

Gini index, inequality comparisons and maximin axiom: a paradox

Louis de Mesnard

Keywords. Gini, maximin, Lorenz, concentration, inequality.

JEL Classification. C43, D63, I32.

Abstract. Inequality is appreciated only through a relevant measure, traditionally the Gini index, but this one presents a weakness that prevents to consider it as valid: to a same Gini may correspond many distributions. For the most simple case, a concentration curve with two linear segments, a criterion is proposed, based on the maximin: beside curves with same Gini, the more egalitarian is those in which each poor has the higher revenue. However, this does not allow to decide for two curves with different Gini: two indeterminate zones appear. This is extended to curves with three linear segments and two kinks.

Author.

Prof. Louis de Mesnard
LATEC (UMR CNRS 5601)
Faculty of Economics
University of Burgundy
2 Bd Gabriel
21000 Dijon, FRANCE
E-mail : louis.de-mesnard@u-bourgogne.fr

I. Introduction

Many authors use Gini index as a measure of inequality. Often, one think that a low concentration, i.e., a low Gini index, goes hand in hand with equality. This is false because several distributions of income, nevertheless well different, can seem to have a same level of inequality measured by the index of Gini. This is the main problem with the Gini index: it doesn't discriminate between some distributions of the variable x even if they are obviously different. The question is not to determine if two different distributions can have the same Lorenz curve, but if two distributions can have the same value of the Gini index. The answer to this question is yes as one see in what follows.

The difference between two distributions, however symmetrical about the second diagonal can be large with the same Gini index. For example, the two distributions of table 1 have the same value of the Gini index ($I_G = 0.3$) and are symmetrical even if they are very different:

Table 1 here

The question is: between two distributions with the same Gini index, what is the more egalitarian. As an example, consider two discrete distributions of revenue, $\{2, 2, 2, 4\}$ and $\{1, 3, 3, 3\}$. For both distributions, the Gini index is $I_G = 0.15$: the two indices are equal ¹. However, what is the most not egalitarian distribution, the first or the second? In table 1, starting from the first distribution to go to the second, a transfer of 15 individuals from the

¹ Remember that Gini Index is insensible to an homothetic change in the revenues. For the same reason, the total number of individuals does not matter: one can always normalize to have the same number of individuals in each distribution.

poorer to the richer class does not affect the measure of inequality by the Gini index: this is confusing! This phenomenon of symmetry is indirectly mentioned by Thon (1982, p. 142) for the axiom "SC" for *Strong Comparability* but Thon does not investigate further this way. To a same value of the Gini index can correspond an infinity of distributions, not only by a symmetry effect about the second diagonal.

Another example. In management science (stock management, personnel management, clientèle management, etc.), the so-called "80-20 law" is known. It is called like this because in management science, 80% of the population are reputed to represent 20% of the revenue and 20% of the population 80% of the revenue (see figure 1, obtained for the following distribution: when each of 80% of individuals have a revenue of 1, each of 20% other have a revenue of 16; here the kink is exactly on the second bisector, with a Gini index equal to 0.6). A first segment has a low slope, a second has a high slope and they are separated by only one kink. In another case, 85% of people earn 25% of revenues and the 15% other earn the 75% remaining revenues (when one of the poorer earns 1, one of the richer earns 17); in a last case, 75% of people earn 15% of revenues and 25% earn 85% (when one poorer earns 1, one richer earns 17). What is the more egalitarian case, knowing that in these three cases, the value of the Gini index is 0,6 (and cases 2 and 3 are symmetrical regarding to the second diagonal; see table 2)?

Figure 1 here

Table 2 here

So, the discrimination between distributions to say what is the less egalitarian is not easy ². In this paper, restricting the analysis to the Gini index alone and leaving aside other measures as the Kuznets index (see Kuznets (1963), McCabe (1974)) or as the entropy (see Bourguignon (1979), Shorrocks (1980), Cowell (1980), Cowell and Kuga (1981)), I will try to introduce a characterization of inequality that takes into account the above critics and tries to respond to it in the most simple case: a concentration curve approximated by a two-linear-segment curve. The case of a concentration curve with three linear segments will be discussed also.

II. The 80-20 model of measure of inequality

For simplicity, one will assume that the distribution have only two groups and the Gini curve is composed only by two linear segments . This corresponds to the *80-20 law*. What is the less egalitarian distribution? What distribution do we prefer? The problem is to distinguish between two distributions that have the same value of the Gini index: the kinks are aligned along a 45° straight line because, denoting (x, y) the coordinates of the kink, the area under the curve is $A = \frac{1}{2}(1 - x + y)$ and the Gini index is $I = 1 - 2A = x - y$. In $\{2, 2, 2, 4\}$ you have one rich facing three persons of low-middle class when in $\{1, 3, 3, 3\}$ you have one poor facing three persons of high-middle class, but the choice is not clear. At first glance, the second distribution seems less egalitarian than the first because the ratio of the higher to the lower revenue is 3

² One must not confuse. This phenomenon, multiple distributions can correspond to the same Gini index, cannot be removed even when the generalized Lorenz curve is used. It is proved that there is a duality between a distribution and the generalized Lorenz curve that comes from it (Thistle 1989) (see also Iritani and Kuga (1983)). However, the generalized Lorenz curve is not suitable to calculate an index like Gini's one.

when it is only 2 in the first one, and at the same time, poorer people are less poor but they are much more numerous in the first case.

For the example indicated by table 2, from the right skewed case to the left skewed case, the ratio of the higher revenue over the lower remains fixed (17:1) so this ratio plays no role. In the left skewed case, each 1% of the poor earns 0.200% of the revenue when in the right skewed case, they earn 0.294% of the revenue (and 0.250% in the not skewed case); so poorer people are less poor in right skewed case than in left skewed case. Moreover, looking at the limit cases of this example, indicated by table 3, we see that the whole population is in one category in the right skewed case where is separated in two in the left skewed case: in a certain sense, the right skewed case is more egalitarian.

