AMARTYA SEN

RAWLS VERSUS BENTHAM: AN AXIOMATIC EXAMINATION OF THE PURE DISTRIBUTION PROBLEM*

ABSTRACT. The decision rules yielded respectively by the Rawlsian 'maximin' conception of justice and by classical utilitarianism are compared and contrasted. The discussion is based on the assumption of a pure distribution problem and sharp differences are brought out. An axiomatic analysis of the two conceptions is undertaken, the result of which is that Rawls and utilitarianism both omit essential aspects of distributional welfare judgments: Rawls leaves out questions of welfare differences, utilitarianism leaves out questions of welfare levels. It is possible to pay attention to the ranking of welfare levels without concentrating exclusively on the welfare levels of worst off persons only, thereby departing from both Bentham and Rawls.

I. INTRODUCTION

In this paper I would like to compare and contrast the decision rules yielded respectively by the Rawlsian 'maximin' conception of justice[1] and by classical utilitarianism. Much of the discussion will take place in the context of a pure distribution problem, typified by the exercise of justly dividing a cake among n persons, which brings out some of the differences sharply.

In Section II a set of axioms is presented which the various choice rules may be expected to follow, and it is examined which of these axioms are satisfied respectively by the Utilitarian, the Rawlsian and other choice rules. The presentation in Section II is informal, but the axioms are more formally stated in Section III, in which the results presented in Section II are fitted with proofs. If the reader is bored by formalities, he can easily move from Section II directly to Section IV, where the results are discussed again in completely informal terms.

II. CHOICE RULES AND AXIOMS OF RANKING

There are n people rather austerely christened $1, \ldots, n$. There is a fixed homogeneous income (cake) to be distributed among them. Each likes more and more income but the gain from an additional unit goes down as he gets richer and richer. We take his welfare to be a function of his

Theory and Decision **4** (1974) 301–309. *All Rights Reserved*

own income only and it increases at a diminishing rate as he gets more and more. The problem is to rank all possible distributions of the cake according to some rule of choice.

The utilitarian rule (henceforth, UR) is to maximize the sum of individual welfares. The simplest version of the Rawlsian maximin rule (henceforth, MR) is to maximize the welfare level of the worst off person. The lexicographic version of the Rawlsian maximin rule (henceforth, LMR) is to follow MR, but if the worst off persons in two distributions are equally well off, then to maximize the welfare of the second worst off person. If the worst off persons are equally well off and so are the second worst off persons in two distributions, then maximize the welfare of the third worst off. And so on, under LMR.

Three axioms on rules of choice are now introduced.

The Symmetry Preference Axiom (SPA):

If everyone has the same welfare function, then any transfer from a richer man to a poorer man, which does not reverse the inequality, is always preferable.

The Weak Equity Axiom (WEA):

If person i is worse off than person j whenever i and j have the same income level, then no less income should be given to i than to j in the optimal solution of the pure distribution problem.[2]

The Joint Transfer Axiom (JTA):

It is possible to specify a situation in which j is [slightly] better off than k (the worst off person), and [strongly] worse off than i, such that some transfer from i to j [sufficiently large], even though combined with a simultaneous transfer [sufficiently small] from k to j, leads to a preferred state than in the absence of the two transfers.

The Symmetry Preference Axiom simply stands in favour of a reduction of inequality if the persons have identical 'needs'. The rationale of this is well discussed; see Kolm (1969). The Weak Equity Axiom demands that a person who is more deprived in non-income respects should not be made to receive less income as well. The Joint Transfer Axiom suggests that an inequality increasing transfer (from k to j) can be outweighed by a sufficiently large inequality decreasing transfer (from i to j). That is, some trade offs are permitted. The words in square brackets are not needed in the statement of JTA and have been included only to motivate the axiom. While SPA is concerned with single transfers, JTA is concerned with pairs of transfers.

The following results are true and are proved in Section III below.

(T.1) The Utilitarian Rule violates the Weak Equity Axiom for some set of permissible individual welfare functions.

(T.2) The Maximin Rule violates the Symmetry Preference Axiom and the Joint Transfer Axiom for some set of permissible individual welfare functions.

(T.3) The Lexicographic Maximin Rule can violate the Joint Transfer Axiom for some set of permissible individual welfare functions.

(T.4) There exist choice rules that can satisfy all three axioms (SPA, WEA and JTA) for all permissible individual welfare functions.

III. FORMAL PRESENTATION

The share of income of person i is y_i, for $i=1,\ldots,n$. The problem is to rank all vectors $\mathbf{y}$, i.e., $(y_1,\ldots,y_n)$, subject to:

$$(1) \qquad \sum_{i=1}^{n} y_i = Y > 0,$$

$$(2) \qquad \forall i: y_i \geqslant 0.$$

Person i's welfare W_i is a monotonically increasing and twice differentiable function of his income y_i and is strictly concave.

$$(3) \qquad W_i = W_i(y_i), \quad \text{with} \quad W_i' > 0 \quad \text{and} \quad W_i'' < 0.$$

The W_i functions can vary from person to person but are interpersonally fully comparable (see Sen, 1970, Chapter 7).

