

Poverty and Equity: Theory and Estimation

by

Jean-Yves Duclos

**Département d'économie and CRÉFA,
Université Laval, Canada**

Preliminary version

This text is in large part an output of the MIMAP training programme financed by the International Development Research Center of the Government of Canada. The underlying research was also supported by grants from the Social Sciences and Humanities Research Council of Canada and from the Fonds FCAR of the Province of Québec. I am grateful to Abdelkrim Araar and Nicolas Beaulieu for their excellent research assistance.

Corresponding address:

Jean-Yves Duclos, Département d'économie, Pavillon de Sève, Université Laval, Québec, Canada, G1K 7P4; Tel.: (418) 656-7096; Fax: (418) 656-7798; Email: jduc@ecn.ulaval.ca

January 2002

Contents

I	Introduction	5
1	Well-being and poverty	6
1.1	The welfarist approach	6
1.2	Non-welfarist approaches	8
1.2.1	Basic needs and functionings	8
1.2.2	Capabilities	10
1.3	A graphical illustration	11
1.3.1	Exercises	14
1.4	Practical measurement difficulties	14
2	Poverty measurement and public policy	16
2.1	Welfarist and non-welfarist policy implications	17
II	Measuring poverty and equity	20
3	Notation	21
3.1	Continuous distributions	21
3.2	Discrete distributions	23
4	The measurement of inequality and social welfare	25
4.1	Lorenz curves	25
4.2	Gini indices	27
4.3	Social welfare	32
4.3.1	Atkinson indices	35
4.3.2	S-Gini indices	36
4.4	Decomposable indices of inequality	37
4.5	Other popular indices of inequality	39
5	Aggregating and comparing poverty	40
5.1	Cardinal versus ordinal comparisons	40
5.2	Aggregating poverty	41
5.2.1	The EDE approach	41
5.2.2	The poverty gap approach	42
5.3	Group-decomposable poverty indices	46

5.4	Poverty and inequality	47
5.5	Poverty curves	48
5.6	S-Gini poverty indices	48
5.7	The normalization of poverty indices	49
5.8	Decomposing differences in poverty	50
6	Estimating poverty lines	53
6.1	Absolute and relative poverty lines	53
6.2	Social exclusion and relative deprivation	54
6.3	Estimating poverty lines	56
6.3.1	Cost of basic needs	56
6.3.2	Cost of food needs	56
6.3.3	Non-food poverty lines	59
6.3.4	Food energy intake	61
6.3.5	Illustration for Cameroon	63
6.3.6	Relative and subjective poverty lines	64
7	The measurement of progressivity, equity and redistribution	68
7.1	Taxes and concentration curves	68
7.2	Indices of concentration	70
7.3	Progressivity comparisons	72
7.3.1	Deterministic tax and benefit systems	72
7.3.2	General tax and benefit systems	73
7.4	Reranking and horizontal inequity	75
7.5	Redistribution	79
7.6	Indices of progressivity and redistribution	80
8	Issues in the empirical measurement of well-being and poverty	82
8.1	Survey issues	82
8.2	Income versus consumption	84
8.3	Price variability	85
8.4	Household heterogeneity	89
8.4.1	Equivalence scales	89
8.4.2	Household decision-making and within-household inequality	93
III	Ethical robustness of poverty and equity comparisons	95

9	Poverty dominance	96
9.1	Primal approach	100
9.2	Dual approach	104
9.3	Assessing the limits to dominance	105
10	Inequality dominance	107
10.1	Primal approach	108
10.2	Dual approach	109
10.3	Inequality and progressivity	109
11	Welfare dominance	111
11.1	Primal approach	112
11.2	Dual approach	113
IV	Poverty and equity: policy design and assessment	115
12	Poverty alleviation: policy and growth	116
12.1	Measuring the benefits of public spending	116
12.2	Checking the distributive effect of public expenditures	116
12.3	The impact of targeting and public expenditure reforms on poverty	117
12.3.1	Group-targeting a constant amount	119
12.3.2	Inequality-neutral targeting	120
12.3.3	Price changes	121
12.3.4	Tax/subsidy policy reform	124
12.3.5	Income-component and sectoral growth	126
12.4	Overall growth elasticity of poverty	126
12.5	The Gini elasticity of poverty	128
13	The impact of policy and growth on inequality	129
13.1	Growth, tax and transfer policy, and price shocks	129
13.2	Tax and subsidy reform	131
V	Estimation and inference for distributive analysis	133
14	Non parametric estimation for distributive analysis	134
14.1	Density estimation	134
14.1.1	Univariate density estimation	134

14.1.2	Statistical properties of kernel density estimation	136
14.1.3	Choosing a window width	137
14.1.4	Multivariate density estimation	139
14.1.5	Simulating from a nonparametric density estimate	139
14.2	Non-parametric regression	141
15	Symbols	144
16	References	151
17	Graphs and tables	164

Part I
Introduction

1 Well-being and poverty

The assessment of well-being for poverty analysis is traditionally characterized according to two main approaches, which, following Ravallion (1994), we will term the welfarist and the non-welfarist approaches. The first approach tends to concentrate in practice mainly on comparisons of "economic well-being", which we will also call "standard of living" or "income" (for short). As we will see, this approach has strong links with traditional economic theory, and it is also widely used by economists in the operations and research work of organizations such as the World Bank, the International Monetary Fund, and Ministries of Finance and Planning of both developed and developing countries. The second approach has historically been advocated mainly by social scientists other than economists and partly in reaction to the first approach. This second approach has nevertheless also been recently and increasingly advocated by economists and non-economists alike as a sound multidimensional complement to the classical standard of living approach.

1.1 The welfarist approach

The welfarist approach is strongly anchored in classical micro-economics, where, in the language of economists, "welfare" or "utility" are generally key in accounting for the behavior and the well-being of individuals. Classical micro-economics usually postulates that individuals are rational and that they can be presumed to be the best judges of the sort of life and activities which maximize their utility and happiness. Given their initial endowments (including time, land and physical and human capital), individuals make production and consumption choices using their set of preferences over bundles of consumption and production activities, and taking into account the available production technology and the consumer and producer prices that prevail in the economy. Under these assumptions and constraints, a process of individual and rational free choice will maximize the individuals' utility; under additional assumptions (including that markets are competitive, that agents have perfect information, and that there are no externalities – assumptions that are thus very restrictive), a society of individuals all acting independently under this freedom of choice process will also lead to an outcome known as Pareto-efficient, in that no one's utility could be further improved by government intervention without decreasing someone else's utility.

Underlying the welfarist approach to poverty, there is a premise that good note should be taken of the information revealed by individual behavior when

it comes to assessing poverty. This says, more particularly, that the assessment of someone's well-being should be consistent with the ordering of preferences revealed by that person's free choices. For instance, a person could be observed to be poor by the total consumption or income standard of a poverty analyst. That same person could nevertheless be able (*i.e.*, have the working capacity) to be non-poor. This could be revealed by the observation of a deliberate and free choice on the part of the individual to work and consume little, when the capability to work and consume more nevertheless exists. By choosing to spend little (possibly for the benefit of greater leisure), the person reveals that he is happier than if he worked and spent more. Although he could be considered poor by the standard of a (non-welfarist) poverty analyst, a welfarist judgement should conclude that this person is not poor. As we will discuss later, this can have important implications for the design and the assessment of public policy.

A pure welfarist approach faces important practical problems. To be operational, pure welfarism requires the observation of sufficiently informative revealed preferences. This is rarely the case, however. For instance, for someone to be declared poor or not poor, it is not enough to know that person's current characteristics and living standard status, but it must also be inferred from that person's actions whether he judges his utility status to be above a certain utility poverty level. Another – more fundamental – problem with the pure welfarist approach is the need to assess levels of utility or "psychic happiness". How are we to measure the actual pleasure derived from experiencing economic well-being? Moreover, it is highly problematic to attempt to compare that level of utility across individuals – it is well known that such a procedure poses serious ethical problems. Preferences are heterogeneous, personal characteristics, needs and enjoyment abilities are diverse, households differ in size and composition, and prices vary across time and space. Besides, it is not clear that we should accept as ethically significant the actual level of utility felt by individuals. Why should a difficult-to-satisfy rich person be judged less happy than an easily-contented poor person? That is, why should a "grumbling rich" be judged "poorer" than a "contented peasant" (see Sen (1983), p.160)?

Hence, welfarist comparisons of poverty almost invariably use imperfect but observable proxies for utilities, such as income or consumption. These money-metric indicators are often adjusted for differences in needs, prices, and household sizes and compositions, but they clearly do remain far-from-perfect indicators of utility and well-being. Indeed, economic theory tells us little about how to use consumption or income to make consistent interpersonal comparisons of well-being. Besides, the consumption and income proxies are rarely able to take full

account of the role for well-being of public goods and non-market commodities, such as safety, liberty, peace, health. In principle, such commodities can be valued using reference or "shadow" prices. In practice, this is very difficult to do accurately and consistently.

1.2 Non-welfarist approaches

1.2.1 Basic needs and functionings

There are two major non-welfarist approaches, the basic-needs approach and the capability approach. The first approach focuses on the need to attain some basic multidimensional outcomes that can be observed and monitored relatively easily. These outcomes are usually (explicitly or implicitly) linked with the concept of functionings, a concept developed in Amartya Sen's influential work:

Living may be seen as consisting of a set of interrelated 'functionings', consisting of beings and doings. A person's achievement in this respect can be seen as the vector of his or her functionings. The relevant functionings can vary from such elementary things as being adequately nourished, being in good health, avoiding escapable morbidity and premature mortality, *etc.*, to more complex achievements such as being happy, having self-respect, taking part in the life of the community, and so on (Sen (1997), p.39).

In this view, functionings can be understood to be *constitutive* elements of well-being. The functioning approach would generally not attempt to compress these elements into a single dimension such as utility or happiness. Utility or happiness is viewed as a single and reductive aggregate of functionings, which are multidimensional in nature. The functioning approach focuses instead on multiple specific and separate outcomes, such as the enjoyment of a particular type of commodity consumption, being healthy, literate, well-clothed, well-housed, not in shape, *etc.*

The functioning approach is closely linked with the well-known basic needs approach, and the two are often difficult to distinguish in their practical application. Functionings, however, are not synonymous with basic needs. Basic needs can be understood as the physical inputs that are usually required for individuals to achieve some functionings. Hence, basic needs are usually defined in terms of means rather than outcomes, for instance, as living in the proximity of providers of health care services (but not necessarily being in good health), as the number

of years of achieved schooling (not necessarily as being literate), as living in a democracy (but not necessarily as participating in the life of the community), and so on. In other words,

Basic needs may be interpreted in terms of minimum specified quantities of such things as food, shelter, water and sanitation that are necessary to prevent ill health, undernourishment and the like (Streeten *et al.* (1981)).

Unlike functionings, which can be commonly defined for all individuals, the specification of basic needs depends on the characteristics of individuals and of the societies in which they live. For instance, the basic commodities required for someone to be in good health and not to be undernourished will depend on the climate and on the physiological characteristics of individuals. Similarly, the clothes necessary for one not to feel ashamed will depend on the norms of the society in which he lives, and the means necessary to travel, on whether he is handicapped or not. Hence, although the fulfillment of basic needs is an important element in assessing whether someone has achieved some functionings, this assessment must also use information on one's characteristics and socio-economic environment. Human diversity is such that equality in the space of basic needs generally translates into inequality in the space in functionings.

Whether unidimensional or multidimensional in nature, most applications of both the welfarist and the non-welfarist approaches to poverty measurement do recognize the role of needs and of socio-economic environments in achieving well-being. Streeten *et al.* (1981) and others have nevertheless argued that the basic needs approach is less abstract than the welfarist approach in recognizing that role. As mentioned above, assessing the fulfillment of basic needs it can also be seen as a useful practical and operational step towards appraising the achievement of the more abstract "functionings".

Clearly, however, there are important degrees in the multidimensional achievements of basic needs and functionings. For instance, what does it mean precisely to be "adequately nourished"? Which degree of nutrition adequacy is relevant for poverty assessment? Should the means needed for the adequate nutrition functioning only allow for the simplest possible diet and for highest nutritional efficiency? These problems also crop up in the estimation of poverty lines in the welfarist approach. A multidimensional approach extends them to several dimensions. In addition, how ought we to understand such functionings as the functioning of self-respect? The appropriate width and depth of the concept of basic needs and

functionings is admittedly ambiguous, as there are degrees of functionings which make life enjoyable in addition to being purely sustainable or satisfactory. Finally, could some of the dimensions be substitutes in the attainment of a given degree of well-being? That is, could it be that one could do with lower needs and functionings in some dimensions if he has high achievements in the other dimensions? Such possibilities of substitutability are generally ignored (and are indeed hard to identify precisely) in the multidimensional non-welfarist approaches.

1.2.2 Capabilities

A second alternative to the welfarist approach is called the capability approach, also pioneered and advocated in the last two decades by the work of Sen. The capability approach is defined by the *capacity* to achieve functionings, as defined above. In Sen's words (1997),

the capability to function represents the various combinations of functionings (beings and doings) that the person can achieve. Capability is, thus, a set of vectors of functionings, reflecting the person's freedom to lead one type of life or another. (p.40)

What matters for the capability approach is the ability of an individual to function well in society; it is *not* the functionings actually attained by the person. Having the capability to achieve "basic" functionings is the source of freedom to live well, and is thereby sufficient in the capability approach for one not to be poor or deprived.

The capability approach thus distances itself from achievements of specific outcomes or functionings. In this, it imparts considerable value to freedom of choice: a person will not be judged poor even if he chooses not to achieve some functionings, so long as he would be able to attain them if he so chose. This distinction between outcomes and the capability to achieve the outcomes also recognizes the importance of preference diversity and individuality in determining functioning choices. It is, for instance, not everyone's wish to be well-clothed or to participate in society, even if the capability is present.

An interesting example of the distinction between fulfilment of basic needs, functioning achievement and capability is given by Townsend's (1979, Table 6.3) deprivation index. The deprivation index is built from answers to questions such as whether someone "has not had an afternoon or evening out for entertainment in the last two weeks", or "has not had a cooked breakfast most days of the week". It

may be, however, that one chooses deliberately not to go out for entertainment (he prefers to watch television), or that he chooses not to have a cooked breakfast (because he does not have time to prepare it), although he does have the capacity to do both. That person therefore achieves the functioning of being entertained without meeting the basic need of going out once a fortnight, and does have the capacity to achieve the functioning of having a good breakfast, although he chooses not to.

The difference between the capability and the functioning or basic needs approach is in fact somewhat analogous to the difference between the use of income and consumption as indicators of living standards. Income shows the capability to consume, and "consumption functioning" can be understood as the outcome of the exercise of that capability. There is consumption only if a person chooses to enact his capacity to consume out a given income. In the basic needs and functioning approach, deprivation comes from a lack of direct consumption or functioning experience; in the capability approach, poverty arises from the lack of incomes and capabilities, which are imperfectly related to the actual functionings achieved.

Although the capability set is multidimensional, it thus exhibits a parallel with the unidimensional income indicator, whose size determines the size of the "budget set":

Just as the so-called 'budget set' in the commodity space represents a person's freedom to buy commodity bundles, the 'capability set' in the functioning space reflects the person's freedom to choose from possible livings (Sen (1997, p. 40)).

This illustrates further the fundamental distinction between the space of achievements, the extents of freedoms and capabilities, and the resources required to generate these freedoms and to attain these achievements.

1.3 A graphical illustration

To illustrate the links between the main approaches to assessing poverty, consider Figure 1. Figure 1 shows in four quadrants the links between income, consumption of two commodities – transportation and clothing goods – and the functionings associated to each of these two goods. The northeast quadrant shows a typical two-good budget set for the two goods T and C , namely, for transportation and clothing respectively, and with a budget constraint $Y1$. The curve $U1$ shows the utility indifference curve along which the consumer chooses his preferred commodity bundle, which is here located at point A .

The southeastern and the northwestern quadrants then transform the consumption of goods T and C into associated functionings F_T and F_C . This is done through the Functioning Transformation Curves TC_T and TC_C , for transformation of consumption of T and C into transportation and clothing functionings. The curves TC_T and TC_C appear in the northwest and the southeast quadrants respectively. These curves thus bring us from the northeastern space of commodities, $\{C, T\}$, into the southwestern space of functionings, $\{F_C, F_T\}$. Using these transformation functions, we can draw a budget constraint $S1$ in the space of functionings from the traditional commodity budget constraint, $Y1$. Since the consumer chooses point A in the space of commodities, he enjoys B 's combination of functionings. But all of the functionings within the constraint $S1$ can also be attained by the consumer. The triangular area between the origin and the line $S1$ thus represents the individual's capability set. It is the set of functionings which he is able to attain.

Now assume that functioning thresholds of z_C and z_T must be exceeded (or must be potentially exceeded) for one not to be considered poor by the non-welfarist approaches. Given the transformation functions TC_T and TC_C , a budget constraint $Y1$ makes the individual *capable* of not being poor in the functioning space. But this does not guarantee that the individual will choose a combination of functionings that will exceed z_C and z_T : this will also depend on the individual's preferences. At point A , the functionings achieved are above the minimum functioning threshold fixed in each dimension. Other points within the capability set would also surpass the functioning thresholds: these points are shown in the shaded triangle to the northeast of point B . Since part of the capability set allows the individual to be non-poor in the space of functionings, the capability approach would also declare the individual not to be poor.

So would conclude, too, the functioning approach since the individual *chooses* functionings above z_C and z_T . Such a concordance does not always have to prevail, however. Consider Figure 2. The commodity budget set and the Functioning Transformation Curves have not changed, so that the capability set has not changed either. But there has been a shift of preferences from $U1$ to $U2$, so that the individual now prefers point D to point A , and also prefers to consume less clothing than before. This makes his preferences for functionings to be located at point E , thus failing to exceed the minimum clothing functioning z_C required. Hence, the person would be considered non poor by the capability approach, but poor by the functioning approach. Whether an individual with preferences $U2$ is really poorer than one with preferences $U1$ is debatable, of course, since the two have exactly the same "opportunity sets", that is, have access to exactly the same

commodity and capability sets.

An important message of the capability approach is that two persons with the same commodity budget set can face different capability sets. This is illustrated in Figure 3, where the Functioning Transformation Curve for transportation has shifted from TC_T to TC'_T . This may be due to the presence of a handicap, which makes it more costly in transportation expenses to generate a given level of transportation functioning (disabled persons would need to expend more to go from one place to another). This shift of the TC_T curve moves the capability constraint to $S1'$ and thus contracts the capability set. With the handicap, there is no point within the new capability set that would surpass both functioning thresholds z_C and z_T . Hence, the person is deemed poor by the capability approach and (necessarily so) by the functioning approach. Whether the welfarist approach would also declare the person to be poor would depend on whether it takes into account the differences in needs implied by the difference between the TC_T and the TC'_T curves.

For the welfarist approach to be reasonably consistent with the functioning and capability approaches, it is thus essential to consider the role of transformation functions such as the TC curves. If this is done, we may (in our simple illustration at least) assess a person's poverty status either in the commodity or in the functioning space. In other words, we may determine whether someone is capability-poor either from observing his consumption of different goods, or from observing his attained functionings.

To see this, consider Figure 4. Figure 4 is the same as Figure 1 except for the addition of the commodity budget constraint $Y2$ which shows the minimum consumption level needed for one not to be poor according to the capability approach. According to the capability approach, the capability set must contain at least one combination of functionings above z_C and z_T , and this condition is just met by the capability constraint $S2$ that is associated with the commodity budget $Y2$. Hence, to know whether someone is poor according to the capability approach, we may simply check whether his commodity budget constraint lies below $Y2$.