Table 3 here

So, we propose to adapt the Rawls's maximin to obtain a criterion, that I call axiom II. Among two distributions (with only two classes of revenue), all things equal by elsewhere (i.e., with the same value of the Gini index), **the distribution in which the poor people have the larger part of the revenue is declared as more egalitarian. Between two 2-classes-distributions with the same value of the Gini index, the distribution that is the more right skewed is qualified as more egalitarian than the other distribution.** For example $\{2, 2, 2, 4\}$ is more egalitarian than $\{1, 3, 3, 3\}$, as the right skewed case of table 2. Looking at the two distributions of table 1, the first distribution (80% of poor but with 50% of revenue instead of 50% of poor but with 20% of revenue) is the more egalitarian (see figure 2). Between two distributions with the same Gini index, this axiom declares that the more egalitarian is the distribution that provides **the higher revenue in percentage to each 1% of**

the population of the poor. This works for two distributions that are both right skewed or both left skewed, with the same Gini index.

Figure 2 here

Generally, one declares as more egalitarian the distribution with the lower Gini index all things equal by elsewhere (axiom I) but we know that angle α is important. What happens for two distributions that have not the same Gini index? In other terms: what do you prefer, a highly concentrated distribution but skewed to the right or a low concentrated distribution but skewed to the left? This seems to be indeterminate.

As shown by figure 3, consider the constant concentration line BC passing by a (along this line, the concentration is constant) and consider the line OA , passing by the origin and a (along this line, the revenue of 1% of the population that is poor is constant; I call this revenue the "unit revenue of poor", or URP). Above BC , the concentration is lower, that is preferable for equality (axiom I). Above OA , the unit revenue of poor is higher, that is preferable for equality (axiom II). So, in the sector BaC , concentration and unit revenue of poor are better for equality: all kinks inside BaC are dominating a , regarding equality. In sector AaC , all kinks are dominated by a because both concentration and unit revenue of poor are less favorable. For example, define p as the intersection of the line of constant unit revenue of poor that passes by a and the constant concentration line that passes by kinks b, c, d (all these kinks belong into OaB and have the same concentration, lower than in a); p dominates a because the concentration in p is lower than in a along a constant unit revenue of poor line (axiom I), when d dominates p along a constant concentration line (axiom II); finally, d dominates a ; the same reasoning holds with p' . Similarly, b , the kink with the same abscissa than a , dominates

d, p and $a; c$ dominates b, d, p and a . Conversely, a dominates p'' , and p'' dominates e , a kink inside AaC .

There remain sectors OaC and BaA , that are indeterminate: in the first sector concentration is better but unit revenue of poor is lower when in the second sector concentration is bad but unit revenue of poor is good. Kink f inside OaC is dominated by p , but p dominates a ; conversely, kink g inside AaB dominates p' but a dominates p' also. Curves that have their kinks in these indeterminate sectors cannot be compared. When a is displacing along the constant concentration line, the indecisiveness zone OaC dwindles and the other, BaA increases when a goes to C , and conversely when a goes to B .

To justify axiom II, simply assume that it is reversed, i.e., the more egalitarian is the kink where each 1% of the population that is poor is less rich for a same Gini index; or, the kink with the lower angle β is chosen. Now, sectors OaC and BaA are more favorable to equality. OaB and AaC sectors are indeterminate: the complete no-concentration (obtained when the kink is on the line OZ) appears as indeterminate compared to a and if the kink is on the X-axis, the case is also indeterminate compared to a when the poor have zero as revenue! This is shocking: axiom II cannot be reversed. This is not a demonstration because another axiomatic rule could have been chosen, as "for a same concentration, all kinks are equivalent", but this indicates how to choose between two kinks with equal Gini.

Figure 3 here

III. The ABC model of measure of inequality

All the above results are obtained for a two-segments concentration curve, i.e., for a two-classes distribution. The following section will analyze what happens for a three-segments concentration curve, that is for a three-classes distribution. This will be a first approach for a generalization of these results to any concentration curve, even if things are rapidly much more complicated.

One will assume that the concentration curve has a special form (figure 4): a first segment quasi-horizontal, named segment A, a second with an intermediary slope, named segment B that we may qualify as the marshland (in French, the "marais", by reference to the political situation during the French Revolution), and a third quasi-vertical, named segment C. The width of the B segment is important: when the width is high, there is a large middle class. In this model, three segments are separated by two kinks. This decomposition is very known in managerial practice: it is the so-called *A-B-C law*. By example for stock management, where parts are managed following a decomposition in three: an important share of parts in stock has a low turnover, another share has a more important turnover, and for a small share the turnover is very high.

Figure 4 here

The variety of situations is larger than with one kink. One may have curves with a tangent zone placed symmetrically regarding to the second diagonal; curves skewed to the left, i.e. with a tangent zone or a B zone situated to left of the second diagonal (inequality in disfavor of the poorer, with a more numerous high-middle class, like in the second distribution given in

example above); curves skewed to the right, i.e. with a tangent zone or a kink situated to right of the second diagonal (inequality in favor of the richer, with more poorest people, like for the first distribution given in example); all admissible combinations of former cases, as a large tangent zone to the left of the second diagonal. Sometimes, the B segment may be astride on the second diagonal, even more to the left or to the right.

For distributions with two kinks, things are more complicated than with only one kink because we have two degrees of freedom in addition: the constant concentration curve cannot be set easily, even if an infinity of three-segments curves can have the same Gini index. Denoting (x_L, y_L) and (x_R, y_R) the coordinates of left and right kinks respectively, the area under the curve and the Gini index are respectively:

$$A = \frac{1}{2} (1 - x_R + y_R - x_L y_R + x_R y_L)$$

$$I = 1 - 2A = x_R - y_R + x_L y_R - x_R y_L.$$

I depends on four numbers by a not linear function, even if, when (x_R, y_R) is given, y_L depends linearly on x_L by:

$$y_L = \frac{x_R - y_R - y_R x_L - I}{x_R}$$

and when (x_L, y_L) is given, y_R depends linearly on x_R by:

$$y_R = \frac{x_R (1 - y_L) - I}{x_L - 1}$$

In this last case, the constant concentration curve is a straight line but only if the relative situation of poor is given. For example, if two curves have the same left kink, the reasoning about the position of the right kink is similar than for the one-kink case with the same type of undetermined zones, but in all other cases, complexity is higher, what does not signify that indecision is lower, on the contrary! Indecision remains when, besides two distributions, one

have the lower Gini but the other provide the higher revenue to, either the poor (at left kink), either the poor plus the middle-class people (at right kink): this will be clarified by the following. For example, assume that the left kink is privileged and consider the distributions of table 4 and figure 5. Distribution 1 and 2 have the same Gini, but in distribution 2 each percent of the population that is poor receives more: it is more egalitarian; distribution 3 and 4 are also equivalent in concentration, but distribution 4 has a better *URP*; anyway there is an indecision between distribution 1 in one hand and 3 and 4 in another hand, but distribution 2 is more egalitarian than distributions 3 and 4; distribution 5 is more egalitarian than distribution 6 and than all other distributions when the situation of distribution 6 cannot be decided compared with all other.