The following notation will be used in addition to standard symbols of algebra: $\rightarrow$ 'if-then'; $\leftrightarrow$ 'if and only if'; & 'and' (conjunction); $\vee$ 'or' (alternation); $\sim$ 'not' (negation); $\forall$ 'for all' (the universal quantifier); and $\exists$ 'for some' (the existential quantifier). Further, R is the binary relation of 'at least as good as', P that of 'better than', and I that of 'indifferent to'.

(4) $$\mathbf{xPy} \leftrightarrow [\mathbf{xRy} \;\&\; \sim \mathbf{yRx}].$$

(5) $$\mathbf{xIy} \leftrightarrow [\mathbf{xRy} \;\&\; \mathbf{yRx}].$$

UR states that:

(6) $$\mathbf{xRy} \leftrightarrow \sum_i W_i(x_i) \geqslant \sum_i W_i(y_i).$$

MR states that:

(7) $$\mathbf{xRy} \leftrightarrow \min_i W_i(x_i) \geqslant \min_i W_i(y_i).$$

In distribution **x**, call the worst off person $x1$, and generally the ith worst off person xi. (In case of ties in the poverty ranking, take the tied persons in either order.) Similarly, yi is the ith worst off person in distribution **y**.

LMR states that:

(8) $$\begin{aligned}\mathbf{xRy} \leftrightarrow [&\{W_{x1} > W_{y1}\} \vee \{W_{x1} = W_{y1} \;\&\; W_{x2} > W_{y2}\} \vee \ldots \vee \\ &\vee \{\forall i\colon i \leqslant n-2\colon W_{xi} = W_{yi} \;\&\; W_{x(n-1)} > W_{y(n-1)}\} \\ &\vee \{\forall i\colon i \leqslant n-1\colon W_{xi} = W_{yi} \;\&\; W_{xn} \geqslant W_{yn}\}]\end{aligned}$$

The axioms are now formally defined.

Symmetry Preference Axiom (SPA):

(9) $$\begin{aligned}[&\{\forall i, j, y\colon W_i(y) = W_j(y)\} \;\&\\ &\&\; \{y_i < x_i \leqslant x_j < y_j\} \;\&\; \{x_i - y_i = y_j - x_j\} \;\&\\ &\&\; \{\forall k \neq i, j\colon x_k = y_k\}] \rightarrow \mathbf{xPy}.\end{aligned}$$

Weak Equity Axiom (WEA):

(10) $$[\{\forall y\colon W_i(y) < W_j(y)\} \;\&\; \{\forall \mathbf{y}\colon \mathbf{xRy}\}] \rightarrow x_i \geqslant x_j.$$

Joint Transfer Axiom (JTA):

(11) $$\begin{aligned}\exists \mathbf{x}, \mathbf{y}, \delta_1, \delta_2\colon [&(y_i > y_j > y_k) \;\&\; (x_i \geqslant x_j \geqslant x_k) \;\&\\ &\&\; (\forall r\colon y_r \geqslant y_k) \;\&\; (\delta_1, \delta_2 > 0) \;\&\; (\forall r \neq i, j, k\colon x_r = y_r) \;\&\\ &\&\; (x_i = y_i - \delta_1) \;\&\; (x_k = y_k - \delta_2) \;\&\\ &\&\; (x_j = y_j + \delta_1 + \delta_2) \;\&\; \mathbf{xPy}].\end{aligned}$$

Now the proofs of (T.1)–(T.4).

Proof of (T.1). Consider $W_i(.) = mW_j(.)$ with $0<m<1$ and $W_i(.)$ always positive. If $\forall \mathbf{y}$: **xRy**, then under UR:

$$(12) \qquad \sum_i W_i(x) = \max_{\mathbf{y}} \sum_i W_i(y_i).$$

In view of the strict concavity and twice differentiability of each W_i, this implies that:

$$(13) \qquad W_j'(x_j) = W_i'(x_i) = mW_j'(x_i).$$

Since $m<1$, and $W_j''<0$, it must be the case that $x_j > x_i$. Thus UR violates WEA.

Proof of (T.2): Consider the antecedent in the statement of SPA, and take a case in which $\exists y_k: y_k < y_i$. Since $\forall k \neq i, j: x_k = y_k$, and $y_i < x_i \leqslant \leqslant x_j < y_j$, evidently $\min_i W_i(x_i) = \min_i W_i(y_i)$. Hence **xIy**.