Even if the actual commodity budget constraint lies above $Y2$, the individual may well choose a point outside the non-poor functioning set, as we discussed above in the context of Figure 2. Clearly then, the minimum total consumption needed for one to be non-poor according to the functioning or basic needs approach generally exceeds the minimum total consumption needed for one to be non-poor according to the capability approach. More problematically, this minimum total consumption depends in principle on the preferences of the individuals. On Figure 2, for instance, we saw that the individual with preference $U2$ was con-

sidered poor by the functioning approach, although another individual with the same budget and capability sets was considered non-poor by the same approach.

1.3.1 Exercises

1. Show on a figure such as Figure 1 the impact of an increase in the price of the transportation commodity on the commodity and capability constraints.
2. On a figure such as Figure 4, show the minimal commodity budget set that ensures that the person
 - (a) is just able to attain one of the two minimum levels of functionings z_C or z_T ;
 - (b) chooses a combination of functionings such that one of them exceeds the corresponding minimum level of functionings z_C or z_T ;
 - (c) is just able to attain both minimum levels of functionings z_C and z_T ;
 - (d) chooses a combination of functionings such that both exceed the corresponding minimum level of functionings z_C and z_T .
 - (e) How do these four minimal commodity constraints compare to each other?

1.4 Practical measurement difficulties

How are we to measure capabilities? Unless a person chooses to enact them in the form of functioning achievements, capabilities are not easily inferred. Achievement of all basic functionings implies non-deprivation in the space of all capabilities; but a failure to achieve all basic functionings does not imply capability deprivation. This makes the monitoring of functioning and basic needs an imperfect tool for the assessment of capability deprivation. Besides, and as for basic needs, there are clearly degrees of capabilities, some basic and some wider.

Non-welfarist (capability and basic needs) approaches to poverty measurement also suffer from some comparability problems. This is because they typically generate multidimensional qualitative poverty criteria: their fulfilment typically takes a simple dichotomic yes/no form. It is unlikely that true well-being is such a dichotomic and discontinuous function of achievement and capabilities. Indeed, for most of the functionings assessed empirically, there are degrees of achievement, such as for being healthy, literate, living without shame, *etc...* It

is important to take into account the varying degrees of poverty in assessing and comparing the intensity of poverty. Besides, how should we assess adequately the degree of poverty of someone who has the capability to achieve two functionings out of three, but not the third? Is that person necessarily "better off" than someone who can achieve only one, or even none of them? Are all capabilities of equal importance when we assess well-being?

The multidimensionality of the non-welfarist criteria also translates into greater implementation difficulties than for the usual proxy indicators of the welfarist approach. In the welfarist approach, the size of the multidimensional budget is ordinarily summarized by income or total consumption, which can be thought of as a unidimensional indicator of freedom. A similar transformation into a unidimensional indicator is more difficult with the capability and basic needs approaches. One possibility solution is to use "efficiency-income units reflecting command over capabilities rather than command over goods and services" (Sen (1984, p.343), as we illustrated above when discussing Figure 4. This, however, is practically difficult to do, since command over many capabilities is hard to translate in terms of a single indicator, and since the "budget units" are hardly comparable across functionings such as well-nourishment, literacy, feeling self-respect, and taking part in the life of the community. On Figure 4, anyone with an income below Y_2 would be judged capability-poor. But by how much does poverty vary among these capability-poor? A natural measure would be a function of the budget constraint. It is more difficult to make such measurements and comparisons in the capability set.

Furthermore, although there are many different combinations of consumption and functionings that are compatible with a unidimensional money-metric poverty threshold, the welfarist approach will generally not impose multidimensional thresholds. For instance, the welfarist approach will usually not require for one not to be poor that both food and non-food expenditures be larger than their respective food and non-food poverty lines. As indicated above, this simplifies the identification of the poor and the analysis of poverty.

2 Poverty measurement and public policy

The measurement of well-being and poverty plays a central role in the discussion of public policy and safety nets in particular. It is used, among other things, to identify the poor and the non-poor, to design optimal poverty relief schemes, to estimate the errors of exclusion and inclusion in the set of the poor (also known as Type I and Type II errors), and to assess the equity of poverty alleviation policy. How many of the poor, for instance, are excluded from safety net programmes? Is it the poorest of the poor who benefit most from public policy? Would a different sort of poverty alleviation policy reduce deprivation further?

An important example of the central role of poverty measurement in the setting of public policy is the optimal selection of safety net targeting indicators. The theory of optimal targeting suggests that it will commonly be best to target individuals on the basis of indicators that are as easily observable and as exogenous as possible, while being as correlated as possible with the true poverty status of the individuals. Indicators that are not readily observable by programme administrators are of little practical value. Indicators that can be changed effortlessly by individuals will be distorted by the presence of the programme, and will lose their poverty-informative value. Whether available indicators are sufficiently correlated with the deprivation of individuals in a population is given by a poverty profile. The value of this profile will naturally be highly dependent on the particular assumptions and the approach used to measure well-being and poverty.

Estimation of the errors of inclusion and exclusion of the poor is also a product of poverty profiling and measurement. These errors are central in the trade-off involved in choosing a wide coverage of the population – at relatively low administrative and efficiency costs – and a narrower coverage – with more generous forms of support for the fewer beneficiaries. However, as Van de Walle (1998) puts it, a narrower coverage of the population, with presumably smaller errors of inclusion of the non-poor, does not inevitably lead to a more equitable treatment of the poor:

Concentrating solely on errors of leakage to the non-poor can lead to policies which have weak coverage of the poor (Van de Walle (1998, p.366)).

The terms of this trade-off are again given by a poverty assessment exercise.

Another lesson of optimal redistribution theory is that it is ordinarily better to transfer resources from groups with a high level of average well-being to those

with a lower one. What matters even more, however, is the distribution of well-being within each of the groups. For instance, equalising mean well-being across groups does not usually eliminate poverty since there generally exist within-group inequalities. Even within the richer group, for instance, there normally will be found some deprived individuals, whom a rich-to-poor cross-group redistributive process would clearly not take out of poverty. The within- and between-group distribution of well-being that is required for devising an optimal redistributive scheme can be again revealed by a comprehensive poverty profile.

2.1 Welfarist and non-welfarist policy implications

The distinction between the welfarist and non-welfarist approaches to poverty measurement often matters (implicitly or explicitly) for the assessment and the design of public policy. As described above, a welfarist approach holds that individuals are the best judges of their own well-being. It would thus in principle avoid making appraisals of well-being that conflict with the poor's views of their own situation. A typical example of a welfarist public policy would be the provision of adequate income-generating opportunities, leaving individuals decide and reveal whether these opportunities are utility maximising, keeping in mind the other non-income-generating opportunities that are open to them.

A non-welfarist policy analyst would argue, however, that raising income opportunities is not necessarily the best policy option. This is partly because individuals are not necessarily best left with their own resolutions, at least in an intertemporal setting, for their educational and environmental choices for instance. In other words, the poor's short-run preoccupations may harm their long-term self-interest. For example, individuals may choose not to attend skill-enhancing programmes because they appear overly time costly in the short-run, and because they are not sufficiently convinced or aware of their long-term benefits.

Hence, if left to themselves, the poor will not necessarily spend their income increase on functionings that basic-needs analysts would normally consider a priority, such as good nutrition and health. Thus, fulfilling "basic needs" cannot be satisfied only by the generation of private income, but may require significant amounts of targeted and in-kind public expenditures on areas such as education, public health and the environment. This would be so even if the poor did not presently believe that these areas were deserving of public expenditures. Furthermore, social cohesion concerns are arguably not well addressed by the maximization of private utility, and raising income opportunities will not fundamentally solve problems caused by adverse intra-household distributions of well-being, for

instance.

An objection to the basic needs approach is that it is clearly paternalistic since it supposes that it is in the absolute interests of all to meet a set of often arbitrarily specified needs. Indeed, as emphasised above, non-welfarist approaches in general may use criteria for identifying and helping the poor that may conflict with the poor's views and utility maximizing options. For poverty alleviation purposes, this could go as far as enforced enrolment in community development programmes. This would not only conflict with the preferences of the poor, but would also clearly undermine their freedom to choose. Freedom to choose may, however, be one of the basic capabilities which contribute fundamentally to well-being.

A further example of the possible tension between welfarist and non-welfarist approaches to public policy comes from optimal taxation theory, which is linked to optimal poverty alleviation theory. In the tradition of classical microeconomics, which values leisure in the production and labour market decisions of individuals, pure welfarists would incorporate the utility of leisure in the overall utility function of workers, poor and non-poor alike. In its support to the poor, the government would then take care of minimizing the distortion of their labor/leisure choices so as not to create overly high "deadweight losses". Classical optimal taxation theory then shows that giving a positive weight to such things as labor/leisure distortions suggests a generally lower benefit reduction rates on the income of the poor than otherwise. Taking into account such abstract things is less typical of the basic needs and functioning approaches. Such approaches would, therefore, usually be less reluctant to target programme benefits more sharply on the poor, and exact steeper benefit reduction rates as income or well-being increases.

Relative to the pure welfarist approach, non-welfarist approaches are also typically less reluctant to impose utility-decreasing (or "workfare") costs as side effects of participation in poverty alleviation schemes. These side effects are in fact often observed in practice. For instance, it is well-known that income support programmes frequently impose participation costs on benefit claimants. These are typically non-monetary costs. Such costs can be both physical and psychological: providing manual labor, spending energy, spending time away from home, sacrificing leisure and home production, finding information about application and eligibility conditions, corresponding and dealing with the benefit agency, queuing, keeping appointments, complying with application conditions, revealing personal information, feeling "stigma" or a sense of guilt, *etc...*

Although non-monetary, these costs have a clear impact on participants' net utility from participating in the programmes. When they are negatively correlated

with unobserved (or difficult to observe) entitlement indicators, they can provide self-selection mechanisms that enhance the efficiency of poverty alleviation programmes, for welfarists and non-welfarists alike. One unfortunate effect of these costs is, however, that many truly-entitled and truly deserving individuals may shy away from the programmes because of the costs they impose. Although programme participation could raise their income and consumption above a money-metric poverty line, some individuals will prefer not to participate, revealing that they find apparent poverty utility dominant over programme participation. Welfarists would in principle take these costs into account when assessing the merits of the programmes. Non-welfarists would typically not do so, and would therefore judge the programmes more favorably.

The width of the definition of functionings is clearly also important for the assessment and the design of public policy. For instance, public spending on education is often promoted on the basis of its impact on productivity and growth. But education can also be seen as a means to attain the functioning of literacy and participation in the community. This provides an additional strong support for public expenditures on education. Analogous arguments also apply, for instance, to public expenditures on health, transportation, and the environment.

Part II

Measuring poverty and equity

3 Notation

In what follows in this book, we will denote living standards by the variable y . The indices we will use will sometimes require these living standards to be strictly positive, and, for expositional simplicity, we will assume that this is always the case. Strictly positive values of y are required, for instance, for the Watts poverty index and for many of the decomposable inequality indices. It is of course reasonable to expect indicators of living standards such as monthly or yearly consumption to be strictly positive. This assumption is less natural for other indicators, such as income, for which capital losses or retrospective tax payments can generate negative values.

Let $p = F(y)$ be the proportion of individuals in the population who enjoy a level of income that is less than or equal to y . $F(y)$ is called the cumulative distribution function (*cdf*) of the distribution of income; it is non-decreasing in y , and varies between 0 and 1, with $F(0) = 0$ and $F(\infty) = 1$. For expositional simplicity, we may assume that $F(y)$ is continuously differentiable and strictly increasing in y (a reasonable assumption for large-population distributions of income). The density function, which is the first-order derivative of the *cdf*, is denoted as $f(y) = F'(y)$ and is strictly positive since $F(y)$ is assumed to be strictly increasing in y .

3.1 Continuous distributions

A useful tool throughout the analysis will be the concept of “quantiles”. Quantiles will help simplify greatly the exposition and the computation of several distributive measures. They will also sometimes serve as direct tools to analyze and compare distributions of living standards (to check first-order dominance in the dual approach for instance). The quantile $Q(p)$ is defined as $F(Q(p)) = p$, or using the inverse distribution function, as $Q(p) = F^{-1}(p)$. $Q(p)$ is thus the living standard level below which we find a proportion p of the population. Alternatively, it is the living standard of that individual whose rank – or percentile – in the distribution is p . A proportion p of the population is poorer than he is; a proportion $1 - p$ is richer than him.

This is illustrated in Figure 14. The horizontal axis shows percentiles p of the population. The quantiles $Q(p)$ that correspond to different p values are shown on the vertical axis. The larger the rank p , the higher the corresponding living standard $Q(p)$. Alternatively, living standards y appear on the vertical axis of Figure 14, and the proportion of individuals whose income is below or equal to

those y are shown on the horizontal axis. At the maximum income level, y_{max} , that proportion $F(y_{max})$ equals 1. The median is given by $Q(0.5)$, which is the living standard which splits the distribution exactly in two halves.

We will define most of the distributive measures (indices and curves) in terms of integrals over a range of percentiles. This is a familiar procedure in the context of continuous distributions. We will see below why this is also generally valid in the context of discrete distributions, even though the use of summation signs is more familiar in that context. Using integrals will make the definitions and the exposition simpler, and will help focus on what matters more, namely, the interpretation and the use of the various indices and curves that we will consider.

The most common summary index of a distribution is its mean. Using integrals and quantiles, it is defined as:

$$\mu = \int_0^1 Q(p)dp. \quad (1)$$

μ is therefore simply the area underneath the quantile curve. This corresponds to the grey area shown on Figure 14. Since the horizontal axis varies uniformly from 0 to 1, μ is also the average height of the quantile curve $Q(p)$, and this is given by μ on the vertical axis. As for most distributions of income, the one shown on Figure 14 is skewed to the left, which gives rise to a mean μ that exceeds the median $Q(p)$. Said differently, the proportion of individuals underneath the mean, $F(\mu)$, exceeds one half.

For poverty comparisons, we will also need the concept of quantiles censored at a poverty line z . These are denoted by $Q^*(p; z)$ and defined as:

$$Q^*(p; z) = \min(Q(p), z). \quad (2)$$

Censored quantiles are therefore just the incomes $Q(p)$ for those in poverty (below z) and z for those whose income exceeds the poverty line. This is illustrated on Figure 15, which is similar to Figure 14. Quantiles $Q(p)$ and censored quantiles $Q^*(p; z)$ are identical up to $p = F(z)$, or up to $Q(p) = z$. After this point, censored quantiles equal z and diverge from income $Q(p)$.

Censoring income at z helps focus attention on poverty, since the precise value of those living standards that exceed z is irrelevant for poverty analysis and poverty comparisons (at least so long as we consider *absolute* poverty; more on this later). The mean of the censored quantiles is denoted as $\mu^*(z)$:

$$\mu^*(z) = \int_0^1 Q^*(p; z) dp. \quad (3)$$

This is again the area underneath the curve of censored incomes $Q^*(p; z)$. The poverty gap at percentile p , $g(p; z)$, is the difference between the poverty line and the censored quantile at p , or equivalently the shortfall (when applicable) of living standard $Q(p)$ from the poverty line:

$$g(p; z) = z - Q^*(p; z) = \max(z - Q(p), 0). \quad (4)$$

When income at p exceeds the poverty line, the poverty gap equals zero. A shortfall $g(q; z)$ at rank q is shown on Figure 15 by the distance between z and $Q(q)$. The larger one's rank p in the distribution – the higher up in the distribution of income – the lower the poverty gap $g(p; z)$. The proportion of individuals with a positive poverty gap is given by $F(z)$ (see the Figure). The average poverty gap then equals $\mu^g(z)$:

$$\mu^g(z) = \int_0^1 g(p; z) dp. \quad (5)$$

$\mu^g(z)$ is the size of the area in grey shown in Figure 15.

3.2 Discrete distributions

To see how to rewrite the above definitions using summation signs and discrete distributions, we need a little more notation. Say that we are interested in a distribution of n living standards. We first order the n observations of y_i in increasing values of y , such that $y_1 \leq y_2 \leq y_3 \leq \dots \leq y_{n-1} \leq y_n$. We then define n discrete quantiles of living standards as $Q(p_i) = y_i$, for $p_i = 1/n, 2/n, 3/n, \dots, (n-1)/n, 1$.

This is illustrated in Table ?? where $n = 3$ and where the incomes in increasing values are 10, 20 and 30. Figure 13 graphs those quantiles as a function of p .

The formulae for discrete distributions are then computed in practice by replacing the integral sign in the continuous case by a summation sign, by summing across all observed sample quantiles, and by dividing the sum by the number of observations n . Thus, the mean μ of a discrete distribution can be expressed as:

$$\mu = \frac{1}{n} \sum_{i=1}^n Q(p_i). \quad (6)$$

As indicated by equation (1), the mean of the discrete distribution of Table ??, which is 20, is simply the integral of the quantile curve shown on Figure 13. In other words, it is the sum of the area of the three boxes each of length 1/3 that can be found underneath the filled curve.

Discrete distributions are in fact what is always observed in samples and in real-life populations of households or individuals, however large those samples or populations may be. For clarity, we will mention from time to time how indices and curves can be estimated using the more familiar summation signs. For more information, you can also consult *DAD's User Guide* where the estimation formulae shown use summation signs and thus apply to discrete distributions.

4 The measurement of inequality and social welfare

4.1 Lorenz curves

The Lorenz curve has been for the several decades the most popular graphical tool for visualizing and comparing the inequality in income. As we will see, it provides complete information on the whole distribution of income as a proportion of the mean. It therefore gives a more comprehensive description of the relative standards of living than any one of the traditional summary statistics of dispersion can give, and it is also a better starting point when looking at the inequality of income than the computation of the many inequality indices that have been proposed. As we will see, its popularity also comes from its use as a device to order distributions in terms of inequality, in such a way as to check whether the ordering is necessarily the same for (and is therefore robust over) all inequality indices within a large class of inequality indices. The Lorenz curve is defined as follows:

$$L(p) = \frac{1}{\mu} \int_0^p Q(q) dq. \quad (7)$$

The numerator shows the absolute contribution to *per capita* income of the bottom p proportion (the $100p\%$ poorest) of the population. μ is average income. $L(p)$ thus indicates the cumulative percentage of total income held by a cumulative proportion p of the population, when individuals are ordered in increasing values of their income. For instance, if $L(0.5) = 0.3$, then we know that the 50% poorest individuals hold 30% of the total income in the population.

A discrete formulation of the Lorenz curve is easily provided. Recall that discrete income y_i are ordered such that $y_1 \leq y_2 \leq \dots \leq y_n$, with percentiles $p_i = i/n$, for $i = 1, \dots, n$. For $i = 1, \dots, n$, the discrete Lorenz curve is then defined as:

$$L(p_i = i/n) = \frac{1}{n\mu} \sum_{j=1}^i Q(p_j). \quad (8)$$

If needed, other values of $L(p)$ in (8) can be obtained by interpolation.

The Lorenz curve has several interesting properties. It ranges from 0 at $p = 0$ to 1 at $p = 1$, since a proportion $p = 1$ of the population must also hold a

proportion $L(p = 1) = 1$ of the aggregate income. It is increasing as p increases, since more and more incomes are then added up. This is also seen by the fact that the derivative of $L(p)$ equals $Q(p)/\mu$,

$$\frac{dL(p)}{dp} = \frac{Q(p)}{\mu}. \quad (9)$$

This is positive if income are strictly positive, as we have assumed. Hence by observing the slope of the Lorenz curve at a particular value of p , we also know the p -quantile relative to the mean, or in other words, the standard of living of an individual at rank p as a proportion of the overall mean standard of living. The slope of $L(p)$ thus portrays the whole distribution of mean-normalised income. The Lorenz curve is also convex in p , since as p increases, the new incomes that are being added up are greater than those that have already been counted. Mathematically, a curve is convex when its second derivative is positive, and the more positive that second derivative, the more convex is the curve. We can show that the second-order derivative of the Lorenz curve equals:

$$\frac{d^2L(p)}{dp^2} = \frac{1}{f(Q(p))} \quad (10)$$

which is positive. The larger the density $f(Q(p))$ of income at a quantile $Q(p)$, the more convex the Lorenz curve at $L(p)$.

