Centro de Estudios Económicos
www.colmex.mx
El Colegio de México, A.C.

Serie documentos de trabajo

ON THE THEORY OF OPTIMAL CONVERGENCE

Joseph J. Persky and Mo-Yin S. Tam
University of Illinois

DOCUMENTO DE TRABAJO

Núm. III - 1982

On The Theory of Optimal Convergence

Joseph J. Persky
Mo-Yin S. Tam

Department of Economics
University of Illinois at Chicago

This material is based upon work supported by the National Science Foundation under Grant SES-8106955

The most basic axiom in the measurement of income inequality is that of anonymity or symmetry. All of the common measures of inequality take this point of view. Such an axiom necessarily implies the irrelevance of group membership. On an ethical basis it is difficult to fault the index makers for their devotion to this principle. Be that as it may, both the historical and political analysis of income equity issues are often couched in terms not of individuals but of groups identified as being rich or poor. Whatever our ethics might argue, there seems to be a natural tendency to discuss the issues of income equality in terms of units far larger than the individual.[1] In many cases this concern with groups is justified by the peculiar history of a particular segment of society. For example where discrimination against a group has been rampant one might be concerned with erasing the effects of that discrimination regardless of the implications for overall inequality. We explicitly rule out such situations here. Rather for the purposes of this paper the single important normative issue is assumed to be the overall degree of inequality as measured under the axiom of anonymity.

Even where the primary interest is the overall level of inequality there is a strong tendency in practice to use group statistics. This approach usually assumes that narrowing of mean differences between groups represents a gain in equity. We will call this general notion the mean-convergence approach.

As a general proposition, the statement that a reduction in mean differences between groups implies an improvement in social welfare or equity is indefensible. For example, in a recent note [1982] we have shown that for the United States the ongoing convergence of Southern family incomes to those of the non-South has had no significant effect on the overall level of inequality as measured by the Gini coefficient. Moreover, if that tendency toward convergence continues to the point of equal mean incomes it will actually produce a slight increase in national inequality as measured by the Gini coefficient. The explanation for this result is not difficult to appreciate. Bringing the group means closer together involves not only a redistribution to the lower tail of the poorer group, but also a redistribution to the upper tail of that group. The Gini index is quite sensitive to the upper tail as well as the lower one. If the poorer group has a much more unequal distribution than the richer group these shenanigans in the upper tail may actually lead to a rise in the overall index even before the more unequal, and poorer group has caught up with the richer one.

Clearly then in some cases there exists a potential conflict between the ethics of inequality indexes and the mean convergence approach. One possible way to deal with such a conflict is to constrain the measure of inequality in a manner which rationalizes an emphasis on mean convergence. Such an approach produces an ad hoc system of ethical judgments that is difficult to defend. The alternative developed here is to derive the inter-group statistics appropriate to given ethical judgments concerning inequality. Thus the basic purposes of this paper are to explore the circumstances under which an emphasis on reducing differences in mean incomes between groups is consistent with the anonymity principle, to develop a notion of optimal convergence for those situations where it is not, and to use that notion in defining a new measure of a group's ability to pay.

In the next section we formalize the notion of optimal convergence. Section III goes on to demonstrate for a broad class of inequality indexes necessary and sufficient conditions under which convergence in inter-group means leads to a reduction in general inequality. The fourth section then uses the notion of optimal convergence to derive analytical results for many common measures of inequality. This section also contains an empirical exercise based on the income distributions of the Southern United States and the rest of the country. The fifth section goes on to discuss the implications of these results, deriving a measure of a group's ability to pay.

II. Optimal Convergence

Consider the following situation. A given population of N individuals is divided into two arbitrary groups of individuals. Group A, consisting of N_A people, has mean income μ_A. Group B consists of N_B people having a mean income of μ_B. Our basic interest concerns making welfare judgments about the changes in inequality produced by the convergence of these two mean incomes. To keep matters simple assume that total income is held constant and thus any change in mean incomes is achieved by a transfer between the two groups.

At this point we face a major question: what assumption to make about how a transfer is collected in the donor group and how it is distributed in the recipient group. We call this mechanism an incremental distribution rule. At a highly general level such a rule is defined as follows:

Incremental Distribution Rule. Let T be a transfer of income from group B to group A. Then the incremental distribution rule H specifies two vector valued differential functions h_A and h_B such that

$$\dot{y}_A = h_A(y_A) > 0 \tag{1}$$

$$\text{and } \dot{y}_B = h_B(y_B) < 0$$

where $y_A = (y_{A1}, \ldots, y_{AN_A})$ is the vector of incomes in group A and $y_B = (y_{B1}, \ldots, y_{BN_B})$ is the vector of incomes in group B.

Note there is the implicit condition on the above definition that for each group the sum of the elements in the vector derivative must be one in absolute value. Moreover we explicitly limit consideration to H's such that incomes of all individuals in a group change in the same direction.

Obviously if one could choose any H at will, there would be little reason to identify the groups at all and the entire issue of inter-group transfers could be reduced to optimizing inter-individual transfers. Quite clearly our problem only becomes interesting when there exist constraints on what can be done as among members of a group. In any given situation this is fundamentally an empirical issue. The range of possibilities is infinite. Both transfers and taxes can take on either a progressive, proportional, or regressive form. In what follows, results will be presented at several levels of generality. At the most general level, we require, as stated above, only that the incomes of all members in a group move in the same direction. Somewhat more restrictive is the family of H's with the following form:

$$\dot{y}_{A_i} = \frac{(y_{Ai})^\alpha}{\sum (y_{Ai})^\alpha} \quad i=1,N_A$$

(1')

$$\dot{y}_{B_j} = \frac{-(y_{Bj})^\beta}{\sum (y_{Bj})^\beta} \quad j=1,N_B$$

This family allows a wide range of alternative rules depending on the values of the parameters α and β. If $\alpha > 1.0$ a transfer is regressive. If $\alpha = 1.0$ a transfer has a proportional effect. If $\alpha < 1.0$ a transfer is progressive. In this third case if α is zero a transfer takes a per capita form. Clearly for $\beta >$ (<) 1.0 a tax is progressive (regressive). Within this family we will pay particular attention to the case where $\alpha = \beta = 1.0$ and that where $\alpha = \beta = 0.0$, i.e., the case when a transfer or a tax is proportional and that when both are per capita. The first suggests itself on the grounds that incremental shares in many cases are likely to follow original shares; the second, because of simplicity.[2]

We can now define what we mean by optimal convergence:

Optimal Convergence. For two groups of individuals characterized by initial income vectors $\overset{o}{y}_A$ and $\overset{o}{y}_B$, given inequality measure I and an incremental distribution rule H, optimal convergence occurs at T^* if I for the entire population of N individuals achieves a minimum with respect to T at T^*.