Table 4 here

Figure 5 here

In some cases, there is a main kink with an acute angle (measured along the curve), and the second kink is not clear with an obtuse angle (measured along the curve). In these cases, we consider only the main kink and we prefer positive angles according to the one-kink cases studied before. In other terms, only the main kink is considered and two consecutive classes are aggregated because they are considered as similar. However, the aggregation of the two first classes into one class (an operation that removes the first kink but conserves the position of the second kink) dwindles concentration because inequality between poor and middle class is erased. So, approximating the right kink as the main kink is correct only if the changing of slope at the first kink is low, i.e., if the revenue of each poor and each middle class person is

very similar ³. For example, marking the main kinks by *a* and *b* (figure 6), curve *I* have a main kick at *a* and curve *II* at *b*; concentration indicated by curves *I* and *II* does not change strongly if the other kink is neglected; if the criterion of maximin is applied, then curve *I* is skewed to the right but not curve *II*.

Figure 6 here

In all other cases, the two kinks must be considered when comparing two curves with the same Gini index. This type of curve describes a population with three classes, with the poorer at the left side of the concentration curve. The left kink describes the inequality between the poor class and the middle class and the right kink describes the inequality between the aggregate of the poor and the middle classes in one hand and the rich class in the other hand: there is no symmetry between kinks. If you think that the first phenomenon is the more important, you consider the left kink as the main kink. If you think that the second phenomenon - inequality between the poor and middle classes in one hand and the rich class in the other hand - is the more important, you must consider the right kink. In both cases, the maximin can be applied, but considering only the poor by the left kink and the poor plus the

³ A symmetrical reasoning holds for the left kink as a main kink (middle class and rich are aggregated) but this is not a problem because the maximin focus on the poorer part of the society.

The distribution inside groups is not taken in account, as in the papers that work on the inequality decomposition analysis (see Satchell (1987), Lambert and Aronson (1993)). If the inequality inside groups is taken in account, the Gini index is increased because the linear segments of the Lorenz curve are replaced by convex curves.

middle class by the right kink. In figure 7, curves are intersecting twice: curve *II* is chosen with the rule of main kick (kink *b*), but distribution *I* is chosen at right kink whereas inequality between poor and middle classes is high in distribution *I* (at left kink). The choice between *I* and *II* depend on what you consider as an important thing. Note that this problem occurs generally with these twice intersecting curves. As a particular case, suppose that the two distributions are such that the curves are not intersecting and curve *I* has the lower Gini.

Figure 7 here

All this can be generalized cautiously to distributions with many categories of revenue or to continuous distributions. Two or more continuous curves can have the same Gini: the more right skewed is the more egalitarian and curves can with different Gini can be compared from the two axioms. The problem is how to detect the kink for continuous functions to approximate them by a linear segment function. Remember that the curve is convex, so the first derivative of the curve is always positive or null and it is increasing (the second derivative is positive or null). So, the problem is to detect where the second derivative is null or near zero (and the curve is a straight line or a quasi straight line) and where it is significantly positive (and there is clearly a kink). To do so, one could just look at the second derivative curve. When there is no clear kink, i.e. when the second derivative of the curve changes to much slowly to be noticeable, what often happens for continuous distributions, a good approach consists into replacing the kink by another significant point and considering this point as if it was a true kink. It seems to be natural to consider the point where the slope of the curve is equal to 1, i.e., the point where the curve is tangent to a parallel to the first diagonal (figure 8).

Figure 8 here

The curve can be skewed to the left or to the right even when no kink. For example, a uniform distribution (i.e., same number of individuals for each revenue e.g. $n_j = n$ for all j , with the same distance between each revenue e.g. $x_j = a + bj$ for all j) corresponds to a Lorenz curve skewed to the left. Here the abscissa 0.5 is at the tangency point with the bisector. In figure 9, $a = 5$, $b = 10$ and $N = 10$.

Figure 9 here

IV. Conclusion

Inequality is a central concept of the social economy but it can be appreciated only by a relevant measure. One uses traditionally the Gini index but many different distributions can correspond to a same value of the Gini index. The most simple case has been first studied: a concentration curve with only two linear segments and one kink. From the maximin principle, an axiomatic criterion of choice between two distributions that have the same value of the Gini index has been proposed: between two distributions with one kink, the concentration curve that is the most twisted to the right is declared as the more egalitarian because it gives the most possible to each of the poor.

This entails a consequence. When one compares two curves with a different Gini index, one has to apply also another axiomatic criterion: the curve that has the lower Gini is reputed to be the more egalitarian. There is therefore a paradox because the two axioms are not compatible everywhere; for the position of the kink, two zones are undetermined: in one zone the Gini is

higher but the unit revenue of poor is higher, in the other it is the opposite. Hence, to remove the indecision contained in the Gini index (to a same Gini correspond an infinity of curves), induces a new indecision (two curves of different Gini are not always comparable, depending on the position of their kink).

The procedure has been extended also, but of a more complicated manner, to the case of concentration curves that have three linear segments and two kinks: excepted in cases where a kink is clearly more marked than the other, one can grant more importance to the right kink than to the left kink, or conversely, following that one gives the pre-eminence to the analysis of inequality between poor people and middle class, or to the analysis of inequality between rich people and the rest of the population. This is also generally applicable to continuous curves or curves with no true kink.