Some measures of central tendency can also be identified by a look at the Lorenz curve. In particular, the median (as a proportion of the mean) is given by $Q(0.5)/\mu$, and thus by the slope of the Lorenz curve at $p = 0.5$. Since many distributions of incomes are skewed to the right, the mean exceeds the median and $Q(p = 0.5)/\mu$ will typically be less than one. The mean living standard in the population is found at the percentile at which the slope of $L(p)$ equals 1, that is, where $Q(p) = \mu$. Again, this percentile will often be larger than 0.5, namely, where the median living standard is located. The percentile of the mode (or modes) is where $L(p)$ is least convex, since by equation (10) this is where the density $F(Q(p))$ is highest.

If all had the same income, the Lorenz curve would equal p : population shares and shares of total income would be identical. An important graphical element of a Lorenz curve is thus its distance, $p - L(p)$, from the line of perfect equality in income.

Simple summary measures of inequality can already be obtained from the graph of the Lorenz curve. The share in total income of the bottom p proportion of the population is given by $L(p)$; the greater that share, the more equal is

the distribution of income. Analogously, the share in total income of the richest p proportion of the population is given by $1 - L(p)$; the greater that share, the more unequal is the distribution of income. These two simple indices of inequality are often used in the literature. An interesting but less well-known index of inequality is given by the minimum (hypothetical) proportion of total income that the government would need to reallocate across the population to achieve perfect equality in income; this proportion is given by the maximum value of $p - L(p)$, which is attained where the slope of $L(p)$ is 1 (*i.e.*, at $L(p = F(\mu))$). It is therefore equal to $F(\mu) - L_F(F(\mu))$ and is also called the Schutz coefficient. We will revert to that coefficient later when we discuss relative poverty.

Mean-preserving equalising transfers of income are often call Pigou-Dalton transfers; in money-metric terms, they involve a marginal transfer of \$1, say, from a richer to a poorer person, and they keep the mean of income constant. All indices of inequality which do not increase (and sometimes fall) following any such equalising transfers are said to obey the Pigou-Dalton principle of transfers. These equalising transfers also have the consequence of moving the Lorenz curve unambiguously closer to the line of equality. Let the Lorenz curve $L_B(p)$ of a distribution B be everywhere above the Lorenz curve $L_A(p)$. We can thus think of B as having been obtained from A through a series of equalising transfers, applied to the distribution A . Hence, inequality indices which obey the principle of transfers will unambiguously indicate more inequality in A than in B . We will come back to this important link in the Section 10 on making robust comparisons of inequality.

4.2 Gini indices

Compared to perfect equality, inequality thus removes a proportion $p - L(p)$ of total income from the bottom $100 \cdot p\%$ of the population. If we aggregate that “deficit” $p - L(p)$ between population shares, p , and shares in income, $L(p)$, across all values of p between 0 and 1, we get half the well-known Gini index:

$$\frac{\text{Gini index of inequality}}{2} = \int_0^1 (p - L(p)) dp. \quad (11)$$

The Gini index thus assumes that all “share deficits” across p are equally important. It thus computes the average distance between cumulated population shares and cumulated shares in income. One can, however, also think of other weights to aggregate the distance $p - L(p)$. The class of *linear* inequality measures is given

by the use of rank- or percentile-dependent weights, say $\kappa(p)$, applied to that distance. A popular one-parameter functional specification for such weights is given by

$$\kappa(p; \rho) = \rho(\rho - 1)(1 - p)^{(\rho-2)} \quad (12)$$

which depends on the value of a single “ethical” parameter ρ which must be greater than 1 for the weights $\kappa(p; \rho)$ to be positive everywhere. The shape of $\kappa(p; \rho)$ is shown on Figure 3 for three different values of ρ . The larger the value of ρ , the larger the value of $\kappa(p; \rho)$ for small p .

Using (12) gives what is called the class of S-Gini (or “Single-Parameter” Gini) inequality indices:

$$I(\rho) = \int_0^1 (p - L(p))\kappa(p; \rho)dp. \quad (13)$$

Note that when $\rho = 2$, we have that $I(2)$ is the standard Gini index. This is because $\kappa(p; \rho = 2) \equiv 2$, which then gives equal weight to all distances $p - L(p)$. When $1 < \rho < 2$, relatively more weight is given to the distances occurring at larger values of p , as shown by Figure 3. Conversely, when $\rho > 2$, relatively more weight is given to the distances occurring at lower values of p . Changing ρ thus changes the “ethical” concern which we feel for the “shares deficits” at various cumulative proportions of the population.

Let $\omega(p; \rho)$ be defined as follows:

$$\omega(p; \rho) = \int_p^1 k(q, \rho)dq = \rho(1 - p)^{\rho-1}. \quad (14)$$

Note that $\omega(p; \rho) > 0$ and that $\partial\omega(p; \rho)/\partial p < 0$ when $\rho > 1$. The shape of $\omega(p; \rho)$ is shown on Figure 2 for ρ equal to 1.5, 2 and 3. Since $\int_0^1 \omega(p; \rho)dp = 1$ for any value of ρ , the area under each of the three curves on Figure 2 equals 1 too.

The functions $\kappa(p; \rho)$ and $\omega(p; \rho)$ can be given an interpretation in terms of densities of the poor, densities which will be useful to interpret some of the relationships to be described below. Assume that r individuals are randomly selected from the population. The probability that the income of *all* of these r individuals will exceed $Q(p)$ is given by $[1 - F(Q(p))]^r$, and thus the probability of finding a living standard below $Q(p)$ in such samples is $1 - [1 - F(Q(p))]^r = 1 - [1 - p]^r$. The density of the lowest rank of income in a sample of r randomly selected income is the derivative of that probability with respect to p , which is

$$r(1 - p)^{r-1}. \quad (15)$$

This helps us interpret the weights $\kappa(p; \rho)$ and $\omega(p; \rho)$. By equation (12), $\kappa(p; \rho)$ is ρ times the density of the lowest living standard in a sample of $\rho - 1$ randomly selected individuals; analogously, by equation (14), $\omega(p; \rho)$ is the density of the lowest living standard in a sample of ρ randomly selected individuals.

Using (14) and by integration by parts of equation (13), we can then show that:

$$I(\rho) = \frac{1}{\mu} \int_0^1 (\mu - Q(p)) \omega(p; \rho) dp. \quad (16)$$

This says that the deviation of income from the mean is weighted by weights which fall with the ranks of individuals in the population. Since, in equation (16), $I(\rho)$ is a (piece-wise) linear function of the income $Q(p)$, it is a member of the class of linear inequality measures, a feature which will prove useful in measuring progressivity and vertical equity later.

We might be interested in determining the impact of some inequality-changing process on the inequality indices of type (16). One such process that can be handled nicely spreads income away from the mean by a proportional factor λ , and thus corresponds to some form of bi-polarization of incomes away from the mean (loosely speaking). This is equivalent to a process that adds $(\lambda - 1)(Q(p) - \mu)$ to $Q(p)$, since

$$\mu - (Q(p) + (\lambda - 1)(Q(p) - \mu)) = \lambda(\mu - Q(p)). \quad (17)$$

As can be checked from equation (16), this changes $I(\rho)$ proportionally by λ , which also says that the elasticity of $I(\rho)$ with respect to λ , when λ equals 1 initially, is equal to 1 whatever the value of the parameter ρ .

This bi-polarization away from the mean is also equivalent to a process that increases the distance $p - L(p)$ by a factor λ . That this gives the same change in $I(\rho)$ can be checked from equation (13). This bi-polarization process thus increases the "deficit" $p - L(p)$ between population shares p and income shares $L(p)$ by a constant factor λ across population shares or ranks. We will see later how this distance-increasing process leads to a nice illustration of the possible impact of changes in inequality on poverty.

As shown on Figure 2, the larger the value of ρ , the greater the weight given to the deviation of low incomes from the mean. When ρ becomes very large, the index $I(\rho)$ equals the proportional deviation from the mean of the lowest living standard. When $\rho = 1$, the same weight $\omega(p; \rho = 1) \equiv 1$ is given to all deviations from the mean, which then makes the inequality index $I(\rho = 1)$ always

equal to 0, regardless of the distribution of income under consideration. Hence, ρ is a parameter of inequality aversion that determines our ethical concern for the deviation of quantiles from the mean at various ranks in the population. In this sense, it is analogous to the parameter ϵ of relative inequality aversion which we will discuss below in the context of the Atkinson indices. For the standard Gini index of inequality, we have that $\rho = 2$ and thus that $\omega(p; \rho = 2) = 2 \cdot (1 - p)$; hence in assessing the standard Gini, the weight on the deviation of one's living standard from the mean decreases linearly with one's rank in the distribution of income. In a discrete formulation, the weights $\omega(p; \rho)$ take the form of:

$$\omega(p_i; \rho) = \frac{(n - i + 1)^\rho - (n - i)^\rho}{n^\rho}. \quad (18)$$

The S-Gini indices of inequality have nice properties. First, they are graphically easily interpreted as a weighted area underneath the Lorenz curve. Second, they range between 0 (when all incomes are equal to the mean or when the ethical parameter ρ is set to 1) and 1 (when incomes are concentrated in the hands of only one individual, or when ρ is large and the lowest living standard is close to 0). Since the Lorenz curve moves towards p when a Pigou-Dalton equalising transfer is exerted, the value of the S-Gini indices also decreases with such transfers. Finally, the S-Gini indices can be shown to be equal to the following covariance formula:

$$I(\rho) = \frac{-cov(Q(p), \rho(1 - p)^{(\rho-1)})}{\mu} \quad (19)$$

which makes their computation simple using common spreadsheet or statistical softwares. The traditional Gini is then simply:

$$I(\rho = 2) = \frac{2 cov(Q(p), p)}{\mu} \quad (20)$$

which is just a proportion of the covariance between incomes and their ranks.

A further useful interpretive property of the standard Gini index is that it equals half the mean-normalised average distance between all incomes:

$$I(\rho = 2) = \frac{\int_0^1 \int_0^1 |Q(p) - Q(q)| dpdq}{2\mu}. \quad (21)$$

Thus, if we find that the Gini index of a distribution of income equals 0.4, then we know that the average distance between the incomes of that distribution is of the order of 80% of the mean.

A final interesting interpretation of the Gini index is in terms of average relative deprivation, which has been linked in the sociological and psychological literature to subjective well-being, social protest and political unrest. For this, it is usual to quote from the classic work of Runciman (1966), who defines relative deprivation as follows:

The magnitude of a relative deprivation is the extent of the difference between the desired situation and that of the person desiring it (as he sees it). (p.10)

Sen (1973), Yitzhaki(1979) and Hey and Lambert(1980) follow Runciman's lead to propose for each individual an indicator of relative deprivation which measures the distance between his income and the income of all those relative to whom he feels deprived. Thus, let the relative deprivation of an individual with income $Q(p)$, when comparing himself to another individual with income $Q(q)$, be given by:

$$\delta(p, q) = \begin{cases} 0, & \text{if } Q(p) \geq Q(q) \\ Q(q) - Q(p), & \text{if } Q(p) < Q(q) \end{cases} \quad (22)$$

The expected relative deprivation of an individual at rank p is then $\bar{\delta}(p)$:

$$\bar{\delta}(p) = \int_0^1 \delta(p, q) dq \quad (23)$$

which, we can show, can be computed as $\bar{\delta}(p) = \mu(1 - L(p)) - Q(p)(1 - p)$. As we did for the "shares deficits" above, we can aggregate the relative deprivation at every percentile p by applying the weights $\kappa(p; \rho)$. We can show that this gives the S-Gini index of inequality:

$$I(\rho) = \frac{1}{\rho \mu} \int_0^1 \bar{\delta}(p) \kappa(p; \rho) dp. \quad (24)$$

Hence, the S-Gini indices are also an indicator of the average relative deprivation felt in a population. By equations (12), (15) and (24), they equal the expected relative deprivation of the poorest individual in a sample of $\rho - 1$ randomly selected individuals. The greater the value of ρ , the more important is the relative deprivation of the poorer in computing $I(\rho)$.

4.3 Social welfare

We now introduce the concept of a social welfare function. Unlike the concept of relative inequality, which considers incomes relative to the mean, the concept of social welfare will allow us to measure and compare the *absolute* incomes of populations. We will see, however, that under some popular conditions on the shape of social welfare functions, the measurement of inequality and social welfare can be nicely linked and integrated, and that the tools used for the two concepts are then similar.

The social welfare functions we will consider will take the form of:

$$W = \int_0^1 U(Q(p))\omega(p)dp \quad (25)$$

where for expositional simplicity we will restrict $\omega(p)$ to be of the special form $\omega(p; \rho)$ defined by equation (14). $U(Q(p))$ is a “utility function” of income $Q(p)$. Social welfare is then the expected utility of the poorest individual in a sample of $(\rho - 1)$ individuals.

The first requirement that we wish to impose on the form of W is that it be *homothetic*. Homotheticity of W is analogous to the requirement on consumer utility functions that expenditure shares of the different consumption goods be constant as income increases, or the requirement on production functions that the ratio of the marginal products of inputs stays constant when all inputs are doubled. For social welfare measurement, homotheticity implies that the ratio of the marginal social utilities of any two individuals in a population stays the same even when all incomes are doubled or halved¹. For (25) to be homothetic, we need $U(Q(p))$ to take the popular form of $U(Q(p); \epsilon)$, where

$$U(Q(p); \epsilon) = \begin{cases} \frac{Q(p)^{1-\epsilon}}{(1-\epsilon)}, & \text{when } \epsilon \neq 1 \\ \ln Q(p), & \text{when } \epsilon = 1 \end{cases} \quad (26)$$

Hence, W in equation (25) will depend on the parameters ρ and on ϵ , and we will denote this as $W(\rho, \epsilon)$.

Homotheticity of a social welfare function has an important advantage: the social welfare function can then be used to measure relative inequality, the most common concept of inequality in the literature on the distribution of income. To see how this can be done, define $\xi(\rho, \epsilon)$ as the equally distributed living standard that is equivalent, in terms of social welfare, to the actual distribution of income

¹The marginal social utility of a living standard $Q(p)$ is given by $\partial W / \partial Q(p) = W^{(1)}$.

(we will refer to ξ as the EDE living standard). $\xi(\rho, \epsilon)$ is then implicitly defined as:

$$\int_0^1 U(\xi(\rho, \epsilon); \epsilon) \omega(p; \rho) dp = \int_0^1 U(Q(p); \epsilon) \omega(p; \rho) dp. \quad (27)$$

Since $\int_0^1 \omega(p; \rho) dp = 1$, $\xi(\rho, \epsilon)$ is such that:

$$U(\xi(\rho, \epsilon); \epsilon) = \int_0^1 U(Q(p); \epsilon) \omega(p; \rho) dp \quad (28)$$

or, alternatively:

$$\xi(\rho, \epsilon) = U_\epsilon^{-1} \left(\int_0^1 U_\epsilon(Q(p)) \omega(p; \rho) dp \right) = U_\epsilon^{-1}(W(\rho, \epsilon)) \quad (29)$$

where $U_\epsilon^{-1}(\cdot)$ is the inverse utility function:

$$U_\epsilon^{-1}(x) = \begin{cases} (1 - \epsilon)s^{\frac{1}{1-\epsilon}}, & \text{when } \epsilon \neq 1, \\ \exp(x), & \text{when } \epsilon = 1, \end{cases} \quad (30)$$

The index of inequality I corresponding to the social welfare function W is then defined as the distance between the EDE living standard and mean income, as a proportion of mean income:

$$I = \frac{\mu - \xi}{\mu} = 1 - \frac{\xi}{\mu}. \quad (31)$$

When using the specific forms $W(\rho, \epsilon)$ and $\xi(\rho, \epsilon)$, this gives $I(\rho, \epsilon)$.

Clearly, then, the EDE living standard is a simple function of average living standard and inequality in its distribution, with $\xi = \mu \cdot (1 - I)$. Compare to the W , ξ also has the advantage of being money metric and thus of being easily understood and compared to other economic indicators that can also be expressed in money-metric terms. To increase social welfare, we can either try to increase μ , or increase equality of income $1 - I$ by decreasing inequality I . Two distributions of income can display the same social welfare even with different average income if these differences are offset by differences in inequality. This is shown in Figure 23, starting initially with two different levels of mean income μ_0 and μ_1 and zero inequality. We then have that $\xi = \mu_0$ and $\xi = \mu_1$. To preserve the same level of social welfare in the presence of inequality, mean income must be higher: this is

shown by the positive slope of the constant ξ functions. Furthermore, as inequality becomes large, further increases in I must be matched by higher and higher increases in mean income for social welfare not to fall.

Defined as in (31), inequality has an interesting interpretation: it measures the difference between the mean level of actual income and the (lower) level needed instead to achieve the same level of social welfare when income is distributed equally across the population. This difference being expressed as a proportion of mean income, I thus shows the *per capita* proportion of income that is wasted in social terms because of its unequal distribution. Society as a whole would be just as well-off with an equal distribution of a proportion of just $1 - I$ of the total actual income. I can thus be interpreted as a money-metric indicator of the social cost of inequality.

Let a distribution A of income just be a proportional re-scaling of a distribution B . In other words, for a constant $\lambda > 0$, we have that $Q_A(p) = \lambda Q_B(p)$ for all p . If the social welfare function used for the computation of I is homothetic, it must be that $I_A = I_B$. This is illustrated in Figure 20 for the case of two incomes y_1^A and y_2^A for the case of an initial distribution A and y_1^B and y_2^B for a "scaled-up" distribution B . Social welfare in A is given by W_A . The social indifference curve W_A shown in Figure 20 also depicts the many other combinations of incomes that would yield the same level of social welfare. One of these combinations, at point F , corresponds to a situation of equality of income where both individuals enjoy ξ_A . ξ_A is therefore the equally distributed living standard that is socially equivalent to the distribution (y_1^A, y_2^A) .

The average living standard in A is given by μ_A , which is point G in Figure 20. Hence two distributions of income, one made of the vector (y_1^A, y_2^A) and the other of the vector (ξ_A, ξ_A) , generate the same level of social welfare, the first with an unequally distributed average living standard μ_A and the other with an equally distributed average living standard ξ_A . Hence, the distance between point F and point G in Figure 20 can be understood as the "cost of inequality" in the distribution A of income. Taking that distance as a proportion of μ_A (see equation (31)) gives the index of inequality in the distribution A .

The fact that $y_1^A = \lambda y_1^B$ and $y_2^A = \lambda y_2^B$ for the same λ can be seen from the fact that the two vectors of income lie along the same ray from the origin. If the function W is homothetic, then inequality in A must be the same as inequality in B . In other words, the distance between points D and E as a proportion of the distance OE must be the same as the distance between points F and G as a proportion of the distance OG .