In this definition a positive value for T^* implies that up to that point transfer from B to A reduces inequality, while any larger transfer under rule H will increase inequality. On the other hand a negative value of T^* tells us that to reduce inequality we must transfer some income from A to B.

The notion of optimal convergence specifically allows for the possibility that inequality for the union of the two groups may not be minimized when two means are equal. Where this is the case some inequality in mean incomes helps to minimize the overall inequality in the system. In any event the point of optimal convergence clearly demarcates the range of mean incomes for which transfers from B to A are desirable from that for which transfers from A to B are desirable.

Obviously the optimal convergence transfer T^* will depend on the three factors mentioned in its definition: the shapes of the initial distributions, the inequality index and the incremental distribution rule. In the next section we take a broad class of inequality indexes and demonstrate that for this class the optimal convergence point must lie within a well-defined range determined by y_A^o, y_B^o and H. This range will always include the point of equal mean income.

III. Conditions on the Underlying Distributions

Inequality indexes generally are derived from more basic ethical axioms about measuring inequality or social welfare. As already noted, the most common of these axioms is that of anonymity or symmetry with respect to individuals. A second particularly powerful axiom is the Pigou-Dalton principle: a transfer from a richer to a poorer individual should reduce inequality. As made clear in the work of Atkinson [1970], and that of Dasgupta, Sen and Starrett [1973] the basic Pigou-Dalton principle is broadly equivalent to the notion that given two distributions having the same mean, one should always prefer one of them to the other if its Lorenz curve everywhere dominates the other. Of course neither principle tells us how to rank distributions whose Lorenz curves cross, or what is equivalent, distributions that cannot be obtained from one another by a set of Pigou-Dalton transfers. Indeed when two Lorenz curves cross, we know that there will be two indexes which rank them differently while ranking all Lorenz dominant pairs the same.

If then we limit ourselves to the class of inequality measures that are Pigou-Dalton we can draw heavily on the work of Atkinson [1970], and that of Dasgupta, Sen and Starrett [1973], to delineate the range of the optimal transfer, T* from group B to group A. We can immediately state a trivial proposition:

Proposition 1. Given two groups A and B, with initial income vectors y_A^o and y_B^o, such that $y_{Ai}^o < y_{Bj}^o$ for all $i \in A$ and $j \in B$ then for any Pigou-Dalton I and any H satisfying (1), $T^* > 0$.

The proof of Proposition 1 is trivial and will not be given here. Quite simply the groups don't overlap and hence any redistribution between them must be equivalent to a set of Pigou-Dalton transfers from richer to poorer people.[3]

Starting from Proposition I it should be clear that if we introduce a slight overlap between two groups we must still be on safe ground. For example, if the richest member of one group is just a little richer than the poorest member of the other group, an inter-group transfer would still break down into a set of Pigou-Dalton transfers. How far can this overlap go and still keep us on safe ground with respect to all the indexes in our group? The answer depends on which specific incremental distribution rule we have.

Before we state the next proposition, it will be useful to have the following definition:

<u>Cumulative Transfer (Tax) share.</u> For a rule H and an income vector Y_A we define the cumulative transfer share for those with incomes below the scalar y^* as

$$Z(y_A, y^*) = \sum_{y_{A_i} < y^*} | \dot{v}_{A_i} |.$$

The function Z tells us for any rule H what proportion of an infinitesimal transfer to A will accrue to members of group A with incomes below y^* and what proportion of the corresponding infinitesimal tax applied to group B will be raised from members of that group with incomes below y^*. The following proposition states conditions under which income transfer from group B to A reduces inequality and those under which a transfer in the reverse direction reduces inequality.

<u>Proposition II.</u> If for the initial income vectors y_A^o and y_B^o

(2a) $Z(y_A^o, y^*) > Z(y_B^o, y^*)$ for all y^*

with strict inequality for at least one y^* then for any Pigou-Dalton I, $T^* < 0$. Alternatively if

(2b) $\quad Z(y_A^o, y^*) \leq Z(y_B^o, y^*)$ for all y^*

with strict inequality for at least one y^* then for any Pigou-Dalton I, $T^* < 0$. If neither (2a) nor (2b) holds then there exist two Pigou-Dalton indexes, I_1 and I_2 such that

$T_2^* \leq 0 \leq T_1^*$.

The proof of Proposition II rests directly on Dasgupta, Sen and Starrett [1973].[4] We must simply show that the condition (2a) or (2b) guarantees the existence of a Lorenz dominant relation between the before and after-transfer distribution of the entire population of N individuals. The absence of such dominance clearly implies the existence of two Pigou-Dalton indexes that conflict in their rankings. That condition (2a) or (2b) implies Lorenz dominance for very small discrete transfers is obvious once we realize that such transfers do not alter the rankings of individuals in the total population.

Now consider the more specific family of H's defined in (1'). In this case condition (2a) in Proposition II becomes

$$(2a') \quad \left| \frac{\sum_{y_{A_i}<y^*} y_{A_i}^{\alpha}}{\sum Y_{A_i}^{\alpha}} \right| > \left| \frac{\sum_{y_{B_j}<y^*} y_{B_j}^{\beta}}{\sum Y_{B_j}^{\beta}} \right|$$

for all y^* with strict inequality for at least one y^*, and similarly for condition (2b). Working then to even more specific and perhaps more intuitive cases, let us set $\alpha = \beta = 0$. This is the case of per capita incremental distribution rule. Here we obtain the following corollary:

Corollary IIa. For two groups A and B with cumulative income distributions F_A and F_B, if

(2") $F_A(y^*) \underset{(\leq)}{\geq} F_B(y^*)$

with strict inequality for at least one y^*,[5] then for the per capita distribution rule

$$T^* \underset{(<)}{>} 0.$$

for all Pigou-Dalton I. If neither of the above conditions (2") holds, there are two Pigou-Dalton indexes I_1 and I_2 such that $T_2^* < 0 < T_1^*$.

It is easy to show that condition (2") is a specific case of condition (2') with $\alpha = \beta = 0$. In this case the condition simply states that as we come up the income ladder we always have a higher (or lower) proportion of people in group A below us. An intuitive interpretation of this corollary can be given in the case where N_A is equal to N_B. Here the condition (2") is equivalent to stating that $Y_{Bi} < Y_{Ai}$ where i refers to the ith richest person in the particular group, $i = 1, \ldots, N_A = N_B$. In this case the Pigou-Dalton nature of the transfers is obvious.