V. Bibliographical references

- IRITANI Jun and Kiyoshi KUGA (1983) "The duality between the Lorenz curve and the income distribution functions", *Economic Studies Quarterly*, April, pp. 9-21.
- LAMBERT Peter J. and J. Richard ARONSON (1993) "Inequality decomposition analysis and the Gini coefficient revisited", *The Economic Journal*, 103, pp. 1221-1227.
- SATCHELL S. E. (1987) "Source and subgroup decomposition inequalities for the Lorenz curve", *International Economic Review*, Vol. 28, n° 2, pp. 323-329.
- THISTLE Paul. D. (1989) "Ranking distributions with generalized Lorenz Curves", *Southern Economic Journal*, 56(1), pp. 1-12.
- THON Dominique (1982) "An axiomatization of the Gini coefficient", *Mathematical Social Science*, 2, pp. 131-143.

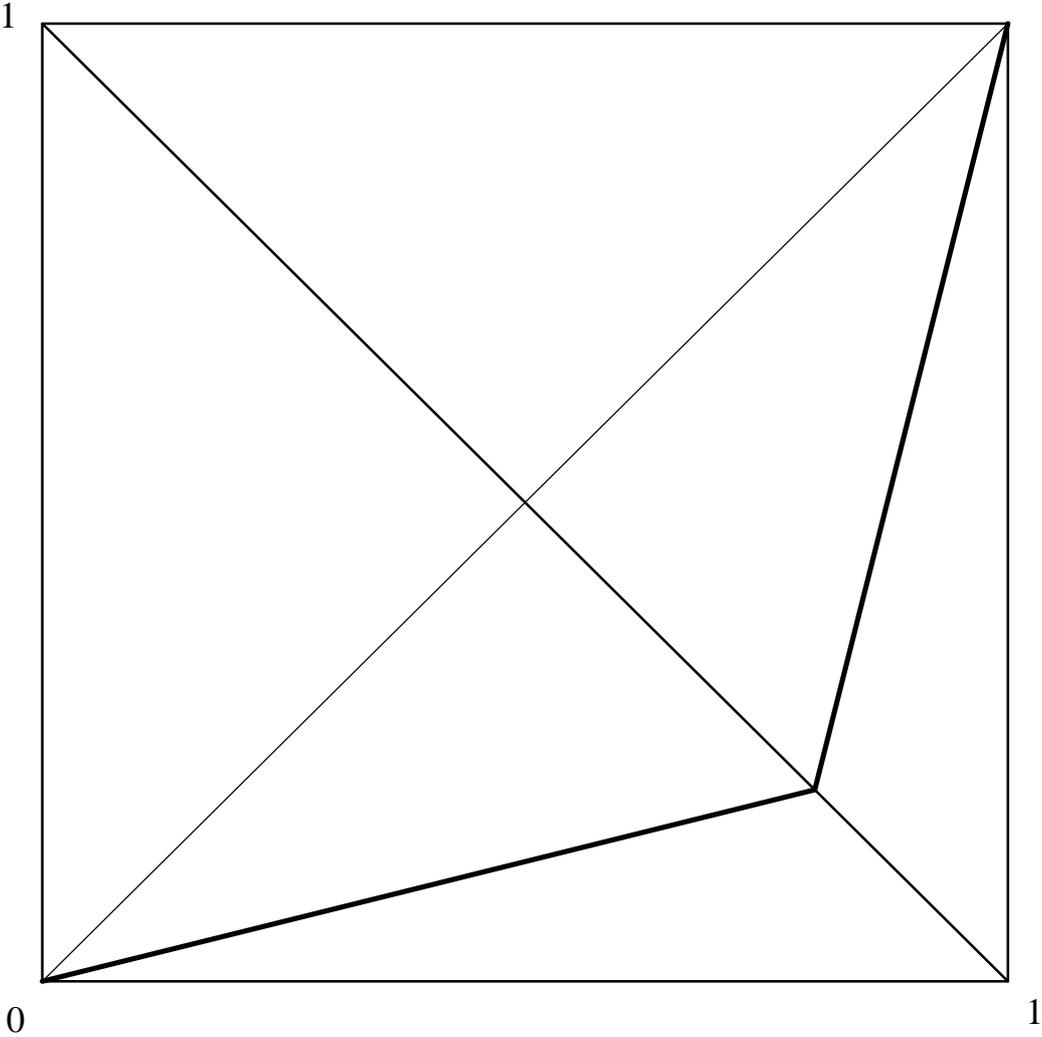


Figure 1. Pure "80-20 law"