4.3.1 Atkinson indices

Two special cases of $W(\rho, \epsilon)$ are of particular interest in assessing social welfare and relative inequality. The first is when the rank of income is not important in computing social welfare: this is obtained when $\rho = 1$, and it yields the well-known Atkinson additive social welfare function, $W(\epsilon)$:

$$W(\epsilon) = W(\rho = 1, \epsilon) = \int_0^1 U(Q(p); \epsilon) dp. \quad (32)$$

The Atkinson social welfare function has often been interpreted as a utilitarian social welfare function, where $U(Q(p); \epsilon)$ is an individual utility function displaying decreasing marginal utilities of income, or where $U(Q(p); \epsilon)$ corresponds to a concave social evaluation of a concave individual utility of income. It can be argued, however, that “it is fairly restrictive to think of social welfare as a sum of individual welfare components”, and that one might feel that “the social value of the welfare of individuals should depend crucially on the levels of welfare (or incomes) of others” (Sen (1973, p.30 and 41)). The unrestricted form $W(\rho, \epsilon)$ allows for such interdependence and is therefore more flexible than the Atkinson additive formulation. In the light of the above, we can also interpret $W(\rho, \epsilon)$ as the expected utility of the poorest individual in a group of ρ randomly selected individuals. This interpretation of the social evaluation function $W(\rho, \epsilon)$ confirms why it is not additive or separable in individual welfare: the social welfare weight on individual utility $U(Q(p); \epsilon)$ depends on the rank p of the individual in the whole distribution of income.

Figure 21 shows the shape of the utility functions $U(y; \epsilon)$ for different values of ϵ^2 . Incomes are shown on the horizontal axis as a proportion of their mean, and utility $U(y; \epsilon)$ can be read on the vertical axis. A normalization $U(\mu; \epsilon) = 1$ has been applied for graphical convenience. Although for all values of ϵ , the slope of $U(y; \epsilon)$ is positive, it is obviously not always the same across all values of y . This is made more explicit on Figure 22 which shows the marginal social utility of income $U^{(1)}(y; \epsilon)$ for different values of ϵ . Again, a normalization of $U^{(1)}(\mu; \epsilon) = 1$ was made. For $\epsilon = 0$, the marginal social utility is constant: increasing by a given amount a poor person’s living standard has the same social welfare impact as increasing by the same amount a richer person’s living standard. For $\epsilon > 0$, however, increasing the poor’s income is socially more desirable than

²This is drawn from Cowell ???.

increasing the rich's. The larger the value of ϵ , the faster the marginal social utility falls with y .

By (29) and (31), the Atkinson inequality index is then given by:

$$I(\epsilon) = I(\rho = 1, \epsilon) = \begin{cases} 1 - \frac{\left(\int_0^1 Q(p)^{(1-\epsilon)} dp\right)^{\frac{1}{1-\epsilon}}}{\mu}, & \text{when } \epsilon \neq 1, \\ 1 - \frac{\exp\left(\int_0^1 \ln(Q(p)) dp\right)}{\mu}, & \text{when } \epsilon = 1. \end{cases} \quad (33)$$

The Atkinson indices are said to exhibit constant relative inequality aversion, since the elasticity of $U^{(1)}(Q(p); \epsilon)$ with respect to $Q(p)$, is constant and equal to ϵ :

$$\frac{Q(p) U^{(2)}(Q(p); \epsilon)}{U^{(1)}(Q(p); \epsilon)} = \epsilon. \quad (34)$$

Figure 18 illustrates graphically the link between the Atkinson social evaluation functions $W(\epsilon)$ and their associated inequality indices. For this, suppose a population of only two individuals, with incomes y_1 and y_2 as shown on the horizontal axis. Mean income is given by $\mu = (y_1 + y_2) / 2$ (the middle point between y_1 and y_2). The "utility function" $U(y; \epsilon)$ has a positive but decreasing slope. $W(\epsilon)$ is then given by $(U(y_1) + U(y_2)) / 2$, the middle point between $U(y_1)$ and $U(y_2)$.

If equally distributed, an average mean living standard of ξ would be sufficient to generate that same level of social welfare, since on Figure 18 we have that $W(\epsilon) = U(\xi, \epsilon)$. The cost of inequality is thus given by the distance between μ and ξ , shown as C on Figure 18. Inequality is the ratio C/μ .

Graphically, the more "concave" the function $U(y; \epsilon)$, the greater the cost of inequality and the greater the inequality indices $I(\epsilon)$. This can be seen on Figure 19 where two functions $U(y; \epsilon)$ have been drawn, with different relative inequality aversion parameters $\epsilon^0 < \epsilon^1$. This difference leads to $\xi^0 > \xi^1$, and therefore to $I(\epsilon^0) < I(\epsilon^1)$. A specification with greater inequality aversion leads to a greater inequality index, and to the judgement that a greater proportion of average income is socially wasted because of the inequality in its distribution.

4.3.2 S-Gini indices

The second special case is obtained when the utility functions $U(Q(p); \epsilon)$ are linear in the levels of living standard, and thus when $\epsilon = 0$. This yields the class of

S-Gini social welfare functions, on which the S-Gini inequality indices are based:

$$W(\rho) = W(\rho, \epsilon = 0) = \int_0^1 Q(p)\omega(p; \rho) dp. \quad (35)$$

Social welfare is thus the expected living standard of the poorest individual in a group of ρ randomly selected individuals. By (29), this is also the EDE living standard. Hence, the inequality indices are then given by:

$$I(\rho, \epsilon = 0) = 1 - \frac{\int_0^1 Q(p)\omega(p; \rho) dp}{\mu} \quad (36)$$

$$= \frac{\int_0^1 (\mu - Q(p))\omega(p; \rho) dp}{\mu} \quad (37)$$

which is seen by (16) to be the same as the S-Gini indices $I(\rho)$. Hence, social welfare and the EDE living standard equal the *per capita* living standard corrected by the extent of relative deprivation in those incomes:

$$W(\rho) = \mu - \frac{1}{\rho} \int_0^1 \bar{\delta}(p)\kappa(p; \rho) dp. \quad (38)$$

A useful curve for the analysis of the distribution of absolute incomes is the Generalised Lorenz curve. It is defined as $GL(p)$:

$$GL(p) = \mu \cdot L(p). \quad (39)$$

The Generalised Lorenz curve has all of the attributes of the Lorenz curve, except for the fact that it does not normalise income by the mean. By (13), (31) and (35), we note that the Generalised Lorenz curve has a nice graphical link with the S-Gini index of social welfare:

$$W(\rho) = \int_0^1 GL(p)\kappa(p; \rho) dp. \quad (40)$$

4.4 Decomposable indices of inequality

A frequent goal is to explain the total amount of inequality in a distribution by the extent of inequality found among socio-economic groups (“intra” or “within” group inequality) and across them (“inter” or “between” group inequality). A useful class of relative inequality indices that allow one to do this is called the class

of decomposable inequality indices. Although that class can be given a justification in terms of social welfare functions, this exercise is less transparent and intuitive than for the class of inequality indices considered above. For all practical purposes, we can express these decomposable inequality indices as Generalised indices of entropy, defined as $I(\theta)$:

$$I(\theta) = \begin{cases} \frac{1}{\theta(\theta-1)} \left(\int_0^1 \left(\frac{Q(p)}{\mu} \right)^\theta dp - 1 \right), & \text{if } \theta \neq 0, 1 \\ \int_0^1 \log \left(\frac{\mu}{Q(p)} \right) dp, & \text{if } \theta = 0 \\ \int_0^1 \frac{Q(p)}{\mu} \log \left(\frac{Q(p)}{\mu} \right) dp, & \text{if } \theta = 1 \end{cases} \quad (41)$$

Some special cases of (41) are worth noting. First, if we replace θ by $1-\epsilon$ (with $\theta \leq 1$), $I(\theta)$ is ordinally equivalent to the family of Atkinson indices. This means that if the use of an Atkinson index $I(\epsilon)$ indicates that there is more inequality in a distribution A than in a distribution B ($I_A(\epsilon) > I_B(\epsilon)$), then the index $I(\theta)$ with $\theta = 1 - \epsilon$ will also indicate more inequality in A than in B ($I_A(\theta) > I_B(\theta)$). Second, $I(\theta = 0)$ gives the Mean Logarithmic Deviation, $I(\theta = 1)$ gives the Theil index of inequality, and $I(\theta = 2)$ is half the square of the coefficient of variation.

Assume that we can decompose the population into K mutually exclusive population subgroups, $k = 1, \dots, K$. The indices in (41) can then be decomposed as follows:

$$I(\theta) = \underbrace{\sum_{k=1}^K \phi(k) \left(\frac{\mu(k)}{\mu} \right)^\theta I(k; \theta)}_{\text{within group inequality}} + \underbrace{\bar{I}(\theta)}_{\text{between group inequality}} \quad (42)$$

where $\phi(k)$ is the share of the population found in group k , $\mu(k)$ is the mean living standard of subgroup k , and $I(k; \theta)$ is inequality within subgroup k . The first term in (42) can be interpreted as a weighted sum of the within group inequalities in the distribution of income. $\bar{I}(\theta)$ is total population inequality when each individual in subgroup k is given the mean living standard $\mu(k)$ of his subgroup (namely, when within subgroup inequality has been eliminated): it can thus be interpreted as the contribution of between subgroup inequality to total inequality. Only, however, when $\theta = 0$ is it the case that the within-group inequality contributions do not depend on mean living standard in the groups; the terms $I(k; \theta)$ are then strictly population-weighted. Otherwise, the within-group inequalities are weighted by weights which depend on the mean living standard in the subgroups k .

4.5 Other popular indices of inequality

There are several other inequality indices that are used in the literature. We list them rapidly here.

A popular descriptive one is the quantile ratio. This is simply the ratio of two quantiles, $Q(p_1)/Q(p_2)$. Popular values of p_1 and p_2 include $p_1 = 0.25$ and $p_2 = 0.75$ (the quartile ratio), as well as $p_1 = 0.10$ and $p_2 = 0.90$ (the decile ratio). Median income is also a popular choice for $Q(p_2)$. For inequality analysis, an arguably better choice for normalizing $Q(p_1)$ is mean income – this can be shown to have a link with first-order restricted inequality dominance.

The coefficient of variation is the ratio of the standard deviation to the mean of income. It is therefore given by $\sqrt{\int_0^1 (Q(p)/\mu - 1)^2 dp}$.

Two popular measures of inequality use sums of logarithms of income. The first one, which we can call the logarithmic variance, is defined as

$$\int_0^1 (\ln Q(p) - \ln \mu)^2 dp \quad (43)$$

and the second, the variance of logarithms, as

$$\int_0^1 \left(\ln Q(p) - \int_0^1 \ln Q(p) dp \right)^2 dp. \quad (44)$$

These two last measures do not, however, always obey the Pigou-Dalton principle of transfers – that is, they will sometimes increase following a spread-reducing transfer of income between two individuals.

Finally, the relative mean deviation is the mean of the absolute deviation from mean income, normalized by mean income:

$$\int_0^1 \frac{|Q(p) - \mu|}{\mu} dp. \quad (45)$$

5 Aggregating and comparing poverty

Making poverty comparisons is essential to determine whether poverty has changed across time, or how it compares across countries, regions, or socio-economic groups. Poverty comparisons are also essential for designing public policy, and for assessing its effects on poverty. They may be used, for instance, to judge whether and by how much a public safety net reduces poverty and whether a reform to its structure could be alleviating poverty further.

5.1 Cardinal versus ordinal comparisons

There are two types of poverty comparisons, cardinal and ordinal. Cardinal poverty comparisons simply involve differences in numerical poverty estimates. Numerical poverty estimates attach a single number to the extent of poverty in a population, for instance, 40% or \$200 per capita. These estimates are valuable when a precise number must be attached to the extent of poverty in a distribution of well-being. Cardinal poverty estimates require specific and precise assumptions, such as the nature of the poverty index, the definition of the indicator of well-being, the value of the poverty line, and how that poverty line varies exactly across household types, regions and time. Once this information is provided, cardinal poverty estimates can tell, for instance, that 30% of the individuals in a population used to have their consumption lie underneath their poverty line, but that a recently-introduced public safety net has decreased that proportion to 25%. Cardinal poverty estimates can also be used to carry out a money-metric cost-benefit analysis of the effects of safety nets. Thus, if the above safety net involved yearly expenditures of \$ 500 million, then we would know immediately that a 1% fall in the proportion of the poor would seem to cost the government on average \$ 100 million. That amount could then be compared to the poverty alleviation cost of other forms of government policy.

The main advantage of cardinal poverty estimates is their ease of communication, their ease of manipulation, and their (apparent) lack of ambiguity. Government officials and the media often want the results of poverty comparisons to be produced in straightforward and precise terms, and can feel annoyed when this is not possible. Cardinal poverty estimates are, however, necessarily (and often highly) sensitive to the choice of a number of arbitrary measurement assumptions. It is clear, for example, that choosing a different poverty line will almost always change the estimated numerical value of any index of poverty. The elasticity of the poverty headcount index to the poverty line is, for example, almost

always significantly larger than 1. This implies that a variation of 10% in the poverty line will change by more than 10% the estimated proportion of the poor in the population; this is a substantial impact for those interested in poverty alleviation, especially since poverty lines are rarely convincingly bounded within a narrow confidence interval. Another source of numerical variability comes from the choice of the form of the poverty index. Many procedures have been proposed to aggregate numerically the poverty of individuals. Depending on the chosen procedure, numerical estimates of poverty will appear large or low. As we will see later, for instance, the estimation of a "socially representative poverty gap" will rest particularly on the weight given to the more deprived among the poor. There is little objective guidance in choosing that weight; the greater that weight, however, the greater the estimated socially representative poverty gap, and the greater the numerical estimate of poverty.

We will return to ordinal comparisons of poverty later. For now, we consider the construction of aggregate cardinal poverty indices.

5.2 Aggregating poverty

Two approaches have been used to devise cardinal indices of poverty. The first uses the concept of the equally distributed equivalent (EDE) living standard of a society where incomes have been censored at the poverty line, and compares it to the poverty line. The second combines income and the poverty line into poverty gaps, and aggregates them in social-welfare like functions to assess overall poverty. We look at these two approaches in turn.

5.2.1 The EDE approach

For the EDE approach to building poverty indices, we simply use the distribution of income $Q(p)$. Since, for poverty comparisons, we want to focus on those incomes that fall below the poverty line (the "focus axiom"), the incomes $Q(p)$ are censored at the poverty line z , to give $Q^*(p; z)$. The censored incomes are then aggregated using one of the many social welfare functions that have been proposed in the literature. A poverty index is obtained by taking the difference between the poverty line and the EDE living standard. For instance, for the social welfare functions proposed in section 4.3, this leads to the following class of poverty indices:

$$P(z; \rho, \epsilon) = z - \xi^*(z; \rho, \epsilon) \tag{46}$$

where $\xi^*(\rho, \epsilon)$ is the EDE living standard of the distribution of censored income $Q^*(p; z)$ and where we need $\rho \geq 1$ and $\epsilon \geq 0$ for the Pigou-Dalton transfer principle not to be violated. $P(z; \rho, \epsilon)$ can then be interpreted as the “socially representative” or EDE poverty gap.

Examples of such poverty indices include a transformation of the Clark, Hemming and Ulph’s (CHU) second class of poverty indices, given by $P(z; \epsilon) = P(z; \rho = 1, \epsilon)$:

$$P(z; \epsilon) = \begin{cases} z - \left(\int_0^1 Q^*(p; z)^{(1-\epsilon)} dp \right)^{\frac{1}{1-\epsilon}}, & \text{when } \epsilon \neq 1, \\ z - \exp \left(\int_0^1 \ln(Q^*(p; z)) dp \right), & \text{when } \epsilon = 1. \end{cases} \quad (47)$$

The CHU indices are then obviously closely related to the Atkinson social welfare functions and inequality indices. When $\epsilon = 1$, the CHU poverty index is also the EDE poverty gap corresponding to the Watts poverty index, which is defined as:

$$PW(z) = \int_0^1 \ln \left(\frac{z}{Q^*(p; z)} \right) dp. \quad (48)$$

For $0 \leq \epsilon \leq 1$, the CHU indices also correspond to the EDE poverty gap of the class of poverty indices proposed by Chakravarty:

$$PC(z; \epsilon) = 1 - \int_0^1 \left(\frac{Q^*(p; z)}{z} \right)^{1-\epsilon} dp, \quad 0 \leq \epsilon \leq 1. \quad (49)$$

Moreover, if we impose $\epsilon = 0$ on the class of indices defined in (46), we obtain the class of S-Gini indices of poverty:

$$P(z; \rho) = P(z; \rho, \epsilon = 0) = z - \int_0^1 Q^*(p; z) \omega(p; \rho) dp. \quad (50)$$

$P(z; \rho = 2)$ is known as the Thon-Chakravarty-Shorrocks index of poverty (it can also be more simply referred to as the “Gini” index of poverty).

5.2.2 The poverty gap approach

The second approach reduces the information necessary to aggregate poverty to the distribution of poverty gaps, $g(p; z) = z - Q^*(p; z)$. Once this distribution is known, no other use of the poverty line is allowed or needed in the aggregation of poverty. Because of this, the poverty gap approach to constructing poverty indices

is more restrictive and puts more structure on the shape of the allowable poverty indices than the previous EDE approach. After the distribution of poverty gaps has been computed, we may use aggregating functions analogous to those used in section 4.3 for the analysis of social welfare. Unlike social welfare functions, however, where we normally wish an increase in someone's living standard to increase social welfare, we would normally wish the poverty indices to be *decreasing* in poverty gaps. Further, whereas an equalizing Pigou-Dalton transfer would often increase the value of a social welfare function, we would typically wish a poverty index to decrease when such an equalizing transfer of income takes place among the poor.

A popular class of poverty gap indices that can obey these axioms is known as the Foster-Greer-Thorbecke (FGT) class. It differentiates its members using an ethical parameter $\alpha \geq 0$ and is generally defined as

$$\bar{P}(z; \alpha) = \int_0^1 \left(\frac{g(p; z)}{z} \right)^\alpha dp \quad (51)$$

for the normalized FGT poverty indices and as

$$P(z; \alpha) = \int_0^1 g(p; z)^\alpha dp \quad (52)$$

for the un-normalized version (which can sometimes be more useful than the usual normalized form). Note that poverty gap indices other than the FGT ones can also be easily proposed, simply by using other aggregating functions of poverty gaps that obey some of the desirable axioms (such as that of being increasing and convex in poverty gaps) discussed in the literature.

When $\alpha = 0$, the FGT index gives the simplest and most common example of a poverty index. This is called the poverty headcount, and is simply the proportion of the poor (those with a positive poverty gap) in a population, $F(z)$. The next simplest and most commonly used index, $\mu^g(z)$, is given by the average poverty gap, $P(z; \alpha = 1)$, and is the average shortfall of income from the poverty line:

$$\mu^g(z) = P(z; \alpha = 1) = \int_0^1 g(p; z) dp. \quad (53)$$

To see how to interpret the form of the FGT indices for general values of α , consider Figure 27. It shows the (absolute) contributions to total poverty $\bar{P}(z; \alpha)$ of individuals at different ranks p . These contributions are given by $(g(p; z)/z)^\alpha$. For $\alpha = 0$, the contribution is a constant 1 for the poor and 0 for the rich (those

whose rank exceeds $F(z)$ on the Figure, or equivalently whose income $Q(p)$ exceeds z). The headcount is then the area of the dotted rectangle on Figure 27. For $\alpha = 1$, the contribution equals precisely the normalized poverty gap, $g(p; z)/z$. The normalized average poverty is then the area underneath the $g(p; z)/z$ drawn on Figure 27. The same reasoning is valid for higher values of α . For instance, the absolute contribution to $\bar{P}(z; \alpha = 3)$ of individuals at rank p is given by $(g(p; z)/z)^3$ on Figure 27.