Cleary the condition stated in corollary IIa will be neither sufficient nor necessary in the case where the incremental distribution rule is proportional rather than per capita i.e. $\alpha = \beta = 1$. Here information about ranking of individuals in the group is insufficient to determine the outcome of a transfer between the groups. One also requires information about the distance separating the income levels. That is, we need information about income shares. This result is clear in Corrollary IIb.

Corollary IIb. For two groups A and B with cumulative income shares S_A and S_B between the groups. One also requires information about the

where $$S_A(y^*) = \frac{\sum_{y_{Ai} \leq y^*} y_{Ai}}{N_A \mu_A}$$

and similarily for S_B, if

$$(2''') \qquad S_A(y^*) \underset{(\leq)}{\geq} S_B(y^*) \text{ for all } y^*$$

with strict inequality for at least one y^*, then for a proportional distribution rule

$$T^* \underset{(>)}{<} 0$$

for all Pigou-Dalton I. If neither of the above conditions (2''') holds there are two Pigou-Dalton indexes I_1 and I_2 such $T_2^* < 0 < T_1^*$.

Again it is clear that Corollary IIb is a straightforward inference of Proposition II. Condition (2''') is condition (2') with $\alpha = \beta = 1.0$. Proposition II and its corollaries underscore the basic conflict between the mean convergence approach and that of anonymous inequality measures. Clearly under the per capita and the proportional incremental distribution rules mean convergence does not imply an unambiguous reduction of inequality as measured by the set of Pigou-Dalton indexes. Indeed we can state this result more generally for all incremental distribution rules which satisfy the inequality conditions of equation (1).

<u>Corollary IIc</u>: There is no incremental distribution rule H, satisfying (1) for which $_B > \mu_A$ implies T* > 0 for all Pigou-Dalton inequality indexes.

The proof of this corollary is quite simple, requiring only the existence of a counter example. Suppose simply that the poorest person in the total population is in group B. By Proposition II it is clear that in this case T* for some I will be less than zero. In terms of the before and after transfer Lorenz Curves for the entire population, the very lowest portion of the latter curve must lie below the very lowest portion of the former in this counter example.

Hence, among the set of Pigou-Dalton inequality measures there is only a weak connection between mean convergence and a reduction in overall inequality. Under these circumstances it is natural to explore constraints on the set of indexes and incremental distribution rules which might guarantee less ambiguous results. On the other hand, if such constraints seem too restrictive it may be useful to explore in more detail alternatives to mean convergence. The next section lays the basis for such an exercise.

IV. <u>Optimal Convergence and Specific Inequality Measures</u>

A discussion of the conditions under which various indexes of inequality may be reduced by an inter-group transfer is ultimately dependent on the derivation of the optimal convergence points under different indexes and incremental rules. Table I presents in summary form first order conditions for minimizing several common measures of inequality under both per capita and proportional incremental distribution rules.[6] Starting with the per capita incremental scheme, only the variance (V) and the (square of the) coefficient of variation (V^*) are minimized at the point of equal means. For all the other measures, additional information about the intra-group distributions is generally required to find the minimum of total inequality.

Most of the resulting rules have straightforward interpretations. For example, the Gini coefficient (G), is particularly sensitive to people's rank in the overall distribution. The effect of per capita transfers from one group to another thus depends on the average rank of members in the two groups. As long as one group has a higher average rank regardless of its mean the Gini will respond to a redistribution. It is easy to see that this point of equal average ranks must occur in the range delimited by Corollary IIa, since first order statistical dominance by either group would necessarily imply a higher average rank. The relative mean absolute deviation (D) presents a similar optimal point. Here the issue is only the proportion of individuals in each group below the mean income. When these proportions are equal optimal convergence is achieved. This condition thus picks out one specific level of income, namely μ, to make the comparisons of condition (2"). If asked when a per capita transfer is appropriate this measure responds by looking at the value of the two cumulative distributions at the mean income of the entire population. A third index with a rather surprisingly simple condition for optimal convergence is Theil's measure of entropy (T). Here, the message is to continue the per capita redistribution until the geometric means of the two groups, $\hat{\mu}_A$ and $\hat{\mu}_B$ are equal. The Atkinson family of indexes (A) reach optimal convergence when the $(-\varepsilon)$-moment of the two groups are equal. This result has a particularly straight-forward interpretation if one is willing to follow Atkinson in identifying inequality measures with social welfare functions. The Atkinson indexes are in fact derived from the additive separable social welfare functions with the following form:

$$W = \sum_{i=1}^{N} [k_1 + \frac{k_2 y_i^{1-\varepsilon}}{1-\varepsilon}]$$

where k_1 and k_2 are constants. In this situation the $(-\varepsilon)$-moment for groups A & B are proportional to the average marginal utility of income in the two groups respectively. Hence for a per capita transfer and an additively separable social welfare function equating the $(-\varepsilon)$-moment of the two groups is obviously the appropriate condition for optimal convergence. For the particular case where $\varepsilon = 1$ the condition simplifies to equating the average income reciprocals.

All of the above indexes (other than V and V^*) present optimal convergence conditions that substitute an alternative income statistic for the mean income when reaching a conclusion as to the desirability of a transfer between groups. Where V and V^* implicitly analyze a transfer according to its effect on the difference in group means, these other indexes are looking at alternative measures which in one fashion or another involve higher moments of the two group distributions. In any of these cases the optimal convergence point could be on either side of the mean convergence point.

Not surprisingly the results presented for the proportional incremental distribution rule are quite different as to specifics. Here, the optimal convergence points for V and V^* are sensitive to higher moments of the intra-group distribution while they hadn't been under the per capita transfer scheme. In particular, optimal convergence now implies a higher mean for whichever group has the lower inequality as measured by V^*. Now the optimal convergence points under both G and D are influenced by income shares, where previously they were only dependent on ordinal aspects of the cumulative distributions. The index T has a particularly straightforward trade-off between differences in group incomes and differences in intra-group inequality as measured by the entropy index applied to each group individually. As in the case of V and V* the optimal convergence condition for T also implies a higher mean for the group with lower inequality. Similarly the variance of the logarithms (L) has an optimal convergence point that is sensitive to intra-group inequality. In this case, however, the optimal point gives a higher mean to the group with more inequality as measured by (C). Hence a compensating notion is apparant here.