Thus MR violates SPA.

Now consider JTA: if all conditions within the square brackets are satisfied except **xPy**, then clearly $\min_i W_i(x_i) \leqslant \min_i W_i(y_i)$. Hence **yRx** according to MR. But this rules out JTA.

Proof of (T.3): The same reasoning holds for LMR as in the case of MR in the latter part of the proof of (T.2).

Proof of (T.4): An example will suffice and many examples do exist. Consider the choice rule of minimizing that widely used measure of inequality, the Gini coefficient, which we know can be written as:[3]

$$(14) \qquad G = 1 + (1/n) - (2/nY)\,[ny_{y1} + (n-1)\,y_{y2} + \cdots + y_{yn}],$$

in which y_{yi}, as defined in the context of LMR, is the income of the ith worst off person in the distribution **y**.

This amounts to the choice rule of maximizing W given by:

$$(15) \qquad W = \sum_i (n+1-i)\,y_{yi}$$

The unique optimum is given by $y_i = y_j$ for all i, j. Evidently WEA is satisfied.

It is clear that SPA is satisfied, since:

$$(16) \qquad W(\mathbf{x}) - W(\mathbf{y}) > 0,$$

given the antecedent of SPA, as is obvious from (15), since the lower the income level the higher the weight on it. Finally, for JTA, consider the effect on W of a joint transfer which does not alter the ranking of poverty of the persons. Let $r(i)$, $r(j)$ and $r(k)$ be the worst off rank positions of i, j and k respectively. Obviously, $r(i)>r(j)>r(k)$.

$$(17) \qquad W(\mathbf{x}) - W(\mathbf{y}) = \delta_1 [r(i) - r(j)] - \delta_2 [r(j) - r(k)]$$

We can easily get $W(\mathbf{x})>W(\mathbf{y})$, by choosing:

$$(18) \qquad \delta_1 > \delta_2 [r(j) - r(k)]/[r(i) - r(j)]$$

IV. DISCUSSION

While neither classical Utilitarianism (UR) nor the Rawlsian maximin rules (MR and LMR) can satisfy all three of the axioms SPA, WEA and JTA, choice rules do exist that satisfy all three. In fact, as was shown, minimizing a standard measure of inequality, viz., the Gini coefficient, is such a rule. It can also be shown that minimizing some other common measures (e.g., the coefficient of variation) also satisfies SPA, WEA and JTA.

It may, however, be noted that the Gini coefficient, the coefficient of variation and other standard measures of inequality are defined independently of the individual welfare functions. While the statements of the conditionals required under WEA and SPA involve individual welfare functions, these references occur only in the respective antecedents, and the rules of minimizing standard measures of inequality such as the Gini coefficient satisfy WEA and SPA since they render the consequents true *irrespective* of the truth value of the antecedents concerned with individual welfare functions. It is easy to propose other axioms involving individual welfare functions such that an automatic fulfilment of the consequents by the choice rules yielded by the inequality measures will be ruled out.[4]

The uses made of interpersonal comparisons of welfare in the different approaches are worth contrasting. For this I shall use an analytical framework that I have presented elsewhere, omitting the formal structure, for which the readers are referred to Sen (1970), Chapters 7 and 7*.

Non-Comparability:

Welfare numbers of different persons cannot be compared in any way.

Level Comparability:

Welfare levels of different persons can be compared but not differences between levels of different persons. E.g., it makes sense to say that *A* is better off than *B*, but none to say that *A*'s welfare gain in moving from **x** to **y** is greater than *B*'s gain (or loss) from the movement.

Unit Comparability:

Welfare level differences of different persons can be compared but not the levels themselves.[5]

Full Comparability:

Both levels and differences can be compared.

What are the requirements of comparability of the three approaches discussed in the earlier sections?

(1) *Inequality measures (e.g., the Gini coefficient):*
None is needed.

(2) *Utilitarianism (UR):*
Unit comparability, or full comparability.

(3) *Rawlsian maximin rules (MR or LMR):*
Level comparability, or full comparability.[6]

A distinctive feature of the Rawlsian maximin rules is their concentration on welfare levels, whereas Utilitarianism concentrates on welfare differences only. If unit comparability holds, MR or LMR cannot even be formulated, but UR can be used, since all it requires is to sum the welfare differences of all the persons in moving from **x** to **y**. And **x** is preferred, or **y**, or both equally, according as the sum is positive, negative, or zero, respectively. On the other hand, it is easily seen that if level comparability holds, UR cannot be formulated, whereas MR or LMR can flourish. The real conflict arises only when full comparability is assumed, when MR, LMR and UR can all be used. This was the framework used in the earlier sections of this paper.

