theta) <$ $g(y_j) - g(y_h)$, or $I(y^1) < I(y)$. Following the same argument, we can show that $I(y^2) < I(y^1), \ldots,$ $I(\tilde{y}) < I(y^{\ell})$, hence, $I(y) > I(\tilde{y})$.

The foregoing Lemma allows an easy proof for the optimal solution based on the evaluation function *I*.

PROPOSITION 1. The unique solution to the problem of minimizing $I(y + t - \tau)$ given *y* and *B* is $t = t^P(B; y)$, $\tau = \tau^R(B; y)$.

PROOF: Based on Lemma 1, the optimal redistribution scheme must be such that further transfers from an upper to a lower part of the distribution should no longer be feasible. Let $B > y_2 - y_1$, and $B > y_N - y_{N-1}$. Suppose $t_h > 0, h > 1$. If $y_h + t_h > (<)$ $y_{h-1} + t_{h-1}$, then there is an amount θ that can be transferred to *h* -1 rather than to *h* (*h* rather than to *h* - 1). This reduces *I*. Therefore we must have $y_h + t_h = y_{h-1} + t_{h-1} =$

---

[2] Vector inequality notation is as follows: $\forall h$, $y^1 \geq y^0$ iff $y_h^1 \geq y_h^0$; $y^1 > y^0$ iff $y^1 \geq y^0$ and $y^1 \neq y^0$; and $y^1 >> y^0$ iff $y_h^1 > y_h^0$.

$... = y_1 + t_1$. Similarly, for $\tau_j > 0$, if $y_j + \tau_j < (>) y_{j+1} + \tau_{j+1}$, then there is a θ that can be levied from *j* + 1 rather than from *j* (*j* rather than from *j* +1). This also reduces *I*. Therefore we must have $y_j + \tau_j = y_{j+1} + t_{j+1} = ... = y_N + t_N$. For the case in which $B < y_2 - y_1$, then only a transfer of *B* to person 1 makes further redistributions from upper to lower incomes impossible; similarly if $B < y_N - y_{N-1}$, then *B* should be levied from person y_N to make further redistribution impossible. The foregoing defines a p-transfer and an r-tax.

The following corollary is an immediate consequence of duality:

CORROLARY. Let $I^M = I(y + t^P(\overline{B}) - \tau^R(\overline{B}))$. The solution to the problem $\underset{t,\tau}{\text{Min}}\, B$ s.t. $I(y + t - \tau) \le I^M$ is given by $t^P(\overline{B}), \tau^R(\overline{B})$.

PROOF. Suppose there is a $\hat{B} \le \overline{B}$ for vectors $\hat{t}, \hat{\tau}$, $u'\hat{t} = u'\hat{\tau} = \hat{B}$ such that $I(y + \hat{t} - \hat{\tau}) \le I^M$. Since *g* is decreasing, any pattern of transfers of a positive difference $\overline{B} - \hat{B}$ is certain to reduce *I* below $\hat{I}$ and therefore below I^M; this is contrary to Lemma 1 which asserts that I^M is minimized with $t^P(B),\, \tau^R(B)$. We must therefore have $\hat{B} = \overline{B}$; by Proposition 1, $(t^P(B),\, \tau^R(B))$ uniquely achieves I^M, hence $(\hat{t}, \hat{\tau}) =$ $(t^P(B),\, \tau^R(B))$.

The dilemma associated with DLP measures arises when the least poor person or persons are "near" the poverty line. This notion of nearness is made precise in the following.

DEFINITION. For *P* in (6), let $G \in^i (1, D]$, $B = z - y_G$, where *B* is allocated either as a p-transfer or as a transfer to person *G*. If

$$\sum_{h=1}^{F(B)}[L(y_h)-L(y_h+t_h^P(B))]\le\delta\ ,$$

then person G is *proximate to z.* Alternatively, we may refer to G as *borderline poor.*

The definition states that a person is "near" the poverty line if the decline in P due to a transfer of B to G exceeds the decline in P due to a transfer of B to the poorest persons. The second Lemma obtains a necessary implication of proximity for persons indexed within the interval $[G, D]$.