In the proportional case the most interesting results are probably those concerning the Atkinson family. Again there is a straightforward interpretation of the optimal convergence condition in terms of the marginal utility of individual incomes.[7] More importantly the optimal convergence condition as stated in Table I provides a clear statement of trade-offs between differences in group means and differences in intra-group inequality. When $\varepsilon < 1$ optimal convergence implies a higher mean income for the group with lower inequality as measured by the Atkinson index itself. Interestingly when $\varepsilon > 1$ the relation reverses and the higher mean goes to the group with greater inequality. This latter case of $\varepsilon > 1$ is a situation where the underlying index is particularly sensitive to the lower tail of the distribution. Raising the mean of the more unequal group is then an attempt to guarantee that its lower tail doesn't extend too far. This view of intergroup transfers is clearly a compensatory one as opposed to that common to the variance, the coefficient of variation, Theil's entropy and Atkinson's index itself when $\varepsilon < 1.0$. Not surprisingly when $\varepsilon = 1.0$ the index simply ignores intra-group inequality in fixing the optimal convergence point by equating μA to μ_B.

The theoretical and ethical questions raised by these results are discussed in more detail in Section V. Before turning to those questions it may be useful to provide an empirical illustration that develops the theme of the last two sections. In this exercise we will only consider the proportional incremental rule. It should be noted that for all the common indexes in Table I except V, such a transfer leaves intra-group inequality unaffected. The exercise uses 1979 data on the distribution of family income in the South and the rest of the United States. In that year the mean income of Southern families was about 92% of the mean income of non-Southern families. Based on these data and simple simulations we can determine the range of optimal convergence for the Pigou-Dalton measures as specified by conditions 2'''. This range is shown in Figure I. For the class of Pigou-Dalton measures optimal convergence of the South and the North implies a Southern mean income between 88% and 152% of the Northern mean.

In Figure I the Southern mean relative to the non-Southern mean for optimal convergence is identified for each index.[8] The results here are clearly influenced by the fact that regardless of the measure used the South is the region with greater inequality (see Table II). As pointed out previously indexes such as V, V^* and T imply higher means for lower inequality at optimal convergence. Hence these indexes suggests an optimal μ_S/μ_N that stops short of full convergence. On the other hand, an index such as L, that has a compensating notion, suggests an optimal μ_S/μ_N greater than one. This dichotomy is also apparent for the family of Atkinson's indexes. Those with $\varepsilon < 1$ have the mean of South (the more unequal region) less than that of the non-South for optimal convergence. The reverse is the case for indexes with $\varepsilon > 1$. Of course, the index (C), (Atkinson's index with $\varepsilon = 1$) gives an optimal convergence point right at the point of equal means.

V. Ability to Pay and Welfare

The discussion of optimal convergence in the last two sections suggests a more formal statement of the mean-convergence approach. This view of inequality can be described as a particular extension of the Pigou-Dalton condition to the group level. More formally an advocate of the mean-convergence approach might justify that position by picking an inequality measure that meets the following condition:

Group Pigou-Dalton Condition. An index I is group Pigou-Dalton if a transfer from group B to group A reduces I whenever $\mu_B > \mu_A$.

It should be clear that there is no Pigou-Dalton index which can meet this condition for all incremental distribution rules. Even if we limit consideration to only the proportional distribution rule, for example, we are left with just one index in Table I that meets this condition, namely (C). This might suggest a special position for this index. That position is clearly related to the fact that C has a straightforward disaggregation property.[9] Be that as it may, we would suggest considerable caution before advocating the special attractiveness of (C). Even assuming that for many practical situations the most likely incremental distribution rule is a proportional one, the emphasis on mean group income in the definition of the group Pigou-Dalton condition implies that differences in intra-group distributions are irrelevant in determining need and ability to pay. For the familiar individual version of the Pigou-Dalton condition richer and poorer are intuitively defined. As soon as we move to groups, the notion of richer or poorer is far less clear. Especially where the groups are somewhat arbitrary and the poor may be lumped together with the rich for little good reason, it is hardly obvious that the mean income of the group is an appropriate measure of its affluence or its ability to pay. Why not the median income? Why not the minimum? If the Pigou-Dalton notion is extended by a definition of affluence other than the group mean, the set of inequality measures meeting the new condition must necessarily be different.

Using the notion of optimal convergence, we offer the following definition of a group's ability to pay.

Relative Ability to Pay: Given initial income vectors y_A^o and y_B^o an index I, and an incremental distribution rule H, group B has a greater ability to pay than group A if $T^* > 0$.

Here we suggest that group B is better off than group A if a transfer to group A from group B results in a reduction in overall inequality. If one takes this approach then it would seem obvious to restrict one's selection of indexes to those which satisfy the basic Pigou-Dalton condition for individual transfers. Beyond this point, however, one is free to explore the implications of alternative indexes in terms of their implicit definitions of groups' abilities to pay. Under a specific incremental distribution rule, some indexes rank groups with high inequality as having greater need, while others tend to rank such groups as having a higher ability to pay. Indexes that concentrate on the lower tail of the distribution are particularly likely to view the world in the former manner. The extreme example of this type of situation is the case of Atkinson's index with a very large ε term. (Note that as ε goes to infinity this index approaches Rawls' criterion.) In common sense language these indexes are saying that a higher level of inequality implies that a group has a relatively larger share of the poor and hence is relatively needy. The other type of index (such as Atkinson's indexes with $\varepsilon < 1.0$) just reverses the argument. A higher level of inequality implies more rich people and hence a greater ability to pay.

The suggestion that an inequality measure implies an ordinal ranking of arbitrary groups naturally raises the question of how such a ranking relates to those given by social welfare functions. Blackorby et al. [1978] has demonstrated that for a broad class of inequality measures there are corresponding well-defined social welfare functions.[10] Thus if we take a particular zero homogenous index I and an arbitrary division of a given population into two groups (A and B) we immediately can apply the corresponding social welfare function W to those two groups. How will the ranking of W_A to W_B relate to the ranking suggested by the notion of ability to pay stated above? In general there is no necessary consistency between these alternative ranking schemes. In Blackorby's approach W is equal to $\mu - \mu I$. Hence, when W_B is greater than W_A a proportional transfer from B to A which lowers the former and raises the latter will not necessarily increase W (or reduce I) for the entire group. The direction of movement in W lepends only on the ability to pay rankings suggested above. Also note in this ituation μ_B may be either greater or smaller than μ_A.