Notwithstanding the above, interpreting the value of FGT indices for $\alpha \neq 0, 1$ can be problematic. We can easily comprehend what a proportion of the population in poverty or an average poverty gap signifies, but what, for instance, can a squared-poverty-gap index actually mean? And how to explain it to a government minister?... A further difficulty with such indices emerges from another look at Figure 27, which suggests that the contribution of the poor (including the poorest) to total poverty *decreases* with α – the contribution curves $(g(p)/z)^\alpha$ move down as α rises. This also implies that the normalized FGT indices necessarily fall as α increases. This is paradoxical since it is often argued that the higher the value of α , the greater the focus on those who suffer most "severely" from poverty.

One partial solution to these interpretive problems is to switch one's focus from the *absolute* to the *relative* contribution to total FGT poverty of individuals with different poverty gaps. Such relative contribution is depicted on Figure 28 for $\alpha = 0, 1$ and 2. It shows the ratio of the absolute contributions $g(p)^\alpha$ to total poverty $P(z; \alpha)$ (these ratios are the same for normalized and un-normalized FGT indices). Since we are graphing here relative contributions to total poverty, the area underneath each of the three curves must in all cases equal 1.

For $\alpha = 0$, each poor contributes relatively the same constant $1/F(z)$ to the poverty headcount. The poor's relative contribution to the average poverty gap naturally increases with their own poverty gap, as shown by the curve $g(p)/P(z; \alpha = 1)$. That relative contribution equals 1 at those individuals whose own poverty gap is precisely equal to the average poverty gap. The rank of such individuals is given by $F(\mu^g(z))$, as also shown on Figure 28. Thus, those located at $p = F(\mu^g(z))$ have a poverty gap that is representative of the average poverty gap in the population. Increasing α from 1 to 2 decreases the relative contribution of not-so-poor, but increases reciprocally the contribution of those with the highest poverty gaps. This is then consistent with the general opinion that, in the aggregation of poverty, higher values of α put more emphasis on those who suffer most severely from poverty – those with lower values of p and higher values of $g(p; z)$.

Figure 28 does not, however, solve the main interpretation problems associated with the FGT indices. As mentioned above, explaining to non-technicians or

policymakers the practical meaning of FGT indices for general values of α can prove hazardous since these indices are averages of *powers* of poverty gaps. They are also not money metric nor unit-free (except for $\alpha = 0$ and 1). An another already-mentioned difficulty is that the usual FGT indices will generally fall with an increase in the value of their poverty-severity parameter. A valuable and intuitive solution to these two problems is to transform the FGT indices into EDE poverty gaps. An EDE poverty gap is that poverty gap which – if it were assigned equally to all individuals – would yield the same aggregate index as that which is currently observed. An EDE poverty gap can then be usefully be interpreted as a socially-representative poverty gap. This transformation provides a money-metric measure of poverty which can be usefully compared across different poverty indices and/or across different values of α . As we will see later, it also allows the analyst to determine the impact of poverty-gap inequality upon the level of poverty. For the un-normalized FGT indices, the EDE poverty gap is given simply by (for $\alpha > 0$)

$$\xi^g(z; \alpha) = [P(z; \alpha)]^{1/\alpha}. \quad (54)$$

For the normalized FGT indices, it is just $\bar{\xi}^g(z; \alpha) = \xi^g(z; \alpha)/z$.

Figure 29 shows such socially-representative poverty gaps $\xi^g(z; \alpha)$ for different values of α . In each case, we obtain a socially-weighted *money-metric* indicator of the distribution of deprivation in the population. This summary aggregate indicator can also be compared to the individual distribution of poverty, given by the $g(p; z)$ curve. Those whose $g(p; z)$ exceeds $\xi^g(z; \alpha)$ experience more poverty than the social average. Those exactly at $\xi^g(z; \alpha)$ are exactly representative of the socially-weighted average poverty gap. Those representative of population poverty are thus found at the rank given by $F(\xi^g(z; \alpha))$, which is also shown on Figure 29.

An important point to note is that an increase in α moves the socially-representative poverty gap closer to that experienced by the poorest individuals. This is because that, contrary to the usual FGT indices, $\xi^g(z; \alpha + 1) \geq \xi^g(z; \alpha)$ for any $\alpha > 0$. Hence, we can interpret increases in α as increases in the relative weight given to the poorer in computing the socially-representative poverty gap. The larger the value of α , the more important are the most severe cases of deprivation in drawing up a representative aggregate picture of poverty.

Note finally that, besides being already in an EDE poverty gap form, the S-Gini index of poverty also has the property of being a poverty gap index. Indeed, by (50), we have that:

$$P(z; \rho) = \int_0^1 g(p; z) \omega(p; \rho) dp \quad (55)$$

5.3 Group-decomposable poverty indices

Much of the literature on the construction of poverty indices has focussed on whether indices are decomposable across population subgroups. This has led to the identification of a subgroup of poverty indices known as the “class of decomposable poverty indices”. These indices have the property of being expressible as a weighted sum (more generally, as a separable function) of the same poverty indices assessed across population subgroups. They most commonly include the FGT and the Chakravarty classes of indices as well as the Watts index.

Let the population be divided into K mutually exclusive population subgroups, where $\phi(k)$ is the share of the population found in subgroup k . For the FGT index, we then have that:

$$P(z; \alpha) = \sum_k^K \phi(k) P(k; z; \alpha) \quad (56)$$

where $P(k; z; \alpha)$ is the FGT poverty index of subgroup k . The Watts and Chakravarty indices are expressible as a sum of the poverty indices of each subgroup in exactly the same way as for the FGT indices in (56).

To illustrate the practical contents of this property, consider the following two-group ($K = 2$) example. Let the first group contain 40% of the total population, and let the poverty in group 1 be 0.8 and in group 2 be 0.4. Poverty in the total population is then a simple weighted mean of the group poverty, and is immediately computable as $0.4 \cdot 0.8 + 0.6 \cdot 0.4 = 0.56$. Estimates of poverty in a population can then be constructed in a decentralized manner, community by community or region by region, without a need for all micro data to be regrouped in one register.

Subgroup decomposability also implies that an improvement in one of the subgroups will necessarily improve aggregate poverty if the living standards in the other groups have not changed. It will further mean that the design of social safety nets and benefit targeting within any given group can be done independently of the distribution of income in the other groups. This enables targeting to be done in a decentralized manner: only the characteristics of a relevant population matter for the exercise. If targeting succeeds in decreasing poverty at a local level, then it must succeed also at the aggregate level.

Subgroup decomposability is therefore useful, although it is certainly not imperative for poverty analysis. In particular, it must be admitted that it is not because an index property facilitates poverty profiling and the analysis of the comparative advantages of various forms of targeting that this property is ethically fine. Among other things, imposing the decomposability and additivity property can mean losing some important ethical aspects to the aggregation of poverty. In that context, Ravallion (1994) notes that when measuring poverty "one possible objection to additivity is that it attaches no weight to one aspect of a poverty profile: the inequality between sub-groups in the extent of poverty". This can be an important flaw if considerations of between-group relative deprivation are significant.

5.4 Poverty and inequality

Expressing poverty indices in the form of EDE poverty gaps allows one to separate the impact of the average level of poverty and of the inequality in income upon the index of poverty. Let $\xi^g(z)$ be the EDE poverty gap, and $\Xi^g(z)$ be the cost of inequality on the level of poverty. Then,

$$\xi^g(z) - \mu^g(z) = \Xi^g(z). \quad (57)$$

For instance, for the popular FGT indices, we have that:

$$\Xi^g(z; \alpha) = \xi^g(z; \alpha) - \mu^g(z). \quad (58)$$

When $\alpha = 1$, the inequality of poverty gaps is not taken into account in computing poverty, and thus the cost of inequality on poverty is nil: $\Xi^g(z; \alpha = 1) = 0$. For $\alpha > 1$, $\Xi^g(z; \alpha)$ is positive, but for $0 < \alpha < 1$, we have that $\Xi^g(z; \alpha) < 0$ since (*ceteris paribus*) the greater the level of inequality, the *lower* the assessed level of poverty. $\xi^g(z; \alpha)$ and $\Xi^g(z; \alpha)$ are both increasing in α : the larger the value of α , the larger the cost of inequality in poverty gaps on the level of aggregate poverty. We can thus interpret α as a parameter of inequality aversion in measuring poverty.

A similar decomposition can be done using (46) and the EDE level of censored income. The EDE poverty gap corresponding to that approach is defined as

$$\begin{aligned} z - \xi^*(z; \rho, \epsilon) &= z - \mu^*(z)(1 - I^*(z; \rho, \epsilon)) \\ &= \mu^g(z) + \Xi^*(z; \rho, \epsilon) \end{aligned} \quad (59)$$

where $\Xi^*(z; \rho, \epsilon) = \mu^*(z) \cdot I^*(z; \rho, \epsilon)$ is the cost of inequality in censored income and where $I^*(z; \rho, \epsilon)$ is the index of inequality in censored income.

5.5 Poverty curves

It is generally informative to portray the whole distribution of poverty gaps on a simple graph, in a way which shows both the incidence and the inequality of the deprivation in income. Particularly useful are the poverty gap curves, which plot $g(p; z)$ as a function of p . The poverty gap curve shows the "intensity of poverty" felt at each rank in the population. The curve naturally decreases with the rank p in the population, and reaches zero at the value of p equal to the headcount ratio. The integral under the curve gives the average poverty gap, and its steepness, the degree of inequality in the distribution of poverty gaps.

Another quantile-based curve that is graphically informative and that is useful for the measurement and comparison of poverty is called the Cumulative Poverty Gap (CPG) curve (also sometimes referred to as the inverse generalized Lorenz curve, the "TIP" curve, or the poverty profile curve). The CPG curve cumulates the poverty gaps of the bottom p proportion of the population. It is defined as:

$$G(p; z) = \int_0^p g(q; z) dq. \quad (60)$$

A CPG curve is drawn on Figure 5. The slope of $G(p; z)$ at a given value of p shows the poverty gap $g(p; z)$. Since $g(p; z)$ is non-negative, $G(p; z)$ is non-decreasing. $G(p = 1; z)$ equals the average poverty gap $\mu^g(z)$, and the percentile at which $G(p; z)$ becomes horizontal (where $g(p; z)$ becomes zero) yields the poverty headcount. Furthermore, the higher his rank p in the population, the richer is an individual, and therefore the lower is his poverty gap. $G(p; z)$ is therefore concave. Because of this, the CPG curve also enjoys for poverty comparisons the same descriptive interest as the Lorenz and Generalized Lorenz curves for inequality and social welfare analysis. The distance of $G(p; z)$ from the line of perfect equality of poverty gaps (namely, the line 0B in Figure 5) displays the inequality of poverty gaps among the population. The distance of $G(p; z)$ from the line of perfect equality of poverty gaps *among the poor* (namely, the line 0A in Figure 5) displays the inequality of poverty gaps among the poor. Finally, the concavity of $G(p; z)$ shows the density of poverty gaps at p .

5.6 S-Gini poverty indices

When weighted by $\kappa(p; \rho)$, the area underneath the CPG curve generates the class of S-Gini poverty indices:

$$P(z; \rho) = \int_0^1 G(p; z) \kappa(p; \rho) dp. \quad (61)$$

$P(z; \rho = 1)$ equals the average poverty gap, $\mu^g(z)$, $P(z; \rho = 2)$ is the poverty index corresponding to the standard Gini index of inequality, and the well known Sen index of poverty is given by:

$$\frac{P(z; \rho = 2)}{F(z) \cdot z}. \quad (62)$$

An interesting feature of the $P(z; \rho)$ indices is their link with absolute and relative deprivation. Let absolute deprivation, $AD(z)$, be given by the average absolute shortfall from the poverty line, $\mu^g(z)$. Relative deprivation in censored income at percentile p is given by :

$$\bar{\delta}^*(p; z) = \int_p^1 (Q^*(q; z) - Q^*(p; z)) dq. \quad (63)$$

Average relative deprivation across the whole population is thus :

$$RD(z; \rho) = \int_0^1 \bar{\delta}^*(p; z) \kappa(p; \rho) dp. \quad (64)$$

We can then show that:

$$P(z; \rho) = AD(z) + RD(z, \rho). \quad (65)$$

The larger the value of ρ , the larger is relative deprivation, and the larger is $P(z; \rho)$ and the contribution of relative deprivation and inequality in assessing poverty. This is an alternative way of expressing the impact of inequality upon poverty.

5.7 The normalization of poverty indices

Most of the poverty indices discussed above have initially been introduced in the literature in a normalized form, that is, by dividing censored income and poverty gaps by the poverty line. The FGT indices, for instance, are generally expressed as:

$$\bar{P}(z; \alpha) = \int_0^1 \left(\frac{z - Q^*(p)}{z} \right)^\alpha dp \quad (66)$$

(see (51)). The normalization of poverty indices will make no substantial difference and little expositional difference for poverty analysis when the distributions

of income being compared have identical poverty lines. Normalizing poverty indices by the poverty line will make the EDE poverty gap lie between 0 and 1, will make poverty indices insensitive to and independent of the monetary units (e.g., dollars or cents) used in assessing income, and will make the indices invariant to an equi-proportionate change in all incomes and in the poverty line. This is particularly useful if the poverty lines computed are to be interpreted as price indices, and used to enable comparisons of nominal income across time and areas (price indices are used to convert nominal incomes into base-year real incomes). We can thus refer to normalized poverty indices as “relative poverty indices”. FGT and other poverty gap indices that are not normalized can be called “absolute” poverty indices since equal additions to all incomes and to the poverty line will not affect their value.

When poverty lines are different across distributions, and when they act as more than simple price indices, the normalization of poverty indices by these poverty lines can, however, be problematic, and is surely open to criticism. This is the case for instance when we are interested in comparing the absolute shortfalls of “real” income from a “real” poverty line, when these real poverty lines vary across populations or population subgroups. Examples can arise, *inter alia*, in comparing the poverty of families of different sizes and composition, or comparing poverty across countries with different socially or culturally defined poverty lines (whose computation is conceptually separate from the construction of cost of living or price indices).

To see this more clearly, consider the following example, where all incomes and poverty lines are expressed in real terms (namely, adjusted for differences in the cost of living). In country *A*, the poverty line is \$1,000, and a poor person *i* has a living standard of \$500. Because, say, of cultural, sociological and/or economic differences, the poverty line in country *B* is larger and equal to \$2,000, and a poor person *j* in it enjoys a living standard equal to \$1,100. Which of *i* and *j* is the poorest? If we adopt the relative view to building poverty indices, *i* will be considered the poorest since as a proportion of the respective poverty lines he is farther away from it than *j*. If, instead, absolute poverty indices are used, *j* will be deemed the poorest since his absolute poverty gap (\$900) is by far larger than that of *i* (\$500).

5.8 Decomposing differences in poverty

It is often useful to determine whether it is mean-income growth or changes in the relative income shares accruing to different parts of the population that are

responsible for the evolution of poverty across time. This can also help us assess whether these two factors, mean-income changes and distributive changes, work in the same or in opposite directions across time when it comes to monitoring poverty. Similarly, we may wish to assess whether differences in poverty across countries are due to differences in inequality or to differences in mean levels of income.

There are several ways to do this. All of them suffer from a same basic problem known in the national-accounts literature as the "index problem". Say that we have household data from two time periods, denoted by A and B . Should we use the relative shares in distribution A or in distribution B to assess the poverty impact of the differences in mean income across A and B ? Analogously, should we use mean income in A or in B to determine the poverty impact of the differences in the relative income shares between A and B ?

The most popular approach until now, due to Datt and Ravallion, uses the initial distribution as a reference "anchor point" for the assessment of the impact of mean-income and distributive changes on poverty. To see this, it is easiest to use the normalized FGT indices $\bar{P}(z; \alpha)$ defined in (51), although the methodology can be used with any poverty indices, additive or not. The change in poverty between A and B can be expressed as a sum of a "growth" (change in mean income) effect and of a "distributive" (change in relative income shares) effect, plus an error term that originates from the above-mentioned index problem:

$$= \underbrace{\left(\bar{P}_A \left(\frac{z\mu_A}{\mu_B}; \alpha \right) - \bar{P}_A(z; \alpha) \right)}_{\text{growth effect}} + \underbrace{\left(\bar{P}_B \left(\frac{z\mu_B}{\mu_A}; \alpha \right) - \bar{P}_A(z; \alpha) \right)}_{\text{distributive effect}} + \text{error term.} \quad (67)$$

The error term equals

$$\bar{P}_B(z; \alpha) + \bar{P}_A(z; \alpha) - \bar{P}_B \left(\frac{z\mu_B}{\mu_A}; \alpha \right) - \bar{P}_A \left(\frac{z\mu_A}{\mu_B}; \alpha \right). \quad (68)$$

It can be shown to be – and interpreted as – either the difference between the growth effect measured using B as a reference distribution and that using A as the reference distribution,

$$\bar{P}_B(z; \alpha) - \bar{P}_B \left(\frac{z\mu_B}{\mu_A}; \alpha \right) - \left(\bar{P}_A \left(\frac{z\mu_A}{\mu_B}; \alpha \right) - \bar{P}_A(z; \alpha) \right), \quad (69)$$

or as the difference between the distributive effect measured using B as the reference distribution and the distributive effect using A as the reference distribution,

$$\bar{P}_B(z; \alpha) - \bar{P}_A\left(\frac{z\mu_A}{\mu_B}; \alpha\right) - \left(\bar{P}_B\left(\frac{z\mu_B}{\mu_A}; \alpha\right) - \bar{P}_A(z; \alpha)\right). \quad (70)$$

6 Estimating poverty lines

Three major issues arise in the discussion of poverty lines. First, we must define the space in which well-being is to be measured. As outlined above, this can be the space of utility, living standards, "basic needs", functionings, or capabilities. Second, we must determine whether we are interested in an absolute or in a relative poverty line in the space considered. Third, we must choose whether it is by someone's "capacity to attain" or by someone's "actual attainment" that we will judge if that person is poor. We consider here the second issue of the choice between an absolute and a relative poverty line.

6.1 Absolute and relative poverty lines

An absolute poverty line can be interpreted as fixed in any of the spaces in which we wish to assess well-being. A relative poverty line depends on the distribution of well-being (including the utilities, living standards, functionings or capabilities) found in a society. Considerable controversy exists on whether absoluteness or relativity is a better property for a poverty threshold. Most analysts would nevertheless agree that a poverty threshold defined in the space of functionings and capabilities should be absolute (but even on this there is no unanimity). An absolute threshold in these spaces would, however, imply relativity of the corresponding thresholds in the space of the commodities and in the level of basic needs required to achieve these functionings.

There are two main reasons for this. First, the relative prices and the availability of commodities depend on the distribution of living standards. For instance, as a society initially develops, the affordability and accessibility of public transportation usually first increases as rising numbers of people need to travel to work and to trade, without first being able to afford the costs of private transportation. As societies become richer on average, however, their citizens make increasing use of private forms of transportation, which causes a fall in the supply and availability of public transportation, and leads to an increase in its price. This makes the capacity to travel (arguably an important functioning) more or less costly, depending on the state of economic development.

Second, not to be deprived of some capability may require the absence of relative deprivation in the space of some commodities. In support of this, there is Adam Smith's famous statement that the commodities needed to go without shame (an oft-mentioned basic functioning) can be to some extent relative to the distribution of living standards in a society:

By necessities, I understand not only the commodities which are indispensably necessary for the support of life but whatever the custom of the country renders it indecent for creditable people, even of the lowest order, to be without. (Smith (1776))

Sen (1984) reinforces this by distinguishing clearly the two dimensions of capabilities and commodities:

I would like to say that poverty is an absolute notion in the space of capabilities but very often it will take a relative form in the space of commodities and characteristics (Sen ((1984, p.335)).