Tables and Fig ii

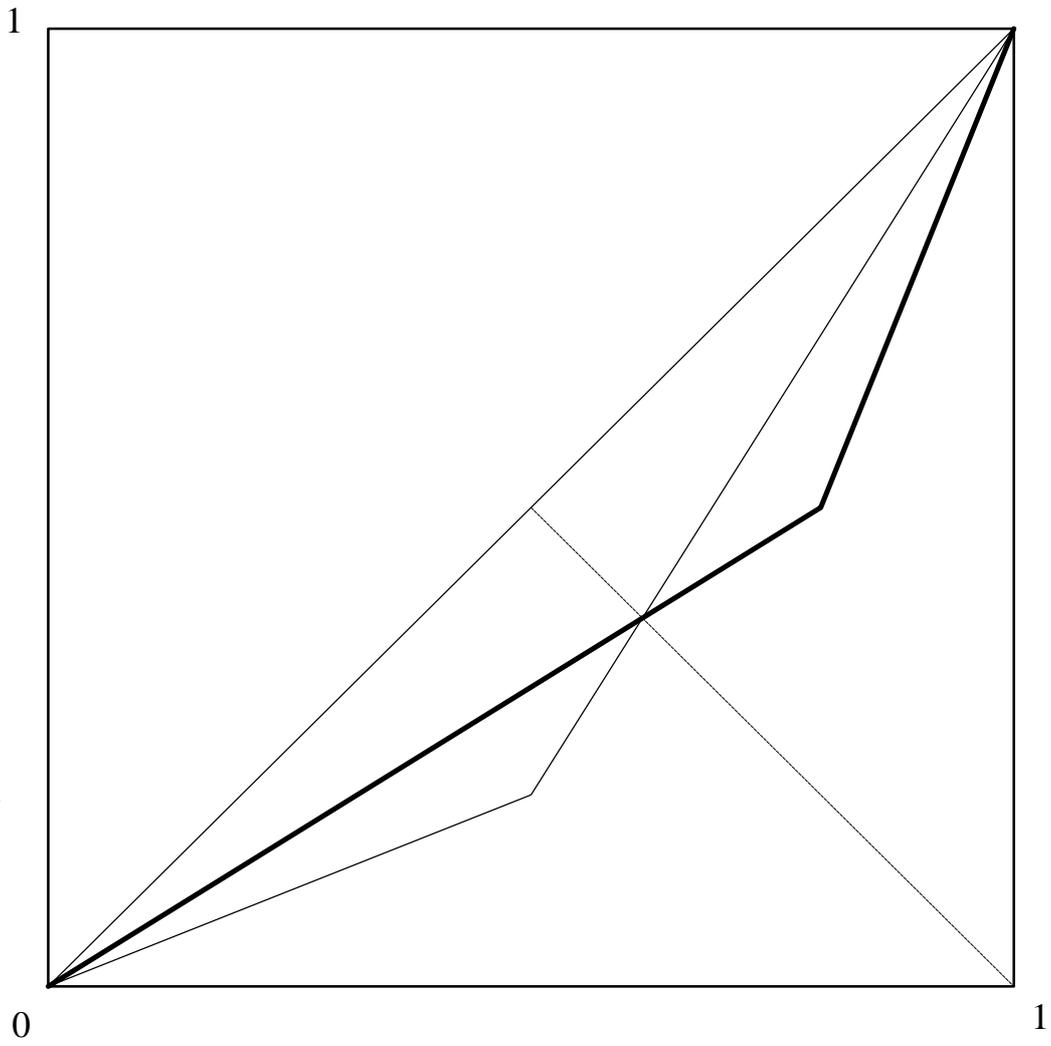


Figure 2. The preferable distribution, skewed to the right (bold line)