Hence there are three alternative group ranking schemes. First, comparing roup means as suggested by the convergence approach. Second, focusing on notion f ability to pay. Third, comparing the levels of social welfare achieved in the roups. We have noted above that these three schemes are not in general onsistant. To make these contrasts clearer take the example of the Atkinson ndex with ε equal to 1/2 and consider the proportional distribution rule. If we ake group means as the basis of a ranking then obviously B is better off than A f μ_B is greater than μ_A. If we apply a social welfare function to the two oups separately, B is better off than A as long as μ is greater than μ $(1-I_A)/(1-I_B)$. However, if we use ^{B} ^{A} our definition of ability to pay, group B better off than A as long as μ_B is greater than $u_A * (1-I_B)/(1-I_A)$. Moreover when $\mu_A = \mu_B$ using separate social welfare functions to rank A and B, B is ranked better if it has less inequality. However, if we use our ability to pay principle, B is ranked better if it has more inequality, because then a transfer comes from richer people.

Hence, where the primary concern of analysis or policy is the level of inequality for a total population group data must be interpreted carefully. Whatever their other implications may be, neither inter-group mean convergence nor inter-group welfare convergence necessarily imply a reduction in inequality for the entire population. When the group mean incomes are approximately the same a lower level of welfare and correspondingly a higher level of inequality in one group implies (relatively) more of both the rich and the poor in that group. A transfer to that group aids not only the poor but also the rich. Short of changing the intra-group distribution rules this dilemma must be faced. Hopefully the definition of optimal convergence and the related notion of a group's ability to pay developed above help to clarify both the nature of this dilemma and the trade-offs implicit in various indexes of inequality. They do not sacrifice the range of ethical judgments to computational considerations. Rather they add an important dimension to those judgments by focusing on the response of inequality indexes to constrained inter-group transfers and therefore help to clarify both the nature of this dilemma and the trade-offs implicit in various indexes of inequality.

Footnotes

1. At the extreme, virtually no one questions the appropriateness of measuring income inequality in terms of family units, despite the fact that these units undoubtedly distribute their effective purchasing power in very different ways among their members. Virtually every poverty program implicitly aggregates individuals by easily identifiable characteristics. Even the negative income tax and related proposals have not suggested looking inside the family unit to force redistribution at this most micro of levels. One way or another aggregation to the group level is implicit in all of these. In many it is blatantly explicit. Programs for the economic development of poorer regions are among the most obvious examples. While various attempts have been made to put at least minimal conditions on the distribution of the gains from such programs (minority hiring quotas for construction projects, etc.), even these restrictions have generally been couched in aggregative terms.

2. Of course the actual form of H is fundamentally an empirical issue. For a given situation there may well be an H implied by the particulars at work. For example, there is a widespread notion that a country's level of inequality is related in an inverse U fashion to its level of per-capita income. Of course, where such a relation exists and is known, it would be the prime candidate for consideration.

3. Obviously in this situation the Lorenz curve moves up with T. Note that we follow Dasgupta, Sen and Starrett [1973] in constructing a Lorenz curve by plotting Lorenz points for each individual and connecting them by straight lines.

4. A formal proof of Proposition 2 is contained in Appendix A.

5. $F_A(y^*)$ is defined as the proportion of people in group A with incomes less than or equal to y^*. It should be noted that condition (2") is a condition of first order stochastic dominance between the two group income distributions.

6. For each index we present in the Table a definition that is commonly used. In Appendix B we give an alternative representation of such a definition, where possible, in terms of the intra-group inequalities. It is the latter form that we use to derive optimal convergence conditions stated in Table I.

7. In this case where transfers and taxes are proportional to income these marginal utilities must be weighted by the intra-group incomes shares. It is not difficult to show that the condition in Table I is equivalent to equating

$$\frac{1}{N_A} \sum_{i \varepsilon A} \left(\frac{y_{A_i}}{\mu_A}\right) y_{Ai}^{-\epsilon} \text{ to } \frac{1}{N_B} \sum_{j \varepsilon B} \left(\frac{y_{B_j}}{\mu_B}\right) y_{Bi}^{-\Sigma}.$$

Clearly this condition generalizes for the family of incremental distribution rules determined by equation 1. For two rules with α and β the optimal convergence point for an Atkinson index with parameter ε is given

$$\text{by } \sum_{i \varepsilon A} \frac{y_{Ai}^{\alpha}}{\sum_{i \varepsilon A} Y_A^{\alpha}} \cdot y_{Ai}^{-\epsilon} = \sum_{j \varepsilon B} \frac{y_{Bj}^{\beta}}{\sum_{j \varepsilon B} Y_{Bj}^{\beta}} y_{Bj}^{-\Sigma}.$$

8. To get a better idea of the degree of sensitivity of the various indexes to differences in the relative means under the proportional shift, consult Appendix C where full simulations are presented.

9. This property has been extensively discussed by Shorrock [1980], Bourguignon [1979] and Theil [1967]. The work of Blackorby et al. [1978] on consistency is also closely related to this issue. Several of these authors have argued that a disaggregation property is highly attractive in its own right. The following discussion suggests that the price of this disaggregation property is quite high in terms of the restrictions it imposes on group rankings.

10. Note that the discussion by Blackory et al. [1978] is an extension of that by Atkinson [1970].

Appendix A

Proof of Proposition II

As stated in the text, the proof of Proposition II follows from the results stated by Dasgupta, Sen and Starrett [1973]. In particular, Dasgupta et al. stated the equivalence of the following two conditions for two income distributions $y_1^o, \ldots, y_N^o$ and $y_1', \ldots, y_N'$: [conditions (ii) and (iv) in their paper]

A.1 $y_1' + \ldots + y_K' \geq y_1^o + \ldots + y_K^o$, all $k \leq N$ (with strict inequality for at least one k) and $y_1' + \ldots + y_N' = y_1^o + \ldots + y_N^o$

and

A.2 for any strict concave function U,

$$U(y_1') + \ldots + U(y_N') > U[y_1^o] + \ldots + U[y_N^o].$$

To prove Proposition II, we show first that condition (2a) is equivalent to condition (A.1) for a range of $o < T \leq \hat{T}$. Consider income vectors y_A and y_B and the corresponding income vector for the total population y such that $\overset{o}{y} = [y_A^o, y_B^o]$. Now order y such that $y_1^o \leq y_2^o \leq \ldots \leq y_N^o$. Now choose income transfer T so that

$$y_k^o > y_\ell^o \rightarrow y_k' > y_\ell' \quad , \; k,\ell = 1,N.$$

where y_k' is the after transfer income of individual k.

In other words, T is chosen such that no individual's income ranking among the population will be reversed. For such a T and any $k \leq N$ by the mean value theorem we know that there exists a position ( ) such that

$$\sum_{i=1}^{k} y_i' = \sum_{i=1}^{k} y_i^o + [Z(y_A^o, y_k^o) + Z(y_B^o, y_k^o)] \cdot T$$

It is clear from the equation above that condition (2a) is equivalent to condition (A.1) which is a statement that the after transfer distribution $y' = [y'_A, y'_B]$ is Lorenz superior to the before distribution y. Hence the transfer T will reduce all Pigou-Dalton Indexes.