This has led some writers (particularly in developed countries) to conclude that attempts to preserve some degree of absoluteness in the space of commodities are untenable:

In summary, it does not seem possible to develop an approach to poverty measurement which is linked to absolute standards. While some analysts are uneasy with relativist concepts of poverty on the grounds that they are difficult to comprehend and can be seen as somewhat arbitrary and open to manipulation, no real practical alternative to relativist concepts exists (Saunders (1994), p. 227)

6.2 Social exclusion and relative deprivation

Complete relativity of the poverty line in the space of commodities would nevertheless draw poverty analysis very close to the analysis of social exclusion (as exemplified by Rodgers et al. (1995) at the International Labor Organization) and relative deprivation (as propounded for instance by Townsend (1979)). Social exclusion entails "the drawing of inappropriate group distinctions between free and equal individuals which deny access to or participation in exchange or interaction" [Silver (1994), p.557]. This includes participation in property, earnings, public goods, and in the prevailing consumption level [Silver (1994), p.541]. Relative deprivation focuses on the inability to enjoy living standards and activities that are ordinarily observed in a society. Townsend (1979, p. 30) defines it as a situation in which

Individuals, families and groups in the population (...) lack the resources to obtain the types of diet, participate in the activities and

have the living conditions and amenities which are customary or at least widely encouraged or approved, in the society to which they belong.

Equating absolute deprivation in the space of capabilities with relative deprivation in the space of commodities can, however, be a source of confusion in poverty comparisons. First, it tends to blur the operational and conceptual distinction between poverty and inequality. Second, it can hinder the identification of "core" or absolute poverty in any of the spaces. Identifying core poverty is, however, probably the most relevant task in the design of public policy in developing countries. Third, although the ethical appeal of Sen's capability approach has variously been invoked to justify the use of an entirely relative poverty line in the space of commodities, Sen himself does not accept this:

Indeed, there is an irreducible core of absolute deprivation in our idea of poverty, which translates reports of starvation, malnutrition and visible hardship into a diagnosis of poverty without having to ascertain first the relative picture. Thus the approach of relative deprivation supplements rather than supplants the analysis of poverty in terms of absolute dispossession (Sen (1981, p.17)).

Furthermore,

(...) considerations of relative deprivation are relevant in specifying the 'basic' needs, but attempts to make relative deprivation the sole basis of such specification is doomed to failure since there is an irreducible core of absolute deprivation in the concept of poverty (Sen (1981, p.17)).

Given the measurement difficulties involved in estimating relative poverty lines that correspond adequately to absolute poverty lines in the space of functionings and capabilities, analysts often find most transparent to use the space of living standards as the space in which to define an absolute threshold below which individuals are considered poor. If this is done, however, it must subsequently be admitted that the procedure will imply a set of thresholds in the space of functionings and capabilities that depend at least partly on the conditions of the society in which an individual lives. Indeed, for a given absolute level of living standard in the space of commodities, an individual's capabilities are relative, that is, they depend on his social environment, at least for functionings such as shamelessness and participation in the life of the community.

6.3 Estimating poverty lines

Methodologies for the estimation of poverty lines have been most developed in the context of the need to fulfil basic physiological functionings. Although such methodologies have often been set in a welfarist framework, they are also important for the basic needs, functioning or capability approaches since these approaches are also concerned with basic physiological achievements. These methodologies have recently been most often applied to developing countries.

6.3.1 Cost of basic needs

The estimation of the "cost of basic needs" usually involves two steps. First, an estimation is made of the minimal food expenditures that are necessary for living in good health; we will denote this by z_F . Second, an analogous estimate of the required non-food expenditures, z_{NF} , is computed and added to z_F to yield a total poverty line, z_T . We consider now in some detail each of these two steps.

6.3.2 Cost of food needs

The first step in the computation of a global poverty line is usually to estimate a food poverty line. The determination of a food poverty line generally proceeds by asking what amount of food expenditures is required to achieve some minimal required level of food-energy intake (or nutrient intake, such as proteins, vitamins, fat, minerals, *etc...*). Early examples of the application of this approach include Rowntree (1901) and Orshansky (1965). A basket of food commodities is designed by "food specialists" such as to provide those minimal required levels of food-energy intake. The cost of that basket yields the food poverty line z_F .

To illustrate how this exercise is carried out in practice, consider Figure 6, which plots consumption $x_1(p)$ and $x_2(p)$ of two goods, goods 1 and 2, over a range of percentiles p . For simplicity, Figure 6 supposes that good 1 is "income-inelastic" ($x_1(p)$ is constant) but that the consumption of good 2 increases with the rank in the distribution of income. The idea then is to select a combination of $x_1(p)$ and $x_2(p)$ that provides a given level of minimum calorie intake. For the purposes of our illustration, assume that this minimum energy intake is 3000 calories per day, and that 1 unit of good 1 and 2 provides 2000 and 1000 calories each respectively. Also assume that each unit of good 1 and 2 costs $q\$$.

The cheapest way to achieve the minimum calorie intake would be to consume only of good 1, since good 1 is the most calorie-efficient (we can think of

good 1 as "cereals" and good 2 as "meat"). Indeed, each calorie provided by the consumption of good 1 costs $q\$/2000$, whereas each calorie provided by the consumption of good 2 costs twice as much, that is, $q\$/1000$. 1.5 units of good 1 (1.5 units * 2000 calories/unit = 3000 calories) would then be required for minimal energy intake to be met, and z_F would then equal $1.5q\%$.

This, however, would suppose a food commodity basket that no (or, at the limit, very few) individual in Figure 6 would be observed to consume. Even at the very bottom of the distribution of income, individuals consume indeed at least some of good 2 at the expense of a diminished consumption of the more calorie-efficient good 1. We should presumably take account of this information if we wished to respect somewhat the cultural and culinary preferences of those whose well-being we aim to evaluate. This raises the obvious question of which preferences we should consider. Note that the preferred ratio of good 2 over good 1 increases continuously with p in Figure 6. For convenience, denote that ratio by $\rho(p) = x_2(p)/x_1(p)$. Simple algebra then shows that the cost of attaining the minimum calorie intake is given by $z_F(p) = 3q\$(1 + \rho(p))/(2 + \rho(p))$, where $z_F(p)$ indicates that z_F depends on the rank p of those whose preferences we use to compute the food poverty line.

Figure 6 plots $z_F(p)$ and shows that is not neutral to the choice of p . Using the preferences of the poorest, we obtain $z_F(p = 0) = 1.8q\%$, but if we use the preferences of the median population, we get $z_F(p = 0.5) = 2.1q\%$. This is in fact just a special example of a more general standard observation in the literature on poverty lines that the choice of reference parameters matters for the estimation of poverty lines. In Figure 6, the farther are the preferences $\rho(p)$ from the 0 calorie-efficient choice, the more costly is the estimated food poverty line $z_F(p)$. Arguably, the preferences $\rho(p)$ should be those of individuals around the total poverty line, but this is a (partly) circular argument since $\rho(p)$ is itself a determinant of that total poverty line. In practice, an arbitrary value of p is often chosen, reflecting some *a priori* belief on the position of those at the edge of the total poverty line. A more common (though arguably less commendable) procedure is to compute an average value of $x_2(p)/x_1(p)$ over a range of p , such as the bottom 25% or 50% individuals of a population.

Even if we were to agree on the position p at which we wish to observe preferences such as $\rho(p)$, there still remains the awkward fact that preferences will often vary significantly even at this given value of p . Said differently, there are in practice many different actual consumption patterns for a group of "typical poor". One solution is simply to ignore these differences and estimate the typical poor's average consumption patterns. Following this line, consumption expenditures on

various food items are regressed against income and the estimated parameters of these regressions are then used to predict the consumption patterns of the typical poor. These regressions have often been parametric – assuming for instance that expenditures on cereals and meat are globally quadratic or log-linear in total expenditures. A better statistical procedure would be to regress consumption expenditures non parametrically on total expenditures.

An additionally important issue then is whether variations in culinary tastes and food habits across socio-economic characteristics should be taken into account. If no account of such variations are taken, then we could choose among our reference group the observed diet that minimizes food cost while providing the minimum required level of food-energy intake. This would typically generate an unreasonably low level of expenditures for many of our reference individuals, with an implied dietary basket of food commodities that could again be very different from those they typically consume.

If, however, full account of diversity in culinary tastes were to be taken, a serious risk would exist of overestimating the poverty lines of those individuals and groups of individuals with a greater taste for expensive foods (*e.g.*, of high quality). This is commonly the case, for instance, for urban households, who customarily have more sophisticated culinary tastes than rural dwellers (for the same overall living standards), and have also greater access to a larger variety of imported and expensive foods. This procedure would then assign greater poverty lines to the urban versus the rural individuals. It would also mean that the equivalents of individual food poverty lines in terms of reference living standards and "utilities" would depend on the peculiarities of the individuals' food preferences. This would generally lead to inconsistent comparisons of well-being across urban and rural inhabitants, and would exaggerate the degree of poverty in the urban as compared to the rural areas.

We can illustrate this using Figure 7. Figure 7 shows baskets of two food commodities, x_1 and x_2 , with three food budget constraints of total food consumption equal to Y_1 , Y_2 , and Y_3 (these total budgets are expressed in units of x_1). Figure 7 also shows a "minimum calorie constraint", along which the total calories provided by the consumption of x_1 and x_2 would equal the required minimum level of calorie intake. If no account whatsoever were taken of preferences, Y_0 would yield the food poverty line. But along the food budget constraint Y_0 , there is only one point which meets the minimum calorie constraint, and it is of course unlikely that individuals will *choose* a food basket to be precisely at that corner. An individual with preferences U_0 , for instance, would not locate himself on the minimum calorie constraint. It is only with the more generous budget constraint Y_1

that this individual will consume the minimally required level of calorie intake, as shown on the Figure.

But not all individuals will necessarily choose to be "calorie-sufficient" even with a total food budget of Y_1 . Individuals with greater preferences – as in the case of U_2 – for the less-calorie efficient good x_2 will not choose a food basket on or above the minimum calorie constraint. Individual with preferences U_2 will instead need Y_2 to be calorie-sufficient. Yet, whether individuals with preferences U_1 and budget Y_1 are as well off as individuals with preferences U_2 and budget Y_2 is highly debatable. Such would be the assumption, however, if we used two distinct poverty lines Y_1 and Y_2 for the two different tastes.

As mentioned above, such comparability assumptions are often implicitly made in practice when individuals living in different regions, rural or urban for instance, are assigned different poverty lines for reasons independent of differences in needs or prices. As illustrated in Figure 7, this supposes that an individual with "sophisticated" preferences (an urban dweller who has been accustomed to food variety) needs a higher budget to be as "well off" as an individual with less expensive preferences (a rural dweller who is content with eating basic food types). Probably more convincing, however, would be the view that U_2 with Y_2 in Figure 7 provides greater utility and well-being than U_1 with Y_1 . Assigning different poverty lines Y_1 and Y_2 would then lead to inconsistent or biased poverty estimates.

Minimally required food expenditures can also be (and are often) adjusted for differences in climate, sex, or age, when such differences impact on needs rather than on tastes (as we discussed above). These expenditures can also be adjusted for variations in activity levels, although activity levels depend on the level of one's well-being, and thus on one's poverty status. Activity-level adjustments would thus involve a poverty line that evolves endogenously with the standard of living of individuals, a slightly awkward feature for comparing well-being across individuals and across time.

6.3.3 Non-food poverty lines

The subsequent step is usually to estimate the non-food component of the total poverty line. The most popular method for doing this is simply to go straight to an estimate of the total poverty line by dividing the food poverty line by the share of food in total expenditures. The intuition for this is as follows. The larger the food share in total expenditures, the closer the food poverty line should be to the total poverty line. Therefore, the smallest should be the necessary adjustment to the food poverty line (the closer to 1 should be the denominator that divides the food

poverty line). The problem of which food share to use is of course an important issue. Popular practices vary, but often make use of:

- A- the average food share of those whose total expenditures equal the food poverty line;
- B- the average food share of those whose food expenditures equal the food poverty line;
- C- the average food share of a bottom proportion of the population (*e.g.*, the 25% or 50% poorest).

In addition to this, another popular method

- D- adds to z_F the non-food expenditures of those whose total expenditures equal z_F .

To see how methods A, B and D work and differ from each other, consider Figure 8. Figure 8 shows (predicted) total expenditures against various levels of food expenditures. The regression can be done parametrically, but a generally better approach would be to predict total expenditures using a non-parametric regression on food expenditures. On each of the two axes we have shown the level of the (previously estimated) food poverty line z_F . These two levels meet at the 45 degree line. As indicated above, method A makes use of the average food share of those whose total expenditures equal the food poverty line. Total expenditures equal the food poverty line, z_F , at point E on Figure 8. Hence, the food share at point E is given by the inverse of the slope of the line OE that goes from the origin to point E. The total poverty line according to method A is therefore given by the height of a line OE that extends to just above a level of food expenditures of z_F . This gives the height of point A in the Figure as the total poverty line according to method A.

Method B makes use of the average food share of those whose food expenditures equal the food poverty line. Those who consume z_F in food are located at point B on Figure 8. Their food share is given by the inverse of the slope of the line that extends from point O to point B. Hence, dividing z_F by that food share brings us back to point B, which is therefore the total poverty line according to method B.

The total poverty line according to method B is more generous than that according to method A since the food share used for B is lower than that used for

A. Indeed, method A focusses on the food share of a rather deprived population: those who, *in total*, only spend the *food* poverty line. Method B focusses on the food share of a less deprived population: those who, *on food only*, spend the food poverty line. Since food shares tend to decline with standards of living, method B's food share is lower than method A's.

Finally, method D considers the non-food expenditures of those whose total expenditures equal z_F . As for method A, these individuals are found at point E on Figure 8. Their non-food expenditures are given by the length of line EG on the Figure. Adding these non-food expenditures to z_F yields a total poverty line given by the height of point D.

The choice of methods and food shares and the estimation of the non-food poverty lines is rather arbitrary, and the resulting estimate of the total poverty line will also be somewhat arbitrary. Moreover, and perhaps more worryingly, some of the estimates will also vary with the distribution of living standards, as in the case of method C where the food share is an average over a range of individuals. To avoid inconsistencies in the comparisons of poverty, it is therefore preferable to use the same food share across the distributions being compared, and to use methods that do not make estimates dependent on a particular distribution of living standards.

6.3.4 Food energy intake

A slightly different method for estimating poverty lines that is popular in the literature is the so-called Food-Energy-Intake (FEI) method. Estimates of the observed calorie intake of persons are first computed and then graphed against their observed (total or food) expenditures. The analyst then estimates the expenditures of those whose calorie intake is just at the minimum required for healthy subsistence. When these expenditures are on food, this provides a food poverty line, which can then be used as described above in Section 6.3.3 to provide an estimate of a global poverty line. When the expenditures are total expenditures, the FEI method provides a direct link between a minimum calorie intake and a total poverty line.

Figure 9 illustrates how this method works. The curve shows the level of expenditure (measured on the vertical axis) that is observed (on average) at a given level of calorie intake (shown on the horizontal axis). The curve is increasing and convex, since calorie intake is expected to increase at a diminishing rate with food or total expenditures. Above z_k , the minimum calorie intake recommended for a healthy life, we read z , the food or total poverty line according to the FEI method.

As just exposed, the FEI method may appear straightforward and simple to implement. A number of conceptual and measurement problems are, however, hidden behind this apparent simplicity. First, the line traced on Figure 9 is the *expected* link between expenditure and calorie intake; there is in real life a significant amount of variability around this line. How are we to interpret this variability? If it is due to measurement errors, then we may perhaps ignore it. If it is due to variability in preferences, then we may wish to model the calorie-intake-expenditure relationship separately for different groups of the population, as is often done in practice, for urban and rural areas for instance. As in the cost-of-basic-needs method, however, we then run the risk of estimating higher poverty lines for those groups that have more expensive or sophisticated tastes for food. This would lead to inconsistent comparisons of well-being and poverty, as discussed in Section 6.3.2.

To compute expected expenditure (given the variability of actual observed spending) at a given calorie intake, we can estimate the parameters of a parametric regression linking expenditures to calorie intake. The regression is often postulated to be log-linear or quadratic. This parametric specification supposes, however, that the functional relationship between expenditures and calorie intake is known by the analyst, up to some unknown parameter values. This is unlikely to be true everywhere, especially for those far from the level of calorie intake of interest (*e.g.*, those at the lower and upper tails of the distribution of spending and calorie intake). In such cases, the parametric procedure will make the estimated expenditure poverty line affected by the presence of "outliers" that are relatively far from the minimum level of calorie intake. This procedure will then generate a biased estimator of the "true" poverty line.

An alternative approach estimates the link between expenditures and calorie intake non parametrically. A non-parametric regression does not impose a fixed functional relationship between calorie intake and expenditures along the entire range of calorie intake. On the contrary, it allows a fair amount of flexibility by estimating the link between the two variables through a local weighting procedure. The local weighting procedure looks at the expenditures of those individuals with a calorie intake in the "region" of the specified minimum calorie intake. It weights those values with weights that decrease rapidly with the distance from the minimum calorie intake. Hence, those with calorie intakes far from the minimum specified level will contribute little to the estimated expenditure needed to attain that minimum level. The results using this method are thus less affected by the presence of "outliers" in the distribution of living standards, and less prone to biases stemming from an incorrect specification of the link between spending and

calorie intake.

6.3.5 Illustration for Cameroon

To see whether these differences in methodologies matter, consider the case of Cameroon. Table 4 shows the result of estimating food, non-food and total poverty lines for the whole of Cameroon and for each of 6 regions separately. Note that the figures are in Francs CFA adjusted for price differences, with Yaoundé being the reference region. The food poverty line was estimated using the FEI method at 2400 calories per day per adult equivalent. A non parametric regression using DAD was performed for the whole of Cameroon and separately for each of the 6 regions. The lower non-food poverty line was obtained (non parametrically) using method D in section 6.3.3, and the upper non-food poverty line using method B. Again, the relevant regressions were carried out for the whole of Cameroon and separately for each of the 6 regions.

As can be seen, the link between calorie intake and food expenditures varies systematically across regions. Expected food expenditure at 2400 calories per day are significantly higher in urban areas (Yaoundé, Douala and Other Cities) than in the rural ones. In Douala, for instance, a household would need 408 Francs CFA per day per adult equivalent to reach an intake of 2400 calories per day. In the Highlands, no more than 170 Francs CFA would on average be needed. The link between food and total expenditures also varies across the regions in Cameroon. Combined with the different estimates for the food poverty lines, this leads to very significant variations across regions in the total poverty lines. Using method D, a lower total poverty line of 589 Francs CFA is obtained for Douala, but that same poverty line is only 235 Francs CFA. Note also that the choice of method B *vs* method D has a very significant impact on the estimate of the total poverty line. For the whole of Cameroon, the lower and the upper total poverty lines are respectively 373 and 534 Francs CFA, a difference of 43%.

Unsurprisingly, these large differences across regions and across methods have a large impact on poverty estimates and on regional poverty comparisons. This is illustrated in Table 6, which shows the proportion of individuals underneath various poverty lines for various indicators of well-being. "Calorie poverty" (first column) is fairly constant across Cameroon. In the whole of Cameroon, 68.1% of the population was observed to consume less than 2400 calories per day per adult equivalent. This proportion varies between 59.9% (for Other Cities) and 86.5% (for Forest) across regions. Roughly the same limited variability and the same poverty rankings appear when food poverty is estimated using for each region its

own food poverty line (third column). However, when a *common* food poverty line is used to assess food poverty in each region (second column), national poverty stays roughly unchanged at around 69% but urban regions now appear significantly less poor than the rural ones. For instance, the poverty headcount in Douala (42.0 %) is now only half that of the Highlands (82.5 %).