LEMMA 2. If person G is proximate to z, $G < D$, then all poor persons with incomes higher than y_G are proximate to z.

PROOF. Without loss of generality, let $G+1<D$ and $y_{G+1} > y_G$. Define $B_0 = z - y_G,\quad B_1 = z - y_{G+1}$. The following is claimed:

$$\sum_{h=1}^{F(B_0)}[L(y_h)-L(y_h+t_h^P(B_0))] > \sum_{h=1}^{F(B_1)}[L(y_h)-L(y_h+t_h^P(B_1))].$$

To demonstrate the claim, observe that $B_0 > B_1$, $t^P(B_0) > t^P(B_1)$, and $F(B_0) \ge F(B_1)$. For the following expression,

$$\left\{\sum_{h=1}^{F(B_1)}L(y_h)-\sum_{h=1}^{F(B_1)}L(y_h)\right\}+\left\{\sum_{h=1}^{F(B_1)}L(y_h+t_h^P(B_1))-\sum_{h=1}^{F(B_1)}L(y_h+t_h^P(B_0))\right\}$$

$$+\left\{\sum_{h=F(B_1)+1}^{F(B_0)}L(y_h)-\sum_{h=F(B_1)+1}^{F(B_0)}L(y_h+t_h^P(B_0))\right\},$$

consider the terms in the curly braces. The first term is zero; the second term positive, since L is nondecreasing; and the third term is nonnegative. Hence the whole expression is positive. Rearranging the expression yields:

$$\sum_{h=1}^{F(B_0)}[L(y_h)-L(y_h+t_h^P(B_0))]-\sum_{h=1}^{F(B_1)}[L(y_h)-L(y_h+t_h^P(B_1))]\geq 0$$

This demonstrates the claim. The foregoing argument also holds for poor persons $G+1,\ G+2,\ \ldots,\ D$, if any.

The proximity property can hold however low the income of the poorest person is (as long as it is positive). More precisely, for an income vector *y*, there exists a y_D such that person *D* is proximate to *z*, as long as $y_1 > 0$. Consider the case where $z - y_D \geq y_2 - y_1$. We exploit the fact that $L'(y_1)$ is finite since twice differentiability of *L* in $(0,\infty)$ implies continuity of $L'(y_1)$, i.e. equality of its finite left and right side limits within that interval. As $\lim_{t_1\to 0}[L(y_1+t_1)-L(y_1)]/t_1 = L'(y_1)$, then $\lim_{t_1\to 0} L(y_1+t_1)-L(y_1) = \lim_{t_1\to 0} t_1 L'(y_1) = 0$. Since $\lim_{t_1\to 0} L(z-t_1)=\delta$, then there exists a $t_1 > 0$ such that and $L(y_1+t_1)-L(y_1) < L(z-t_1)$. For the case in which $z - y_D \leq y_2 - y_1$, then the foregoing argument must be modified to take into account the allocation of $z - y_D$ as a p-transfer.

The optimal distribution scheme implied by *P* can now be characterized:

PROPOSITION 2. The solution to the problem of minimizing $P(y+t-\tau)$ given *B* consists of $\tau = \tau(B)$ as well as:

i) $t = t^P(B)$ if no person is borderline poor.

ii) $t = t^R(B)$ if person *G* is borderline poor and $B = \sum_{h=G}^{D}(z-y_h)$.

iii) a mixed transfer of *B* if $\sum_{h=G-1}^{D}(z-y_h) < B < \sum_{h=G}^{D}(z-y_h)$, where person *G* is borderline poor.

iv) a mixed transfer if $B > \sum_{h=G}^{D}(z - y_h)$ and person *G* is has the lowest income among the borderline poor.