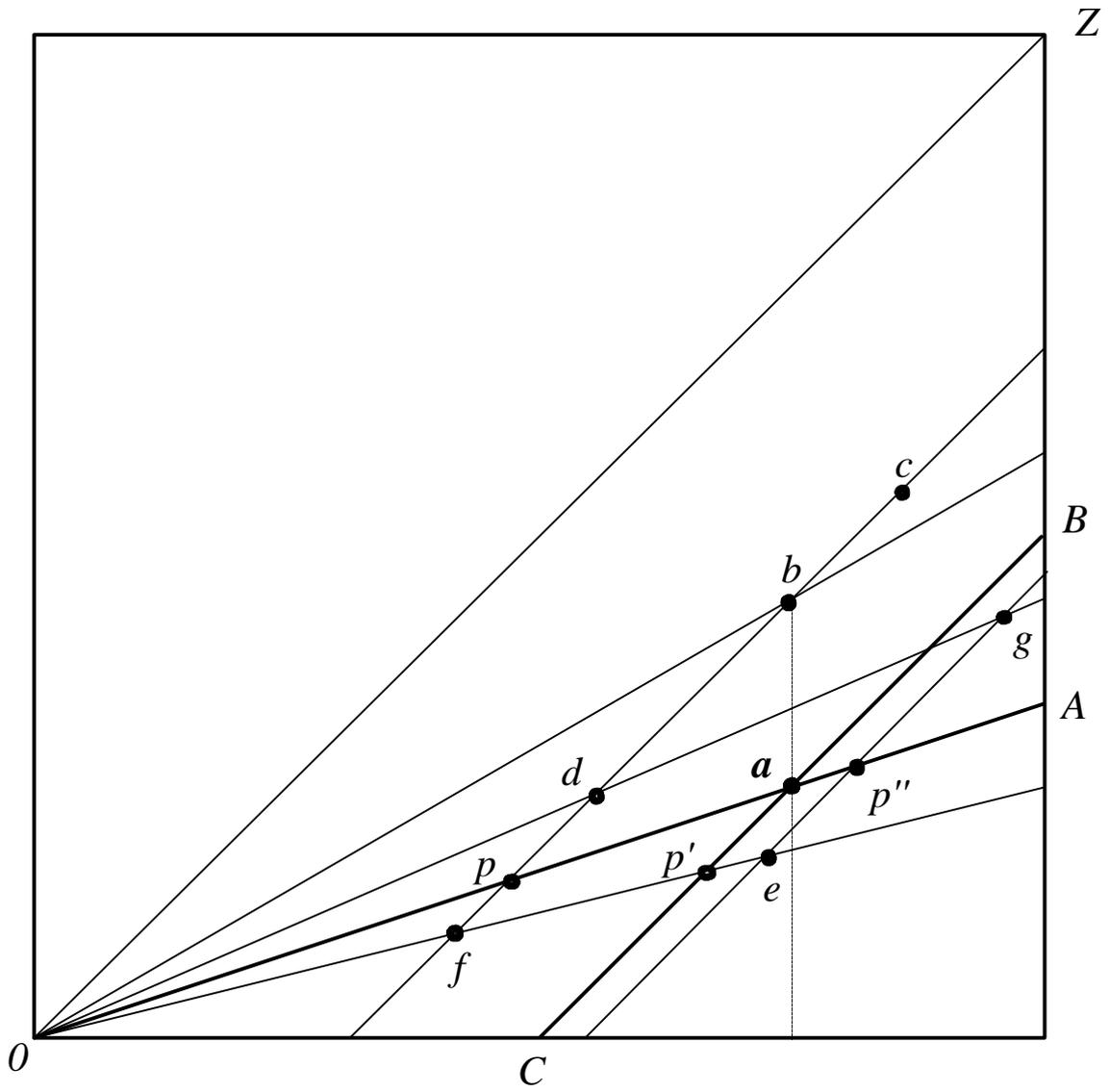


Figure 3. Comparison at a reference kink, a

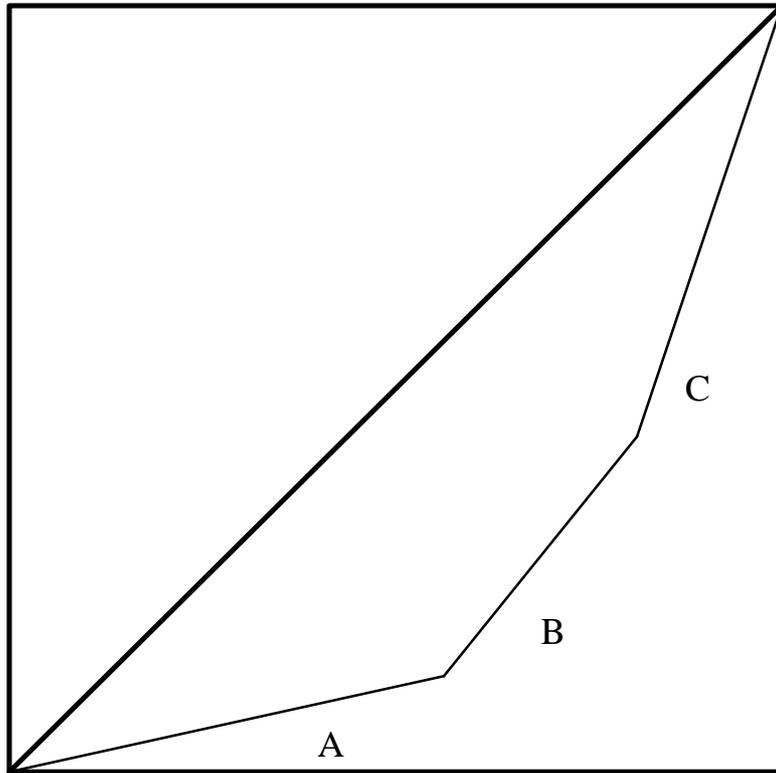
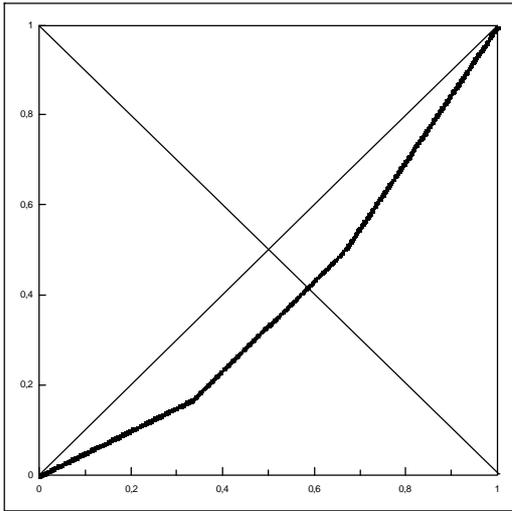
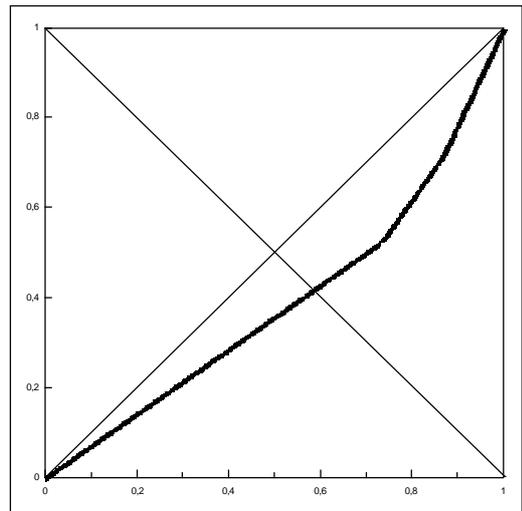


Figure 4. The "A-B-C law".

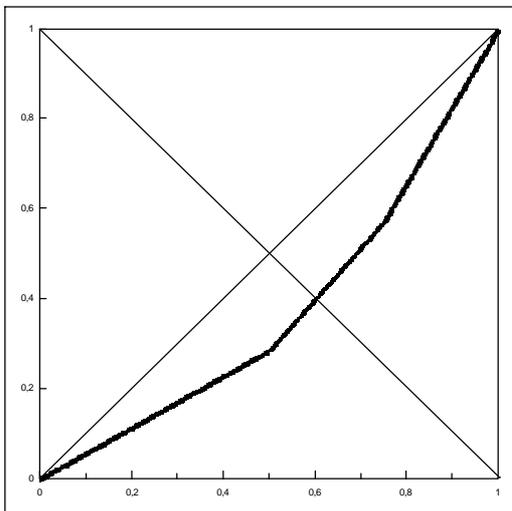
Tables and Fig v



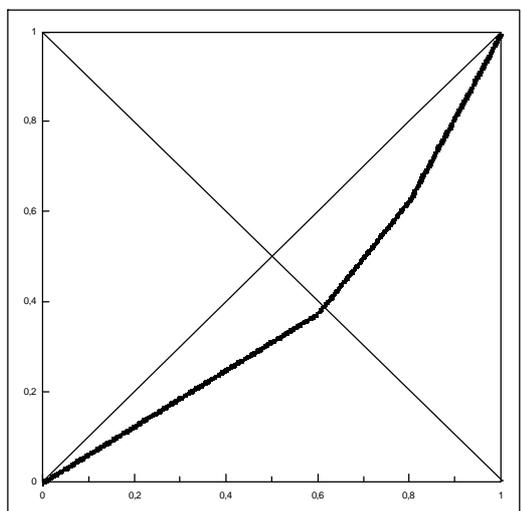
Distribution 1: $I_G = 0.222$, $URP = 0.500$



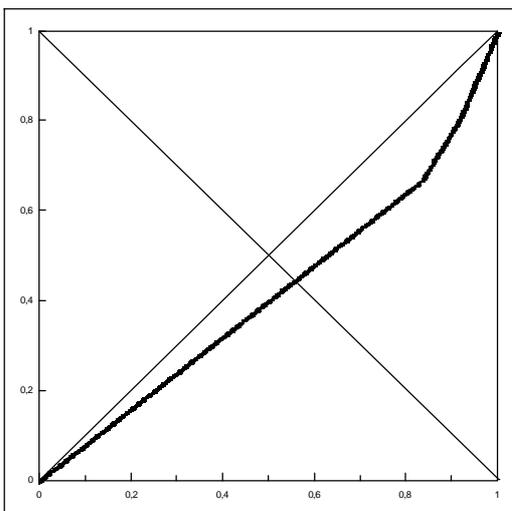
Distribution 2: $I_G = 0.222$, $URP = 0.714$