The rest of Table 6 confirms these lessons. When a common poverty line is used to compare the regions, rural areas are very significantly poorer than urban ones. When region-specific poverty lines are used, these differences are much reduced, and the regional rankings are often even reversed. For example, using a common lower total poverty line (fourth column), the Highlands have a headcount ratio more than three times that of the urban regions. When regional lower total poverty lines are used instead, the Highlands become prominently the least poor of all regions. Setting common as opposed to regional poverty lines can thus have a crucial impact on poverty rankings and the determination of subsequent poverty alleviation policies. The choice of a lower as against an upper total poverty line also makes a difference. For the whole of Cameroon, the proportion of the Cameroonian population in poverty increases from 43.9% to 68.0% when we move from a common lower total poverty line (fourth column) to a common upper total poverty line (sixth column). Clearly, this changes significantly one's understanding of the incidence of poverty in Cameroon.

These results also implicitly caution that the choice of well-being indicators is not neutral to the identification of the poor. In our context, this is because the correlation between calorie intake, food expenditure and total expenditure is imperfect. Table ?? indicates, for example, that in bidimensional poverty analyses using any two of these three indicators of well-being, around 20% to 25% of the population is characterized as poor in one dimension but non poor in the other. In the first part of ??, we note for instance that 11.2% of the population would be judged poor in terms of calorie intake but not poor in terms of food expenditure. Conversely, 9.6% of the population would be deemed non poor in terms of calorie intake but poor in terms of food expenditure. These proportions are slightly higher for the other bidimensional poverty analyses, which compare food with total expenditure poverty, and calorie with total expenditure poverty, respectively.

6.3.6 Relative and subjective poverty lines

Relative poverty lines

There are two other popular methodologies for the estimation of poverty lines. The first deals with relative poverty lines, which, as we saw above, can be useful

to determine the commodities needed for "living without shame" and for participating in the "prevailing consumption level". A relative poverty line is typically set as an arbitrary proportion (often around 50%) of the mean or the median of living standards. Clearly, such a poverty line will vary with the central tendency of the distribution of living standards, and will not be the same across regions and time. One awkward feature of the use of a relative poverty line approach is that a policy which raises the living standards of all, but proportionately more those of the rich, will increase poverty, although the absolute living standards of the poor have risen. Conversely, a natural catastrophe which hurts absolutely everyone will decrease poverty if the rich are proportionately the most hurt. When used alone, relative poverty lines can thus be shown to drift the analysis towards the concept of relative inequality, and away from absoluteness of deprivation in any of the poverty measurement spaces define above. Because of this, they are probably best used in conjunction with absolute living standard thresholds, at least when the aim is to capture both absolute deprivation in basic physiological capabilities and social exclusion and relative deprivation in more social capabilities.

Subjective poverty lines

The second alternative poverty line approach relies on the use of subjective information on the link between living standards and well-being. One source of information comes from interviews on what is perceived to be a sound poverty line, using a query found for instance in Goedhart *et al.* (1977):

We would like to know which net family income would, in your circumstances, be the absolute minimum for you. That is to say, that you would not be able to make both ends meet if you earned less.

The answers are subsequently regressed on the living standards of the respondents. The subjective poverty line is given by the point at which the predicted answer to the minimum income question equals the living standard of the respondents. The idea is that unless someone earns that poverty line, he will not truly know that it is indeed the appropriate minimum income needed to "make both ends meet".

This method is illustrated on Figure 10. Each point represents a separate answer to the above query, namely, the minimum income judged to be needed to make both ends meet as a function of the actual income of the respondents. The filled line shows the predicted response of individuals at a given level of income. For low income levels, this predicted minimum subjective income is well above the respondents' income. The predicted minimum subjective income increases with actual income, but not as fast as income itself, such that at z^* (on the 45-degree line), the predicted minimum subjective income equals actual income.

One difficulty with the subjective approach is the sensitivity of poverty line estimates to the formulation of interview questions. Perhaps a more fundamental disquieting output is the considerable variability in the answers provided, even within groups of relatively socio-economically homogeneous respondents. The presence of this variability is apparent on Figure 10 with points sometimes quite far away from the predicted response line. This variability has some awkward consequences. On Figure 10, for instance, point a is a point at which someone would be judged poor according to the subjective income method, since his income falls below z^* . An individual at a feels, however, that his income exceeds the minimum income he feels to be needed (point a is to the right of the 45-degree line). *He* would therefore feel that he is not poor. Conversely, someone at point b feels that he is poor, since his reported minimum income exceeds his actual income, but he would be judged not to be poor by the subjective poverty line method.

How, therefore, ought we to interpret this variability? Is it due to measurement errors? If so, then we may best ignore it. Is it rather that the link between living standards and real well-being varies systematically within homogeneous groups of people? If so, then we should not attempt to use living standards or other direct or indirect indicators of well-being to classify the poor and the non poor. Instead, we should perhaps take individuals at word on whether they declare themselves to be poor or not. But this would alternatively raise important practical problems for the assessment and the implementation of public policy. Can public policy rely appropriately and confidently for its implementation on the provision of subjective information on the part of individuals?

Subjective poverty lines with discrete information

An alternative approach to estimating subjective poverty lines is to ask respondents whether they feel that their living standards are below the poverty line, without direct indications of what the value of that poverty line may well be. Answers are coded 0 or 1 – according to whether respondents feel that they are poor or not – alongside the respondents' living standards. This is illustrated in Figure 11. Each "dot" is an observation of whether a respondent of a certain income level felt poor (1) or not (0). The implicit assumption is that respondents compare their income to a *common* subjective poverty line z^* . z^* is unobserved and must be estimated. Not everyone with an income below z^* says that he is poor; conversely, not everyone above z^* says that he is not poor. These "classification errors" would be explained by measurement and/or misreporting errors. Hence, on Figure 11, there are "false poor" and "false rich", as shown within the ellipses at the bottom left and at the top right of the Figure.

The estimation of z^* proceeds by minimizing the probability of observing false

poor and false rich. Said differently, the estimator of z^* is that which minimizes the likelihood of observing observations within the ellipses in Figure 11.

7 The measurement of progressivity, equity and re-distribution

7.1 Taxes and concentration curves

Let X and N represent respectively gross and net incomes. Gross income is pre-tax and/or pre-transfer income, and net income is post-tax and/or post-transfer income. The mapping of gross incomes into net incomes is done by T , with $N = X - T$. For expositional simplicity, we assume throughout that gross incomes are exogeneous. We can expect a part of T to be a tax and transfer function, $T(X)$, that depends on the value of gross income X . For several reasons, we also expect T to be stochastically related to X . In real life, taxes and transfers depend on a number of variables other than gross incomes, such as family size and composition, age, sex, area of residence, sources of income, consumption and savings behaviour, and the ability to avoid taxes or claim transfers. Thus, we can think of T as being a stochastic function of X ,

$$T = T(X) + \nu, \quad (71)$$

where ν is a stochastic tax determinant. We will denote by $F_{X,N}(\cdot, \cdot)$ the joint cumulative distribution function (*cdf*) of gross and net income. Let $X(p)$ be the p -quantile function for gross incomes, and $N(q)$ the q -quantile function for net incomes. Let $F_{N|X=x}(\cdot)$ be the *cdf* of N conditional on $X = x$. The q -quantile function for net incomes conditional on a p -quantile value for gross incomes is then technically defined as $N(q|p) = \inf\{s > 0 | F_{N|X=X(p)}(s) \geq q\}$ for $q \in [0, 1]$. $N(q|p)$ thus gives the net income of the individual whose net income rank is q among all those whose rank is p in the distribution of gross incomes.

The expected net income of those at rank p in the distribution of gross income is given by:

$$\bar{N}(p) = \int_0^1 N(q|p) dq. \quad (72)$$

The expected net tax of those at rank p in the distribution of gross income is then given by:

$$\bar{T}(p) = \bar{N}(p) - X(p). \quad (73)$$

The concentration curve for T is:

$$C_T(p) = \frac{\int_0^p \bar{T}(q) dq}{\mu_T} \quad (74)$$

where $\mu_T = \int_0^1 \bar{T}(p) dp = \mu_N - \mu_X$ is average taxes across the population. $C_T(p)$ shows the proportion of total taxes paid by the p bottom proportion of the population. In practice, concentration curves are often estimated by ordering a finite number of sample observations $(X_1, T_1), \dots, (X_n, T_n)$ in increasing values of gross incomes, such that $X_1 \leq X_2 \leq \dots \leq X_n$, with percentiles $p_i = i/n$, $i = 1, \dots, n$. For $i = 1, \dots, n$, the discrete (or “empirical”) concentration curve for taxes is then defined as:

$$C_T(p = i/n) = \frac{1}{n\mu_T} \sum_{j=1}^i T_j. \quad (75)$$

As for the empirical Lorenz curves, other values of $C_T(p)$ can be estimated by interpolation. The concentration curve $C_N(p)$ for net incomes is analogously defined as:

$$C_N(p) = \frac{\int_0^p \bar{N}(q) dq}{\mu_N} \quad (76)$$

and typically estimated as:

$$C_N(p = i/n) = \frac{1}{n\mu_N} \sum_{j=1}^i N_j \quad (77)$$

where the N_j have been ordered in increasing values of the associated X_j . Note that $C_N(p)$ is different from the Lorenz curve of net incomes, $L_N(p)$, defined as:

$$L_N(p) = \frac{\int_0^p N(q) dq}{\mu_N}. \quad (78)$$

Empirically, the Lorenz curve for net income is typically estimated as

$$L_N(p = i/n) = \frac{1}{n\mu_N} \sum_{j=1}^i N_j \quad (79)$$

where the observations have been ordered in increasing values of *net incomes*. Thus, $C_N(p)$ sums up the expected value of net incomes (conditional on a percentile of gross incomes) up to gross income percentile p . $L_N(p)$, however, sums up net incomes up to a net income percentile p . We can show that $C_N(p)$ will never be lower than $L_N(p)$, and will be strictly greater than $L_N(p)$ for at least one value of p if there is “reranking” in the redistribution of incomes (this feature will prove useful later in building indices of reranking)³. Intuitively, $C_N(p)$ cumulates some net incomes whose percentiles in the net income distribution exceed p . These are net incomes that are high compared to the expected net income at a gross income of percentile p or lower. Such high incomes are nevertheless possible, however, due to the stochastic term ν . $L_N(p)$ only cumulates the net incomes whose percentile in the net income distribution is p or lower. Hence, $C_N(p) \geq L_N(p)$. This can also be seen by comparing the estimators in equations (77) and (79). In (79), observations of N_j are cumulated in increasing values of N_j , but in (77), observations of N_j are cumulated in increasing values of X_j , which means that some higher values of N_j may be cumulated before some lower ones.

Denote the average tax rate as a proportion of average gross income as t , with $t = \mu_T/\mu_X$. When $t \neq 0$, we can show the following:

$$C_N(p) - L_X(p) = \frac{t}{1-t} [L_X(p) - C_T(p)]. \quad (80)$$

For a positive t , this indicates that the more concentrated are the taxes among the poor (the smaller the difference $L_X(p) - C_T(p)$), the less concentrated among the poor will net incomes be. The reverse is true for transfers (negative t): the more concentrated they are among the poor, the more concentrated net income is among the poor. This link will prove useful later in defining indices of tax progressivity.

7.2 Indices of concentration

As for the Lorenz curves and the S-Gini indices of inequality, we can aggregate the distance between p and the concentration curves $C(p)$ to obtain indices of concentration. These indices of concentration are useful to compute aggregate indices of progressivity and vertical equity. More generally, they can also serve to decompose the inequality in total income or total consumption into a sum of the inequality in the components of that total income or consumption, such as different sources of income (different types of earnings, interests, dividends, capital

³In a continuous distribution, a sufficient condition for reranking is that ν is not degenerate, namely, that it is not a constant.

gains, taxes, transfers, etc.) or different types of consumption (of food, clothing, housing, etc.). To define indices of concentration, we can simply weight the distance $p - C(p)$ by a weight $\kappa(p)$, of which a popular form is again given by $\kappa(p; \rho)$ in equation (12). This gives the following class of S-Gini indices of concentration:

$$IC(\rho) = \int_0^1 (p - C(p))\kappa(p; \rho)dp. \quad (81)$$

For instance, let X_1 and X_2 be two types of consumption, and let $X = X_1 + X_2$ be total consumption. Let $C_{X_1}(p)$ and $C_{X_2}(p)$ be the concentration curves of each of the two types of consumption (using X as the ordering variable). The concentration indices for X_1 and X_2 are as follows :

$$IC_{X_1}(\rho) = \int_0^1 (p - C_{X_1}(p))\kappa(p; \rho)dp \quad (82)$$

and

$$IC_{X_2}(\rho) = \int_0^1 (p - C_{X_2}(p))\kappa(p; \rho)dp. \quad (83)$$

Inequality in X can then be decomposed as a sum of the inequality in X_1 and in X_2 . The Lorenz curve for total consumption is given by:

$$L_X(p) = \frac{(\mu_{X_1}C_{X_1}(p) + \mu_{X_2}C_{X_2}(p))}{\mu_X} \quad (84)$$

which is a simple weighted sum of the concentration curves for each of the two types of consumption. The index of inequality in total consumption is similarly a simple weighted sum of the concentration indices of each of the two types of consumption:

$$I_X(\rho) = \frac{\mu_{X_1}IC_{X_1}(\rho) + \mu_{X_2}IC_{X_2}(\rho)}{\mu_X}. \quad (85)$$

For given μ_{X_1} and μ_{X_2} , the higher the concentration indices $IC_{X_1}(\rho)$ and $IC_{X_2}(\rho)$, the larger the S-Gini index of inequality in total consumption. The higher the share μ_{X_i}/μ_X of the more highly concentrated expenditure, the higher the inequality in total expenditures.

7.3 Progressivity comparisons

7.3.1 Deterministic tax and benefit systems

Let us for a moment assume that the tax system is non-stochastic (or deterministic), namely, that ν equals a constant zero. Suppose also for now that the deterministic tax system does not rerank individuals, or equivalently that $T^{(1)}(X) \leq 1$. Denote the average rate of taxation at gross income X by $t(X) = T(X)/X$. A net tax (possibly including a transfer or subsidy) $T(X)$ is said to be progressive if the average rate of taxation increases with X , that is, if $t^{(1)}(X) > 0$; it is proportional if $t^{(1)}(X) = 0$, and it is regressive if $t^{(1)}(X) < 0$. There are two popular measures to capture the change in taxes and net income as gross income increases. One is the elasticity of taxes with respect to X , and is called Liability Progression, $LP(X)$:

$$LP(X) = \frac{X}{T(X)} T^{(1)}(X) = \frac{T^{(1)}(X)}{t(X)} \quad (86)$$

and is therefore simply the ratio of the marginal tax rate over the average tax rate. The larger this measure at every X , the more concentrated among the richer are the taxes. It is possible to show that a tax system is uniformly progressive (namely, $t^{(1)}(X) > 0$ everywhere) if $LP(X) > 1$ everywhere. One problem with $LP(X)$ is that it is not defined when $T(X) = 0$, and that it is awkward to interpret when a net tax is sometimes negative and sometimes positive across gross income. Another problem is that it is linked to the relative distribution of taxes, not with the relative distribution of the associated net incomes. These problems are avoided by a second measure, called Residual Progression ($RP(X)$), which is the elasticity of net income with respect to gross income:

$$RP(X) = \frac{\partial(X - T(X))}{\partial X} \cdot \frac{X}{N} = \frac{1 - T^{(1)}(X)}{1 - t(X)}. \quad (87)$$

Unlike $LP(X)$, $RP(X)$ is well defined and easily interpretable even when taxes are sometimes negative, positive or zero, so long as gross and net incomes are strictly positive.

It is then possible to show that a tax system is uniformly progressive (again, this means that $t^{(1)}(X) > 0$ everywhere) if $RP(X) < 1$ everywhere. This will make the distribution of net incomes unambiguously more equal than the distribution of gross incomes, regardless of that actual distribution of gross incomes. Moreover, if the residual progression for a tax system A is always lower than that of a tax system B , whatever the value of X , then the tax system A is said to be

uniformly more progressive than the tax system B , and the distribution of net incomes will always be more equal under A than under B , again regardless of the distribution of gross incomes.

Hence, an important distributive consequence of progressive taxation is to make the inequality of net incomes lower than that of gross income. Analogously, proportional taxation will not change inequality, and regressive taxation will increase inequality. The more progressive the tax system, the more inequality-reducing it is. To check whether a deterministic tax system is progressive, proportional or regressive, we may thus simply plot the average tax rate as a function of X and observe its slope. Alternatively, we may estimate and graph its liability progression or its residual progression at various values of X . To check whether a tax system is more progressive (and thus more redistributive) than another one, we simply plot and compare the elasticity of net incomes with respect to gross incomes. All of this can be done using non-parametric regressions of $T(X)$ and N against X .

Another informative descriptive approach is to compare the share in taxes and benefits to the share in the population of individuals at various ranks in the distribution of income. This is most easily done by plotting on a graph the ratios $T(X)/\mu_T$ or $\bar{T}(p)/\mu_T$ for various values of gross income X or ranks p in the population. If these ratios exceed 1, then those individuals with those incomes or ranks pay a greater share of total taxes than their population share. A similar intuition applies when $T(\cdot)$ is a benefit: a ratio $T(X)/\mu_T$ or $\bar{T}(p)/\mu_T$ that exceeds 1 indicates that the benefit share exceeds the population share.

A competing descriptive tool is to plot the ratio of taxes or benefits $T(X)$ over gross income X , that is, $T(X)/X$, perhaps assessed at some rank p to give $\bar{T}(p)/X(p)$. Such a graph shows how the average tax rate evolves with gross income or ranks. This can show quite a different picture from that shown by $T(X)/\mu_T$ or $\bar{T}(p)/\mu_T$.

7.3.2 General tax and benefit systems

Although graphically informative, the above simple descriptive approaches present, however, some problems. First, if $T^{(1)}(X) > 1$, the tax system will induce reranking, even if it is a deterministic function of X . As we will see below, reranking (and, more generally, horizontal inequity) decreases the redistributive effect of taxation, besides being of course of significant ethical concern in its own right. Second, and more importantly in empirical applications, taxes are typically not a deterministic function of gross income, and randomness in taxes will introduce

greater variability and inequality in net incomes than the above deterministic approach would predict. $X - T(X)$ may then be an unreliable guide to the distribution of net incomes. Randomness in taxes will also introduce further reranking. These two features will reduce the redistributive effect of the tax, and may even in the most extreme cases increase inequality even when the “deterministic trend” of the tax is progressive, or such that $t^{(1)}(X) > 0$. Finally, the actual redistribution effected by taxes depends on the distribution of gross incomes, and not only on the shape of the tax function T . Said differently, the actual redistributive effect of liability or residual progression will depend on the actual distribution of gross incomes.

To deal with these difficulties, we can use the actual distribution of taxes T and net incomes N (instead of their predicted values $T(X)$ and $X - T(X)$) to determine whether the actual tax system is progressive and inequality-reducing. There are two leading approaches for this exercise. The first is the Tax-Redistribution (TR) approach, and the second is the Income-Redistribution (IR) approach.