PROOF: In the following, Δx refers to the absolute value of the change in *x* due to *t* or τ. Observe that properties of *P* differ from those of *I* only in that *P* has a jump discontinuity at *z*. Additive separability implies that minimizing *P* can be divided into the problem of minimizing *P* by choice of τ and that of minimizing *P* by choice of *t*. For choice of τ the least increase in *P* should be selected. Any τ_h that sends a nonpoor person down to just below *z* will not incur the least increase in *P*; therefore, in levying contributions, ΔP equals ΔL and minimizing *L* minimizes *P*. This obtains at $\tau^R(B)$ based on Proposition 1.

As for choice of *t*, we note that maximum *P* deviates from maximum *L* only if ΔP differs from ΔL for some choice of *t*; by Proposition 1, maximum ΔL is achieved at $t^P(B)$. For poor person *h* and $t_h = z - y_h$, the resulting ΔP is δ. Now if no person is borderline poor, then ΔP cannot exceed ΔL. Hence $t^P(B)$ is optimum, establishing i). Now if *G* is borderline poor, then all persons indexed in the interval $[G, D]$ are borderline poor. Hence ΔP for $t^R(B)$ given $B = \sum_{h=G}^{D}(z - y_h)$ is $\delta(D - G)$. This exceeds ΔL for $t^P(B)$, establishing ii). Now if $\sum_{h=G-1}^{D}(z - y_h) < B < \sum_{h=G}^{D}(z - y_h)$ then $B - \sum_{h=G}^{D}(z - y_h)$ cannot be allocated as an r-transfer; hence $\sum_{h=G}^{D}(z - y_h)$ of *B* is allocated as an r-transfer (if *G* is borderline poor) and the remainder as a p-transfer. This establishes iii). Finally, if $B < \sum_{h=G}^{D}(z - y_h)$ where person *G* has the lowest income among the borderline poor, then $B-\sum_{h=G}^{D}(z-y_h)$ should be allocated as an r-transfer and the remainder as a p-transfer. This establishes iv).

The results from Propositions 1 and 2 may now be combined to characterize the optimal redistribution scheme, i.e. the solution to (5).

PROPOSITION 3. Given *B*, *y*, suppose B_1 is such that $I(y+t^P(B_1)-\tau^R(B_1))$ $=I^*$. Let $B_2=B-B_1$ and $\tilde{y}=y+t^P(B_1)+\tau^R(B_1)$. Then the optimal redistribution scheme consists of:

i) $(t^P(B),\tau^R(B))$ if $B\leq B_1$.

ii) the solution to the problem of minimizing $P(\tilde{y}+t-\tau)$ given transfer budget B_2 if $B>B_1$.

PROOF: If $B\leq B_1$, then I^* is not feasible given *B* by the Corollary to Proposition 1; then the solution to (5) is equivalent to minimizing *I*; by Proposition 1, this is given by $(t^P(B),\tau^R(B))$. This establishes i). If $B>B_1$, then I^* is feasible. Solving (5) is equivalent to minimizing *P* subject to the constraint that $I\leq I^*$. First we note that *P* is strictly decreasing in any y_h , hence a higher budget for minimizing *P* leads to a lower *P*. Thus the budget for achieving *I* should be minimized, which, by Corollary 2, is precisely the definition of B_1 . Now the solution of minimizing *P* does not solve for (5) only if that solution violates the constraint $I\leq I^*$. However, by Lemma 1, the solution to minimizing *P*, as described in Proposition 2, cannot increase *I* above I^*. This establishes ii).

Discussion

The reason for reversing Atkinson's priorities can now be explained. *P* captures concern for poverty, which at the margin motivates transfers. *The criterion sensitive to poverty is therefore less important than the criterion sensitive to inequality*. As Georgescu-Roegen (1954) has pointed out:

> Choice aims at satisfying the greatest number of wants starting with the most important and going down their hierarchy. Therefore choice is determined by the least important want that could be reached. This is why when we ask for the reason of choice we get answers which seem prima facie silly. An individual may give as the reason for which he bought a particular car 'the nice emerald green color of the panel'; another would say that he bought his house because 'it offered a nice location for a bird house.' But what both individuals mean is that after eliminating all available cars, and all available houses, by comparing them from the point of view of other more important wants they gradually came down to the color of the panel and to the bird house.