To prove the last part of Proposition II we simply draw on the equivalence of condition (A.1) and (A.2) (implying the equivalence of conditions (2a) and (A-2)) and the existance of a correspondence between U and I.

Appendix B

$$V = \sum_{J=1}^{Q} \frac{N_J}{N} V_J^2 + \sum_{J=1}^{Q} \frac{N_J}{N} (\mu_J - \mu)^2$$

$$V^* = \sum_{J=1}^{Q} \frac{N_J}{N} \left(\frac{\mu_J}{\mu}\right)^2 V_j^* + \sum_{J=1}^{Q} \frac{N_J}{N} \left(\frac{\mu_J}{\mu} - 1\right)^2$$

$$G = 1 + \frac{1}{N} - \frac{1}{N^2 \mu} (y_1 + 2y_2 + \ldots + Ny_N)$$

$$y_1 > y_2 > \ldots > y_N$$

$$D = \frac{1}{N\mu} \sum_{i=I} \sum_{j \varepsilon J} |y_{ij} - \mu|$$

$$T = \sum_{J=1}^{Q} \frac{N_J \mu_J}{N\mu} T_J + \sum_{J=1}^{Q} \frac{N_J \mu_J}{N_\mu} \log \frac{N_J \mu_J}{N_\mu}$$

$$L = \sum_{J=1}^{Q} \frac{N_J}{N} L_J + \sum_{J=1}^{Q} \frac{N_J}{N} (\log \hat{\mu}_J - \log \hat{\mu})^2$$

$$C = \sum_{J=1}^{Q} \frac{N_J}{N} C_J + \sum_{J=1}^{Q} \frac{N_J}{N} (\log \mu - \log \mu_J)$$

$$A = 1 - \left[\sum_{J=1}^{Q} \frac{N_J}{N} \left(\frac{\mu_J}{\mu}\right)^{1-\varepsilon} (1-A_J)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}$$

where Q is the # of groups

Appendix C

$\frac{\mu S}{\mu N}$	V*	G	D	T	L	ε = .50	ε = .75
0.80	0.5069	.36827	0.5181	-17.6562	0.5884	0.1132	0.1699
0.82	0.5046	.36742	0.5171	-17.6580	0.5856	0.1127	0.1693
0.84	0.5027	.36667	0.5162	-17.6580	0.5832	0.1124	0.1687
0.86	0.5012	.36602	0.5155	-17.6598	0.5811	0.1120	0.1682
0.88	0.5000	.36547	0.5149	-17.6598	0.5793	0.1118	0.1678
0.90	0.4991	.36501	0.5145	-17.6598	0.5778	0.1116	0.1675
0.92	0.4986	.36464	0.5142	-17.6598	0.5765	0.1114	0.1672
0.94	0.4983	.36436	0.5139	-17.6615	0.5754	0.1113	0.1671
0.96	0.4984	.36418	0.5138	-17.6615	0.5745	0.1112	0.1669
0.98	0.4987	.36411	0.5137	-17.6615	0.5739	0.1112	0.1669
1.00	0.4993	.36412	0.5138	-17.6598	0.5735	0.1112	0.1669
1.02	0.5002	.36423	0.5139	-17.6598	0.5732	0.1113	0.1669
1.04	0.5013	.36445	0.5141	-17.6598	0.5731	0.1113	0.1670
1.06	0.5026	.36474	0.5143	-17.6598	0.5732	0.1115	0.1672
1.08	0.5042	.36511	0.5147	-17.6598	0.5735	0.1116	0.1674
1.10	0.5060	.36553	0.5151	-17.6598	0.5738	0.1118	0.1676
1.12	0.5080	.36600	0.5156	-17.6580	0.5744	0.1120	0.1679
1.14	0.5103	.36652	0.5161	-17.6580	0.5750	0.1123	0.1682
1.16	0.5127	.36709	0.5167	-17.6580	0.5758	0.1126	0.1686
1.18	0.5153	.36770	0.5173	-17.6562	0.5767	0.1129	0.1690
1.20	0.5181	.36834	0.5180	-17.6562	0.5777	0.1132	0.1695
1.22	0.5211	.36903	0.5188	-17.6544	0.5788	0.1135	0.1699
1.24	0.5242	.36974	0.5196	-17.6544	0.5800	0.1139	0.1704
1.26	0.5275	.37049	0.5204	-17.6526	0.5813	0.1143	0.1710
1.28	0.5309	.37127	0.5213	-17.6526	0.5827	0.1147	0.1715
1.30	0.5345	.37207	0.5223	-17.6508	0.5842	0.1152	0.1721
1.32	0.5383	.37290	0.5233	-17.6490	0.5857	0.1156	0.1727
1.34	0.5421	.37376	0.5243	-17.6490	0.5874	0.1161	0.1734
1.36	0.5461	.37464	0.5254	-17.6472	0.5891	0.1166	0.1740
1.38	0.5503	.37554	0.5265	-17.6455	0.5909	0.1171	0.1747
1.40	0.5545	.37646	0.5276	-17.6435	0.5927	0.1176	0.1754

Appendix C (cont.)

ε=1 A $\|\bar{c}\|$	ε = 1.50	ε = 2.00	ε = 3.0	ε = 4.0	ε = 21.0
0.2572	0.3399	0.4482	0.6240	0.7308	0.9195
0.2561	0.3386	0.4466	0.6217	0.7283	0.9181
0.2551	0.3375	0.4452	0.6197	0.7260	0.9167
0.2543	0.3366	0.4439	0.6179	0.7238	0.9153
0.2536	0.3358	0.4428	0.6162	0.7218	0.9139
0.2530	0.3351	0.4418	0.6147	0.7200	0.9126
0.2526	0.3345	0.4410	0.6134	0.7183	0.9112
0.2522	0.3341	0.4403	0.6122	0.7168	0.9099
0.2520	0.3337	0.4398	0.6112	0.7154	0.9086
0.2518	0.3335	0.4393	0.6103	0.7142	0.9074
0.2518	0.3333	0.4390	0.6096	0.7131	0.9061
0.2518	0.3333	0.4388	0.6089	0.7121	0.9049
0.2520	0.3333	0.4387	0.6084	0.7113	0.9036
0.2522	0.3334	0.4386	0.6080	0.7105	0.9024
0.2525	0.3336	0.4387	0.6077	0.7099	0.9013
0.2528	0.3338	0.4389	0.6075	0.7094	0.9001
0.2532	0.3342	0.4391	0.6074	0.7090	0.8990
0.2537	0.3346	0.4394	0.6073	0.7087	0.8979
0.2543	0.3350	0.4398	0.6074	0.7084	0.8968
0.2549	0.3355	0.4402	0.6075	0.7083	0.8958
0.2556	0.3361	0.4407	0.6078	0.7082	0.8949
0.2563	0.3367	0.4413	0.6080	0.7083	0.8940
0.2570	0.3374	0.4419	0.6084	0.7083	0.8932
0.2579	0.3381	0.4426	0.6088	0.7085	0.8925
0.2587	0.3389	0.4434	0.6093	0.7087	0.8919
0.2596	0.3397	0.4441	0.6098	0.7090	0.8914
0.2606	0.3406	0.4450	0.6104	0.7093	0.8911
0.2616	0.3414	0.4458	0.6110	0.7097	0.8909
0.2626	0.3424	0.4467	0.6116	0.7102	0.8908
0.2637	0.3433	0.4477	0.6124	0.7107	0.8908
0.2648	0.3443	0.4487	0.6131	0.7112	0.8910