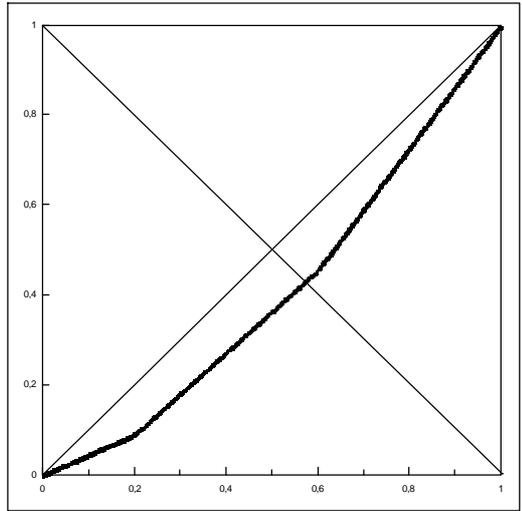
Distribution 3: $I_G = 0.250$, $URP = 0.571$



Distribution 4: $I_G = 0.250$, $URP = 0.625$



Distribution 5: $I_G = 0.172$, $URP = 0.800$



Distribution 6: $I_G = 0.182$, $URP = 0.455$

Figure 5.

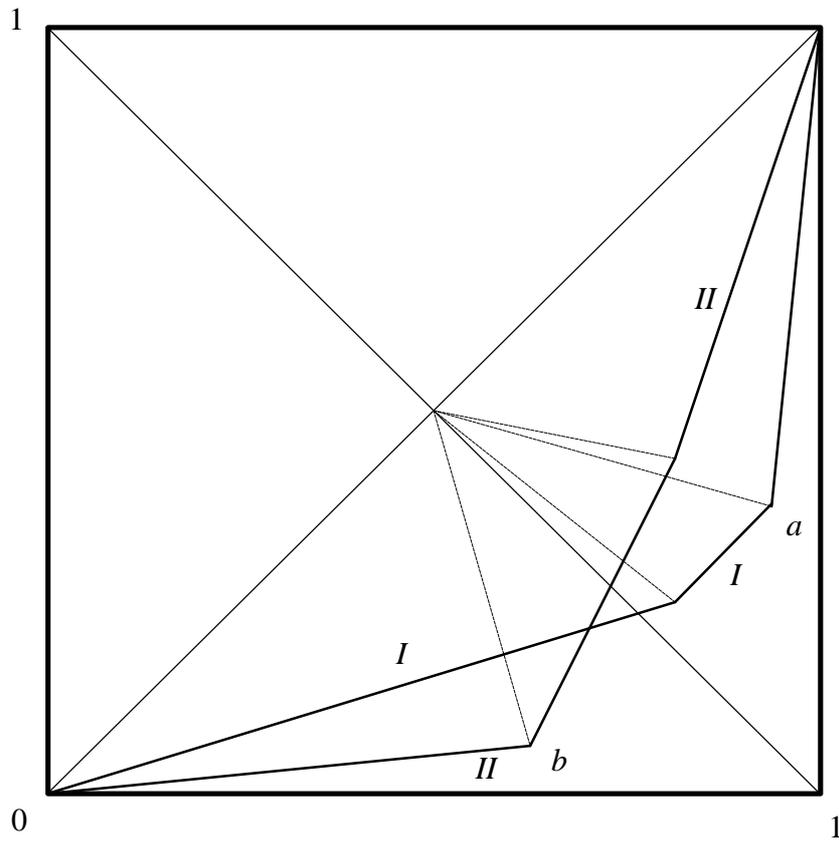


Figure 6. Concept of main kink:

a is the main kink of curve I and b is the main kink of curve II

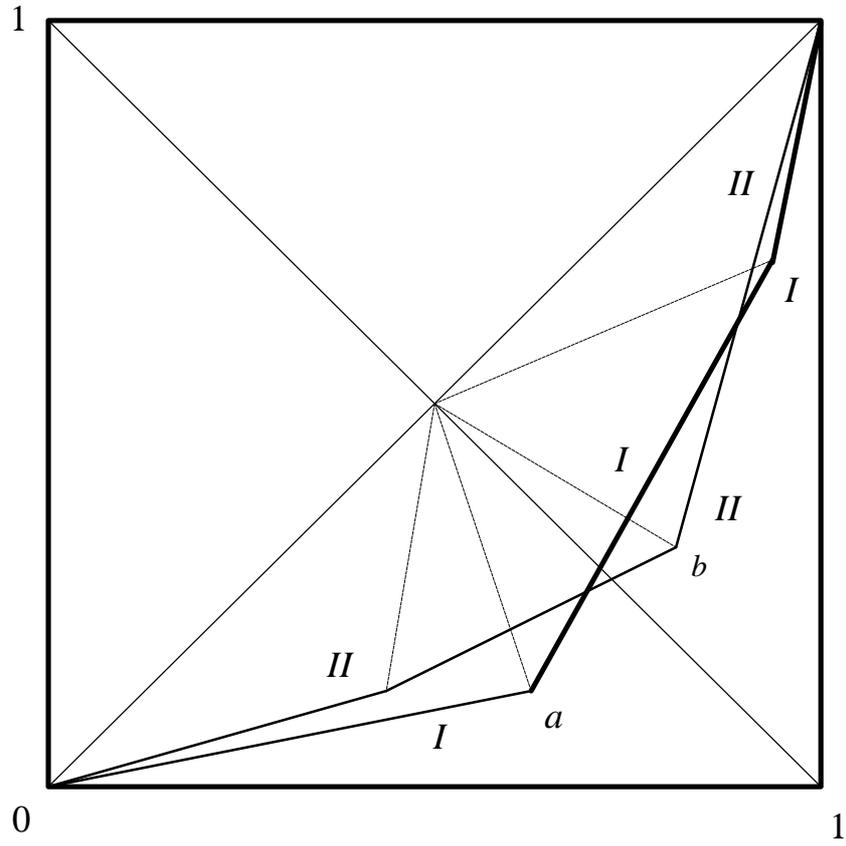


Figure 7.