In the model, the more important social criterion is inequality-based, but two distributions are assessed the same way by this criterion once I^* is reached – equivalently, inequality ceases to be an important basis for choice. Hence the size of I^* determines the degree of society's commitment to prioritize the poorest persons, as well as to wipe out extreme deprivation. Once I^* is reached, society then examines another aspect of welfare, which is poverty. Then, and only then, does it become possible to consider the borderline poor.

The model we propose is very flexible: If a society takes the extreme position that the lowest income persons must always be favored in the redistribution scheme, then in effect it is setting I^* equal to zero. If it takes the other extreme position that pure inequality does not matter and antipoverty transfers should be devoted to getting as many of the poor out of poverty, it is in effect selecting an arbitrarily large I^* and δ. By tuning the parameters I^* and δ, we can represent a gradation of norms regarding the treatment of inequality and deprivation. Among these intermediate norms (corresponding to some value of I^* and δ), is the one that mandates that transfers must be given first to the poorest whose deprivation is especially severe. Hence the dilemma associated with DLP measures is resolved.

4. CONCLUSION

Why is there a special concern for the poor? As a person's income falls from above a poverty line to just below it, she undergoes a discrete switch in status (from nonpoor to poor). Society is concerned with this change in status, over and above the fact that her income is declining. A discrete welfare loss at the poverty line captures this concern very well. Such drops however are ruled out in conventional representations of welfare in terms of real-valued, continuous functions of individual well-being. What such functions do capture, under the appropriate curvature restrictions, is society's concern for inequality rather than poverty.

Simply incorporating a DLP feature in the real-valued welfare function however faces a problem: under this formulation, society would value a reduction of the poverty headcount over assistance to the most severely deprived as long as the least poor are closely clustered around the poverty line. This exposes the dilemma associated with attempts to attach formal normative significance to the poverty line.

This paper resolves the dilemma by adopting a more flexible framework of choice. The framework is a lexicographic model in which social welfare is evaluated over a succession of criteria. In this paper's specification, there are two criteria, one which is inequality-based, another which is poverty-based. The measure for the former criterion exhibits the properties of the standard welfare function, while the measure for the latter criterion exhibits the DLP feature. The inequality-based measure is assigned the first rank, though subject to a maximum (satisficing) value. Under this model, the well-being of the most severely deprived can take precedence over the reduction of the headcount ratio even in the presence of the borderline poor. The degree to which the poorest are valued by society can be represented by the model's parameters (i.e. the magnitudes of the satisficing value and of the DLP.)

We therefore show that normative welfare analysis can be considerably enriched if it transcends the confines of the real-valued objective function. With acceptance of multiple criteria by which to evaluate social welfare, formalization of poverty-related discontinuities becomes tractable and even appealing.

REFERENCES

Atkinson, A., 1971. On the measurement of inequality. *Journal of Economic Theory* 2:244-263.

Atkinson, A., 1987. On the measurement of poverty. *Econometrica* 55:749-64.

Bourguignon, F., and G. Fields, 1997. Discontinuous losses from poverty, generalized P_α-measures, and optimal transfers to the poor. *Journal of Public Economics* 63:155-75.

Chipman, J., 1960. The foundations of utility. *Econometrica* 28:193-224.

Creedy, J., 1997. Evaluating alternative tax and transfer schemes with endogenous earnings. *Oxford Economic Papers* 49:43-56.

Encarnación, J., 1964. A note on lexicographical preferences. *Econometrica* 32:215-17.

Foster, J. and A. Shorrocks. (1991). Subgroup consistent poverty indices. *Econometrica* 59: 687-209.

Georgescu-Roegen, N. (1954). Choice, expectations and measurability. *Quarterly Journal of Economics* 58:503-534.

Lewis, G. and D. Ulph, 1988. Poverty, inequality, and welfare. *Economic Journal* 98: 117-31.

Madden, D. (2000). Relative or absolute lines: a new approach. *Review of Income and - Wealth* 46:181-99.

Ravallion, M., 1996. Issues in measuring and modeling poverty. *Economic Journal* 106:1328-43.