I Conditions for Optimal Convergence Under
Per Capita and Proportional Distribution Rule

ex	Per Capita Rule	Proportional Rule
$\sum_j (y_j - \mu)^2$	$\mu_A = \mu_B$	$\mu_A(1+V_A^\star) = \mu_B(1+V_B^\star)$
$\sum_i (\frac{y_i - \mu}{\mu})^2$	$\mu_A = \mu_B$	$\mu_A(1+V_A^\star) = \mu_B(1+V_B^\star)$
$\frac{1}{N^2\mu} \sum_i \sum_j (y_i - y_j)$	$\bar{R}_A = \bar{R}_B$	$\frac{1}{N_A\mu_A} \sum_{i\varepsilon A} y_{Ai}R_i = \frac{1}{\mu_B\mu_B} \sum_{j\varepsilon B} y_{Bj}R_j$
$\frac{1}{\mu} \sum_i \lvert y_i - \mu \rvert$	$F_A(\mu) = F_B(\mu)$	$S_A(\mu) = S_B(\mu)$
$\frac{y_i}{N\mu} \log \frac{y_i}{N_\mu}$	$\hat{\mu}_A = \hat{\mu}_B$	$\log (N_A\mu_A) + T_A = \log(N_B\mu_B) +$
$\sum(\log y_i - \log \hat{\mu})^2$	$\frac{1}{N_A} \sum_{i\varepsilon A} \log y_{Ai} \frac{1}{y_{Ai}} - \frac{\log \hat{\mu}}{N_A} \sum_{i\varepsilon A} \frac{1}{y_{Ai}}$ $= \frac{1}{N_B} \sum_{j\varepsilon B} \log y_{Bj} \frac{1}{y_{Bj}} - \frac{\log \hat{\mu}}{N_B} \sum_{j\varepsilon B} \frac{1}{y_{Bj}}$	$\hat{\mu}_A = \hat{\mu}_B$
$\log \mu - \log \hat{\mu}$	$\frac{1}{N_A} \sum_{i\varepsilon A} \frac{1}{y_{Ai}} = \frac{1}{N_B} \sum_{j\varepsilon B} \frac{1}{y_{Bj}}$	$\mu_A = \mu_B$
$- [\sum \frac{1}{N} (\frac{y_i}{\mu})^{1-\varepsilon}]^{\frac{1}{1-\varepsilon}}$	$\frac{1}{N_A} \sum_{i\varepsilon A} y_{Ai}^{-\varepsilon} = \frac{1}{N_B} \sum_{j\varepsilon B} y_{Bj}^{-\varepsilon}$	$\frac{\mu_A}{(1-A_A)^{\frac{1-\varepsilon}{\varepsilon}}} = \frac{\mu_B}{(1-A_B)^{\frac{1-\varepsilon}{\varepsilon}}}$

(continued)

Table 1 (cont.)

V = variance; V^* = square of the coefficient of variation;

G = Gini coefficient; D = relative mean absolute deviation;

T = Theil's Entropy; L = variance of the logarithms;

C = The log of the ratio of the arithmatic mean to the geometric mean; and

A = Atkinson's Index.

[note: C is a specific A with $\varepsilon = 1.0$]

R_i = ranking of individual i in either group A or group B in the overall distribution, $\bar{R}_A = \frac{1}{N_A} \sum_{i \varepsilon A} R_i$, and $\bar{R}_B = \frac{1}{N_B} \sum_{j \varepsilon B} R_j$; $F_A(\mu) = \frac{N^A(\mu)}{N_A}$

where $N^A(\mu)$ = # of persons in A with income equal to or below μ, and

$S_A(\mu) = \frac{\sum_{yA_i \leq \mu} yA_i}{N_A \mu_A}$; similarily for $F_B(\mu)$ and $S_B(\mu)$