Critical comments on Rawls' maximin rules have mainly been concerned with their extreme nature concentrating only on the worst off individual and ignoring the rest (or the *k*th worst off, if the more worse off persons tie in the poverty scale, under LMR). This certainly is a significant aspect of Rawls' conception of justice, and this does differ sharply from other approaches such as utilitarianism. On the other hand, there is another, possibly more serious, aspect of the contrast between the two approaches, and this concerns the concentration on *levels* of welfare under the Rawlsian approach in contrast with the concentration on welfare *differences* in the utilitarian scheme of things. The contrast can be easily seen if it is asked: Under what circumstances should a transfer of income from person i to person j be recommended under the two approaches? Under UR, such a transfer should take place if and only if the welfare *gain* of j is greater than the welfare *loss* of i from the transfer. Under MR, it should take place if and only if i has a higher *level* of welfare than j, who is the worst off person (and i stands to gain *something* from the transfer, it does *not* matter how much). The extreme nature of the MR (or LMR) criterion in concentrating on the welfare level of the *worst off* person (or the kth worst off, in the case of ties, under LMR) can be removed in a more general approach, which could still retain the concentration on welfare *levels* as opposed to the exclusive concern with marginal gains or losses in the utilitarian approach. The Weak Equity Axiom is an example of a partial rule that is not extremist in the sense in which MR and LMR are, but which uses the same type of information as the Rawlsian criteria.

Finally, it is reasonable to argue that in making ethical judgements on distributional issues (and in other types of social choices as well), one is typically concerned *both* with comparisons of levels of welfare as well as with comparisons of welfare gains and losses. It is not surprising that the utilitarian approach and the maximin approach both run into some fairly straightforward difficulties, since each leaves out completely one of the two parts of the total picture. Given the powerful hold that utilitarianism has had on thinking on public policy for centuries, it is understandable, and in most ways entirely welcome, that Rawls has concentrated totally on the other half of the information set. But a more complete theory is yet to emerge.

London School of Economics

NOTES

* I have benefited from the comments of Partha Dasgupta.

[1] In this paper I shall not be concerned with the contractual conception of fairness developed by Rawls (1958), (1971), and the justification of the maximin rule in terms of choices in the 'original position'. I have tried to argue elsewhere (Sen, 1970, Chapter 9) that the contractual conception may be more readily acceptable than the maximin rule as such. See also Harsanyi (1955) and Pattanaik (1971).

[2] The Weak Equity Axiom was defined in Sen (1973) in a somewhat more demanding form, requiring that person i should receive *more* (and not merely, no less) income than j under the circumstances specified.

[3] For the derivation see Dasgupta *et al.* (1973), p. 186. Note that in that paper $Y = 1$, and G is taken to be the negative of the Gini coefficient, i.e. $-G$ here.
[4] In fact neither the Gini coefficient nor the coefficient of variation can satisfy the stricter form of WEA proposed in Sen (1973).
[5] Under level comparability, $W_i(\mathbf{x}) > W_j(\mathbf{y})$ makes sense, but not $[W_i(\mathbf{x}) - W_i(\mathbf{y})] > [W_j(\mathbf{y}) - W_j(\mathbf{x})]$. Under unit comparability, the latter makes sense but not the former. Note that level comparability holds if individual welfare functions are all 'ordinal' (order homomorphic to the real numbers) without being 'cardinal' (group homomorphic), but are entirely comparable. Unit comparability holds if individual welfare functions are all 'cardinal' and any change of unit of the numerical representation of the welfare function of one person must be combined with a change of unit in the same ratio for all persons, but an arbitrary constant can be added to any individual's welfare function without a similar constant being added to the welfare functions of others. On all this, see Sen (1970), Chapter 7*, which also discusses a continuum of 'partial comparability'.
[6] Some uses of UR, MR or LMR are possible even when the required type of comparability is only partial, but not under all circumstances; see Sen (1970), Chapters 7 and 7*.

BIBLIOGRAPHY

Dasgupta, P., Sen, A. K., and Starrett, D., 'Notes on the Measurement of Inequality', *Journal of Economic Theory*, **6** (1973).

Harsanyi, J., 'Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility', *Journal of Political Economy* **63** (1955).

Kolm, S. Ch., 'The Optimal Production of Social Justice', in J. Margolis and H. Guitton (eds.), *Public Economics*, Macmillan and St. Martin's Press, 1969.

Pattanaik, P. K., *Voting and Collective Choice*, Cambridge University Press, 1971.

Rawls, J., 'Justice as Fairness', *Philosophical Review*, **67** (1958).

Rawls, J., *A Theory of Justice*, Harvard U. Press and Clarendon Press, 1971.

Sen, A. K., *Collective Choice and Social Welfare*, Holden-Day and Oliver and Boyd, 1970.

Sen, A. K., *On Economic Inequality*, Clarendon Press and Norton, 1973.