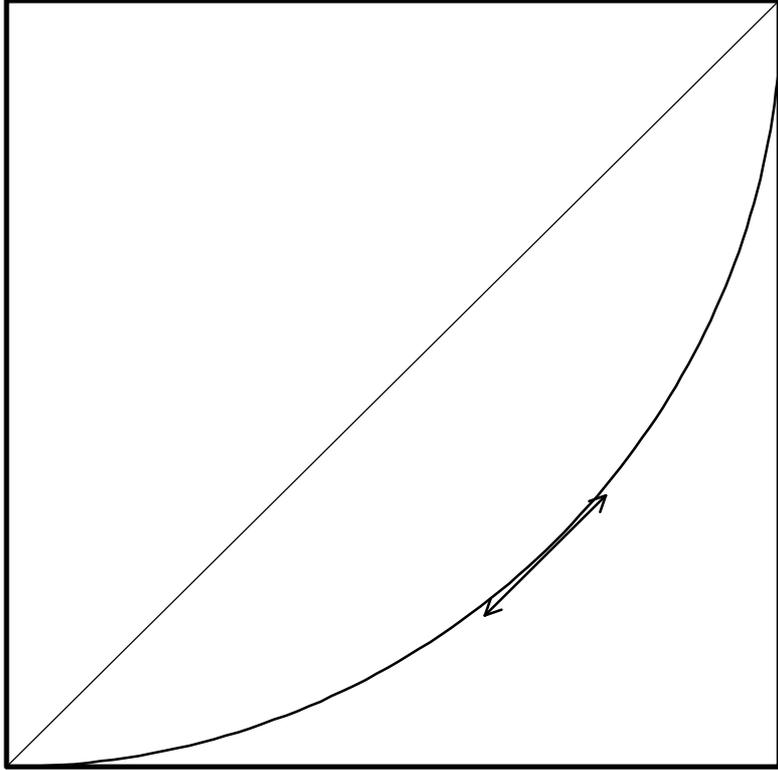


Figure 8. Replacement of the kink.

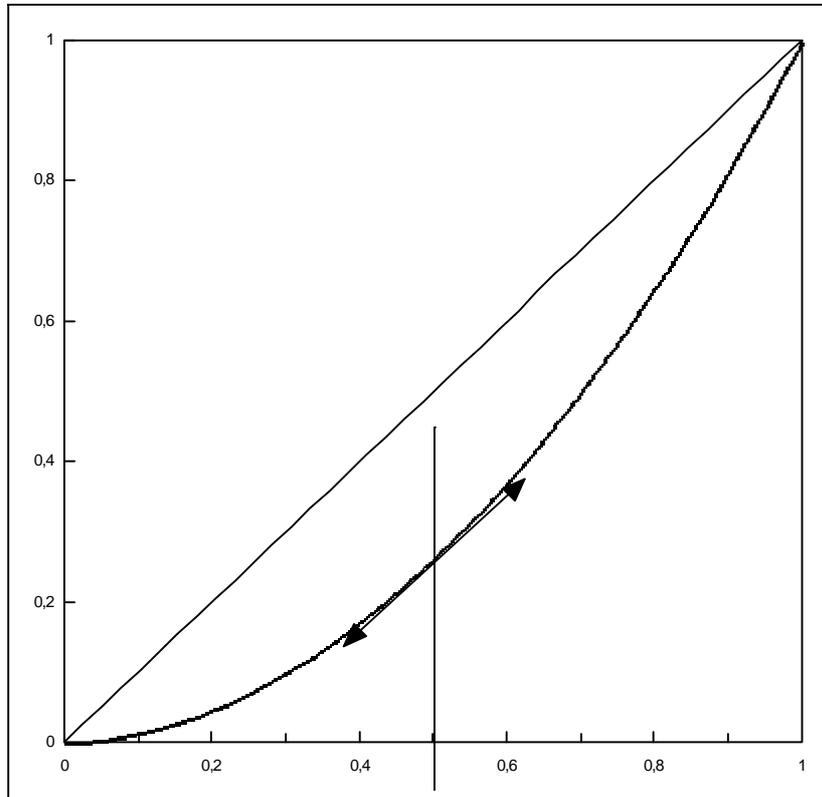


Figure 9. Uniform distribution: curve skewed to the left

Tables and Fig x

Revenues	Number of individuals
1	40
4	10

Revenues	Number of individuals
1	25
4	25

Table 1.

Tables and Fig xi

Not skewed		Right skewed		Left skewed	
Revenues in %	% of indiv.	Revenues in %	% of indiv.	Revenues in %	% of indiv.
20	80	25	85	15	75
80	20	75	15	85	25

Table 2.

Tables and Fig xii

Right skewed		Left skewed	
Revenues in %	% of indiv.	Revenues in %	% of indiv.
40	100	0	60
60	0	100	40

Table 3. Limit cases.

Tables and Fig xiii

Distribution 1		Distribution 2	
$I_G = 0.222$ $URP = 0.500$		$I_G = 0.222$ $URP = 0.714$	
prop. of revenues	prop. of indiv.	prop. of revenues	prop. of indiv.
1/6	1/3	0.5238	0.733
1/3	1/3	0.1905	0.133
1/2	1/3	0.2857	0.133

Distribution 3		Distribution 4	
$I_G = 0.250$ $URP = 0.571$		$I_G = 0.250$ $URP = 0.625$	
prop. of revenues	prop. of indiv.	prop. of revenues	prop. of indiv.
0.2857	1/2	3/8	3/5
0.2857	1/4	1/4	1/5
0.4286	1/4	3/8	1/5

Distribution 5		Distribution 6	
$I_G = 0.172$ $URP = 0.800$		$I_G = 0.182$ $URP = 0.455$	
prop. of revenues	prop. of indiv.	prop. of revenues	prop. of indiv.
2/3	5/6	1/11	1/5
2/15	1/12	4/11	2/5
1/5	1/12	6/11	2/5

Table 4.