Table II

Inequality Indexes - 1979

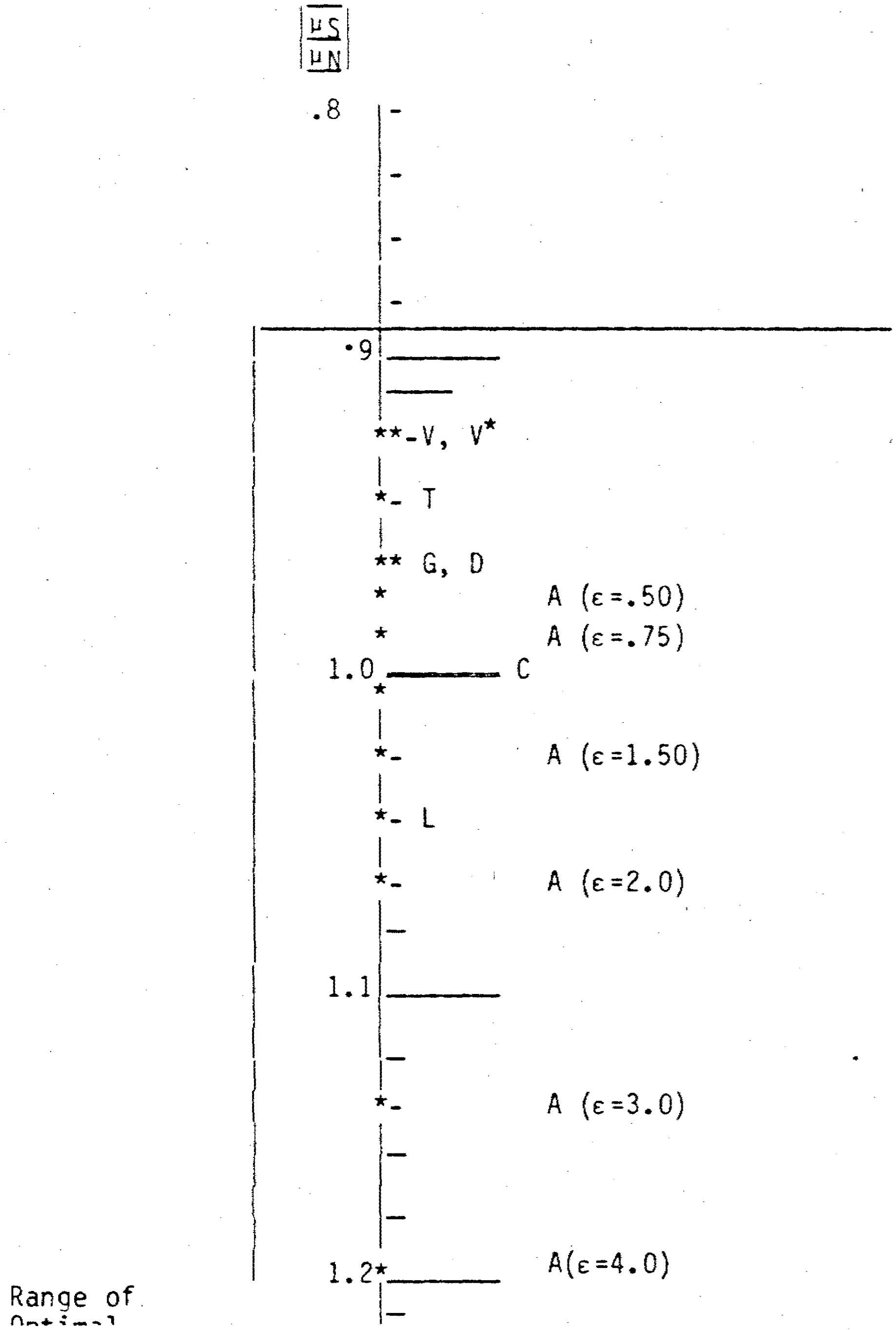
	V^*	G	D	T^a	L	A: ε=.5	ε=.75	ε=1.0 (c)	ε=1.5	ε=2.0	ε=3.0	ε=4.0	ε=21.0
U.S.	0.499	0.365	0.514	-17.66	0.578	0.112	0.167	0.253	0.335	0.442	0.615	0.720	0.912
NON-SOUTH	0.472	0.356	0.502	-17.67	0.540	0.106	0.159	0.239	0.318	0.418	0.581	0.681	0.877
SOUTH	0.555	0.380	0.537	-17.62	0.640	0.121	0.182	0.278	0.363	0.477	0.653	0.752	0.911

a. T ranges between $\log \frac{1}{N}$ (<0) [complete equality] and 0

Source: U.S. Bureau of the Census [1981] Current Population Reports, Consumer Income, Series P-60, No. 129, November, Table 15 pp. 59-60.

Range of Optimal Convergence for all Pigou-Dalton I
and Optimal Convergence Points for Various Indexes.

Figure I



References

1. A. B. Atkinson, "On the Measurement of Inequality," Journal of Economic Theory 2 (1970), pp. 244-263.

2. C. Blackorby and D. Donaldson, "Measures of Relative Equality and Their Meanings in Terms of Social Welfare," Journal of Economic Theory 18, (1978), pp. 59-80.

3. C. Blackorby, D. Primout and R. Russell, Duality, Separability, and Functional Structure: Theory and Economic Applications, American Elsevier/North Holland, New York/Amsterdam, 1978.

4. F. Bourguignon, "Decomposable Income Inequality Measures," Econometrics, Vol. 47, No. 4 (July, 1979), pp. 901-920.

5. P. Dasgupta, A. Sen and D. Starrett, "Notes on the Measurement of Inequality," Journal of Economic Theory 6 (1973), pp. 180-187.

6. Kolm S.-C., "Unequal Inequalities. I," Journal of Economic Theory 12, (1976), pp. 416-442.

7. Kolm S.-C., "Unequal Inequalities. II," Journal of Economic Theory, 13 (1976), pp. 82-111.

8. A. F. Shorrocks, "The Class of Additively Decomposable Inequality Measures," Econometrica, Vol. 48, No. 3 (April 1980), pp. 613-625.

9. M. Y. Tam and J. Persky, "Regional Convergence and National Inequality," Review of Economics and Statistics, Vol. LXIV, No. 1 (1982), pp. 161-164.

10. H. Theil, Economics and Information Theory, Amsterdam: North Holland (1967).

El Centro de Estudios Económicos de El Colegio de México, ha creado la serie "Documentos de Trabajo" para difundir investigaciones que contribuyen a la discusión de importantes problemas teóricos y empíricos aunque estén en versión preliminar. Con esta publicación se pretende estimular el análisis de las ideas aquí expuestas y la comunicación con sus autores. El contendio de los trabajos es responsabilidad exclusiva de los autores.

Editor: José Luis Alberro

Serie Documentos de Trabajo 1982

No. I — Ize, Alain "Disequilibrium Theories, Imperfect Competition and Income Distribution: A Fix Price Analysis"

No. II — Levy, Santiago "Un Modelo de Simulación de Precios para la Economía Mexicana"

No. III — Persky, Joseph and Tam, Mo-Yin S. "On the Theory of Optimal Convergence"

No. IV — Kehoe, Timothy J., Serra-Puche, Jaime y Solís, Leopoldo "A General Equilibrium Model of Domestic Commerce in Mexico"

No. V — Guerrero, Víctor M. "Medición de los Efectos Inflacionarios Causados por Algunas Decisiones Gubernamentales: Teoría y Aplicaciones del Análisis de Intervención"

No. VI — Gibson, Bill, Lustig, Nora and Taylor, Lance "Terms of Trade and Class Conflict in a Computable General Equilibrium Model for Mexico"

No. VII — Dávila, Enrique "The Prices System in Cantillon's Feudal-Mercantile Model"

No. VIII — Ize, Alain "A Dynamic Model of Financial Intermediation in a Semi-Industrialized Economy"

No. IX — Seade, Jesús "On Utilitarianism and Horizontal Equity: When is the Equality of Incomes as such Desirable"

No. X — Cárdenas, Enrique "La Industrialización en México Durante la Gran Recesión: Política Pública y Respuesta Privada"