A tax T is be TR -progressive if

$$C_T(p) < L_X(p) \text{ for all } p \in]0, 1[. \quad (88)$$

A benefit B is TR -progressive if

$$C_B(p) > L_X(p) \text{ for all } p \in]0, 1[. \quad (89)$$

A tax T' is more TR -progressive than a tax T^0 if

$$C_{T'}(p) < C_{T^0}(p) \text{ for all } p \in]0, 1[. \quad (90)$$

A benefit B' is more TR -progressive than a benefit B^0 if

$$C_{B'}(p) > C_{B^0}(p) \text{ for all } p \in]0, 1[. \quad (91)$$

A tax T is more TR -progressive than a benefit B if

$$L_X(p) - C_T(p) > C_B(p) - L_X(p) \text{ for all } p \in]0, 1[. \quad (92)$$

A tax (and/or a transfer) T is IR -progressive if

$$C_N(p) > L_X(p) \text{ for all } p \in]0, 1[. \quad (93)$$

A tax (and/ or a transfer) T' is more IR -progressive than a tax (and/or a transfer) T^0 if

$$C_{N'}(p) > C_{N^0}(p) \text{ for all } p \in]0, 1[. \quad (94)$$

These two *TR* and *IR* approaches are consistent with the use above of liability and residual progression in a deterministic tax system. If $t^{(1)}(X) > 0$ and $T^{(1)}(X) \leq 1$ (namely, no reranking), then, whatever the actual distribution of gross incomes, $T(X)$ is both *TR*- and *IR*-progressive.

Note that these progressivity comparisons have as a reference point the initial Lorenz curve. In other words, a tax is progressive if the poorest individuals bear a share of the total tax burden that is less than their share in total gross income. As mentioned above, another view is that the poor should pay a tax burden that is proportional to their share in the population. This is more often argued in the context of benefits – the reference point to assess the equity of public expenditures is population share. The analytical framework above can easily allow for this alternative view – simply by replacing $L_X(p)$ by p in the above definitions of *TR* progressivity. This will make more stringent the conditions to declare a benefit progressive, but it will also make it easier for a tax to be declared progressive – see (88) and (89).

7.4 Reranking and horizontal inequity

As is well-known, the assessment of tax and transfer systems draws on two fundamental principles: efficiency and equity⁴. The former relates to the presence of distortions in the economic behavior of agents, while the latter focuses on distributive justice. Vertical equity as a principle of distributive justice is rarely questioned as such, although the extent to which it must be precisely weighted against efficiency is of course a matter of intense disagreement among policy analysts. In this section, we examine in more detail a more neglected aspect of the notion of equity: horizontal equity (*HE*) in taxation (including negative taxation).

Two main approaches to the measurement of *HE* are found in the literature, which has evolved substantially in the last twenty years. The classical formulation of the *HE* principle prescribes the equal treatment of individuals who share the same level of welfare before government intervention. *HE* may also be viewed as implying the absence of reranking: for a tax to be horizontally equitable, the ranking of individuals on the basis of pre-tax welfare should not be altered by a fiscal system.

Why should concerns for horizontal equity influence the design of an optimal tax and transfer system? Several answers have been provided, using either of

⁴This section draws heavily from a joint paper written with Vincent Jalbert and Abdelkrim Araar.

two approaches. The traditional or “classical” approach defines *HE* as the equal treatment of equals (see Musgrave (1959)). While this principle is generally well accepted, different rationales are advanced to support it. First, a tax which discriminates between comparable individuals is liable to create resentment and a sense of insecurity, possibly also leading to social unrest. This is supported by the socio-psychological literature which shows that exclusion and discrimination have an impact both on individual well-being and on social cohesion and welfare. For instance, status/role structure theory indicates that one’s relative socio-economic position (and its variability) “may give rise to definable and measurable social and psychological reactions, such as different types of alienation” (Durant and Christian (1990), p.210).

Second, the principles of progressivity and income redistribution, which are key elements of most tax and transfer systems, are generally undermined by *HI* (as we shall see in our own treatment below). This has indeed been one of the main themes in the development of the reranking approach in the last decades (see for instance Atkinson (1979) and Jenkins (1988)). Hence, a desire for *HE* may simply derive from a general aversion to inequality, without any further appeal to other normative criteria. Feldstein (1976) also notes that when utility functions are identical across individuals, a utilitarian social welfare function is maximized when equal incomes are taxed equally. This result then makes the principle of *HE* become a corollary of the principle of *VE*. A separate justification of *HE* would, however, generally be required when preferences are heterogeneous (they usually are), and because in some circumstances a random tax can otherwise be found to be optimal (Stiglitz (1982)). *HI* may moreover suggest the presence of imperfections in the operation of the tax and transfer system, such as an imperfect delivery of social welfare benefits, attributable to poor targeting or to incomplete take-up (see Duclos (1995b)). It can also signal tax evasion, which can *inter alia* cost the government significant losses of tax revenue (see Bishop *et al.* (1994)).

Third, *HE* can be argued to be an ethically more robust principle than *VE* (*VE* relates to the reduction of welfare gaps between unequal individuals). The *HE* principle is often seen as a consequence of the fundamental moral principle of the equal worth of human beings, and as a corollary of the equal sacrifice theories of taxation. Depending on the retained specification of distributive fairness, the requirements of vertical justice can vary considerably, while the principle of horizontal equity remains essentially invariant (Musgrave (1990)). Plotnick (1982) also supports this view by arguing that *HI* in the redistributive process would cause a loss of social welfare relative to an horizontally equitable tax, regardless of any *VE* value judgments on the final distribution. This has led several authors

(including Stiglitz (1982), Balcer and Sadka (1986) and Hettich (1983)) to advocate that *HE* be treated as a separate principle from *VE*, and thus to form one of the objectives between which an optimal trade-off must be sought in the setting of tax policy.

The value of studying classical *HI* has nonetheless been questioned by a few authors, among whom figures Kaplow (1989, 1995), who rejects the premise that the initial distribution is necessarily just (see also Atkinson (1979) and Lerman and Yitzhaki (1995)) and adds that utilitarianism and the Pareto principle may justify the unequal treatment of equals (as seen above)⁵. A number of authors have also expressed dissatisfaction with the classical approach to *HE* because of the implementation difficulties it was seen to present. Indeed, since no two individuals are ever exactly alike in a finite sample, it was argued (see *inter alia* Feldstein (1976) and Plotnick (1982,1985)) that analysis of equals had to proceed on the basis of groupings of unequals which were ultimately arbitrary and which represented “an artificial way to salvage empirical applicability” (Plotnick (1985), p. 241). The proposed alternative was then to link *HI* and reranking and to note that the absence of reranking *implies* the classical requirement of *HE*: “the tax system should preserve the utility order, implying that if two individuals would have the same utility level in the absence of taxation, they should also have the same utility level if there is a tax” (Feldstein (1976), p.94)⁶.

Various other ethical justifications have also been suggested for the requirement of no-reranking. For normative consistency, King (1983) argues for adding the qualification “and treating unequals accordingly” to the classical definition of *HE*, by which it then becomes clear that classical *HE* also *implies* the absence of reranking. Indeed, if two unequals are reranked by some redistribution, then it could be argued at a conceptual level that at a particular point in that process of redistribution, these two unequals became equals and were then made unequal (and reranked), thus violating classical *HE*. Hence, from the above, “a necessary and sufficient condition for the existence of horizontal inequity is a change in ranking between the ex ante and the ex post distributions” (King (1983), p. 102). King (1983) then proposes an additively separable social welfare function that is characterized by a parameter of aversion to horizontal inequity and ver-

⁵The same is true for the reranking of individuals, which is discussed below. See also King(1983), who sees this implication as a flaw of strict utilitarianism since it ignores the fairness of the redistributive process.

⁶The requirement of no reranking further implies that marginal tax rates should not exceed 100%, which can be taken as a basic economic requirement for incentive preservation and efficiency. On this, see *inter alia* Lambert and Yitzhaki (1995).

tical inequality. This function decreases with the distance of each net income from its order-preserving level of net income. Chakravarty (1985) also argues that reranking causes an individual's utility to differ from what it would have been otherwise, and moreover suggests that this difference reduces the final level of individual utility, creating a loss of social welfare as measured again by a utilitarian welfare function.

The theory of relative deprivation also suggests that people often specifically compare their relative individual fortune with that of others in similar or close circumstances. The first to formalize the theory of relative deprivation, Davis (1959), expressly allowed for this by suggesting how comparisons with similar vs dissimilar others lead to different kinds of emotional reactions; he used the expression "relative deprivation" for "in-group" comparisons (*i.e.*, for *HI*), and "relative subordination" for "out-group" comparisons (*i.e.*, for *VE*) (Davis (1959), p.283). Moreover, in the words of Runciman (1966), another important contributor to that theory, "people often choose reference groups closer to their actual circumstances than those which might be forced on them if their opportunities were better than they are" (p.29).

In a discussion of the post-war British welfare state, Runciman also notes that "the reference groups of the recipients of welfare were virtually bound to remain within the broadly delimited area of potential fellow-beneficiaries. It was anomalies within this area which were the focus of successive grievances, not the relative prosperity of people not obviously comparable" (p.71). Finally, in his theory of social comparison processes, Festinger (1954) also argues that "given a range of possible persons for comparison, someone close to one's own ability or opinion will be chosen for comparison" (p.121). In an income redistribution context, it is thus plausible to assume that comparative reference groups are established on the basis of similar gross incomes and proximate pre-tax ranks, and that individuals subsequently make comparisons of post-tax outcomes across these groups. Individuals would then assess their relative redistributive ill-fortune in reference groups of comparables by monitoring *inter alia* whether they are overtaken by or overtake these comparables in income status, thus providing a plausible "micro-foundation" for the use of no-reranking as a normative criterion.

This suggests that comparisons with close individuals (but not necessarily exact equals) would be at least as important in terms of social and psychological reactions as comparisons with dissimilar individuals, and thus that analysis of *HI* and reranking in that context should be at least as important as considerations of *VE*. It also says that, although classical *HI* and reranking are both necessary and sufficient signs of *HI*, they are (and will be perceived as) different manifestations

of violations of the *HE* principle⁷. Hence, it seems reasonable that *VE* and the no-reranking requirement be assessed separately from the classical requirement of equal treatment of equals when assessing the impact of taxes and transfers on social welfare⁸.

How do we analyze this empirically?

Show the variability of $T(X)$: $\sigma_T(X)$ and $\sigma_T(X(p))$

Also: a tax (and/or a transfer) T causes reranking (and hence horizontal inequity) if $C_N(p) > L_N(p)$ for at least one value of $p \in]0, 1[$.

7.5 Redistribution

It is possible to decompose the net redistributive effect of taxes and transfers into components that are called vertical equity (*VE*) and horizontal inequity (*HI*). The *VE* effect measure the tendency of a tax system to “compress” the distribution of net incomes. As we will see, it is linked to the progressivity of the tax system. The *HI* term, which contributes negatively to the net redistributive effect of the tax system (hence the use of horizontal “inequity”), accounts for the “unequal tax treatment of equals”. Depending on the choice of the underlying social welfare function or inequality index, the horizontal inequity term can either take the form of a “classical” horizontal inequity effect or of a “reranking” effect. Our use of Lorenz and concentration curves and of the associated S-Gini indices of inequality and redistribution will make us focus here on reranking indices of horizontal inequity.

The difference between the Lorenz curve of net and gross incomes is given by:

$$L_N(p) - L_X(p) = \underbrace{C_N(p) - L_X(p)}_{VE} - \underbrace{(C_N(p) - L_N(p))}_{HI}. \quad (95)$$

The first term, *VE*, is clearly linked to the definition of *IR*-progressivity in equation (93). As shown in equation (80), it can also be expressed in terms of *TR*-progressivity when $t \neq 0$:

⁷This is nicely discussed in the survey by Jenkins and Lambert (1999) and by Plotnick’s (1999) comments on it. See also Kaplow (1989), who observes that the goal of limiting reranking may conflict with that of limiting the unequal treatment of equals.

⁸In the words of Feldstein (1976): “The problem for tax design is therefore to balance the desire for horizontal equity against the utilitarian principle of welfare maximisation. Balancing these two goals requires an explicit measure of the departure from horizontal equity.” (p.83).

$$C_N(p) - L_X(p) = \frac{t}{1-t} [L_X(p) - C_T(p)]. \quad (96)$$

Furthermore, if there is more than one tax and/or benefit that make up T , we can decompose total VE as a sum of the IR and TR progressivity of each tax and transfer. Let t_j be the overall average rate of the tax T_j , with $j = 1, \dots, J$, such that $\sum_{j=1}^J t_j = t$, and let $C_{T_j}(p)$ and $C_{N_j}(p)$ be the concentration curves of net income and taxes corresponding to tax T_j . Then:

$$C_N(p) - L_X(p) = \sum_{j=1}^J \frac{(1-t_j)}{(1-t)} (C_{N_j}(p) - L_X(p)) \quad (97)$$

where $C_{N_j}(p) - L_X(p)$ is the vertical equity of tax or transfer j at percentile p , and again can be easily seen to be an element of the definition of IR -progressivity. Each of these VE contributions can also be expressed as a function of TR progressivity at p (when $t_j = 0$):

$$C_{N_j}(p) - L_X(p) = \frac{t_j}{1-t_j} [L_X(p) - C_{T_j}(p)]. \quad (98)$$

7.6 Indices of progressivity and redistribution

As for comparisons of inequality and concentration, it is often useful to summarise the progressivity, vertical equity, horizontal inequity as well as the redistributive effect of taxes and transfers into summary indices. We can do this by weighting the differences expressed above by the weights of the S-Gini indices ($\kappa(p; \rho)$) to obtain S-indices of TR -progressivity ($IT(\rho)$), IR -progressivity and vertical equity ($IV(\rho)$), horizontal inequity and reranking ($IH(\rho)$), and redistribution ($IR(\rho)$):

$$IT(\rho) = \int_0^1 (L_X(p) - C_T(p)) \kappa(p; \rho) dp \quad (99)$$

$$IV(\rho) = \int_0^1 (C_N(p) - L_X(p)) \kappa(p; \rho) dp \quad (100)$$

$$IH(\rho) = \int_0^1 (C_N(p) - L_N(p)) \kappa(p; \rho) dp \quad (101)$$

$$IR(\rho) = \int_0^1 (L_N(p) - L_X(p)) \kappa(p; \rho) dp. \quad (102)$$

These indices can also be computed as differences between S-Gini indices of inequality and concentration:

$$IT(\rho) = IC_T(\rho) - I_X(\rho) \quad (103)$$

$$IV(\rho) = I_X(\rho) - IC_N(\rho) \quad (104)$$

$$IH(\rho) = I_N(\rho) - IC_N(\rho) \quad (105)$$

$$IR(\rho) = I_X(\rho) - I_N(\rho). \quad (106)$$

Many of these indices have first been proposed with $\rho = 2$, which corresponds to the case of the standard Gini index. $IT(\rho = 2)$ is known as the Kakwani index of TR progressivity, $IV(\rho = 2)$ is known as the Reynolds-Smolensky index of IR progressivity and vertical equity, and $IH(\rho = 2)$ is known as the Atkinson-Plotnick index of reranking. The general S-Gini formulations have been proposed by Kakwani and Pfähler (TR and IR progressivity) and Duclos (reranking).

8 Issues in the empirical measurement of well-being and poverty

Poverty assessment is customarily carried out using data on households and individuals. These data can be administrative (*i.e.*, stored in government files and records), they can come from censuses of the entire population, or (most commonly) they can be generated by probabilistic surveys on the characteristics and living conditions of a population of households.

8.1 Survey issues

There are several aspects of the surveying process that are important for poverty assessment. First, there is the coverage of the survey: does it contain representative information on the entire population of interest, or just on some socio-economic subgroups? Whether the representativeness of the data is appropriate depends on the focus of the poverty assessment. A survey containing observations drawn exclusively from the cities of a particular country may be perfectly fine if the aim is to design poverty alleviation schemes within these cities; its representativeness will, however, be clearly insufficient if the objective is to assess the allocation of resources between the country's urban and rural areas.

Then, there is the sample frame of the survey. Surveys are usually multi-staged, and built upon strata and clusters. Stratification ensures that a certain minimum amount of information is obtained from each of a given number of areas within a population of interest. Population strata are often geographic and can represent, for instance, the different regions or provinces of a country. Clustering facilitates the interviewing process by concentrating sample observations within particular population subgroups or geographic locations. Strata are thus often divided into a number of different levels of clusters, representing, say, cities, villages, neighborhoods, or households. A complete listing of the clusters in each strata is used to select randomly within each strata a given number of clusters. The selected clusters can then be subjected to further stratification or clustering, and the process continues until the last sampling units (usually households or individuals) have been selected and interviewed.

Fundamental in the use of survey data is the role of the randomness of the information that is generated. Because households and individuals are not all systematically interviewed (unlike in the case of censuses), the information generated from survey data will depend on the particular selection of households and indi-

viduals made from a population. A poverty assessment of a given population will then vary across the various samples that can be selected from this same population. For that reason, poverty assessments carried out using survey data will be subject to so-called "sampling errors", that is, to sampling variability. When generating poverty assessments from sample survey data, it is therefore important to recognise and assess the statistical imprecision of the sampling results obtained.

By ensuring that a minimum amount of information (typically, a certain number of observations) is obtained from each of a number of strata, stratification decreases the extent of sampling errors. A similar effect is obtained by increasing the total size of the sample: the greater the number of households surveyed, the greater on average is the precision of the estimates obtained. Conversely, by bundling observations around common geographic or socio-economic indicators, clustering tends to reduce the informative content of the observations made and thus to increase the size of the sampling errors (for a given number of observations). The sampling structure of a survey also impacts on its ability to provide accurate information on certain population subgroups. For instance, if the clusters within a strata represent regions, and between-region variability is large, it would not be reasonable to try to use the information generated by the selected regions to depict poverty in the other, non-selected, regions.

Survey data are also fraught with measurement and other "non-sampling" errors. For instance, even though they may have been selected for appearance in a sample, some households will not be interviewed, either because they cannot be reached or because they refuse to be interviewed. Such "non-response" will raise difficulties for poverty assessments if it is correlated with observable and non-observable household characteristics. Even if interviewed, households will sometimes consistently misreport their characteristics and living conditions, either because of ignorance, mischief or self-interest. This tends to make poverty assessments built from survey data diverge systematically from the true (and unobserved) population poverty assessment that would be carried out if there were no non-sampling errors. Clearly, such a shortcoming can bias the understanding of poverty and the consequent design of public policy.

The empirical analysis of vulnerability and poverty dynamics is particularly "data demanding". In general, it requires longitudinal (or panel) surveys, which follow each other in time and which interview the same final observational units. Because they link the same units across time, longitudinal data contain in some sense more information than transversal (or cross-sectional) surveys, and they are particularly useful for measuring vulnerability and for understanding poverty dynamics – in addition to facilitating the assessment of the temporal effects of public

policy on well-being. It must be stressed, however, that measurement error problems render the analysis of vulnerability and mobility very difficult, and its results must be interpreted with caution.

8.2 Income versus consumption

It is frequently argued that consumption is better suited than income as an indicator of living standard, at least in many developing countries. One reason is that consumption is believed to vary more smoothly than income, both within any given year and across the life cycle. Income is notoriously subject to seasonal variability, particularly in developing countries, whereas consumption tends to be less variable. Life-cycle theories also predict that individuals will try to smooth their consumption across their low- and high-income years (in order to equalise their "marginal utility of consumption" across time), through appropriate borrowing and saving. In practice, however, consumption smoothing is far from perfect, in part due to imperfect access to commodity and credit markets and to difficulties in estimating precisely one's "permanent" or life-cycle income.

For the non-welfarist interested in outcomes and functionings, consumption is also preferred over income because it is deemed to be a more "direct" indicator of achievements and fulfilment of basic needs. A caveat is, however, that consumption is also an outcome of individual free choice, an outcome which may differ across individuals of the same income and ability to consume, just like actual functionings vary across people of the same capability sets. At a given capability to spend, some individuals may choose to consume less (or little), preferring instead to give to charity, to vow poverty, or to save in order to give important bequests to their children.

Consumption is also held to be more easily observable and measurable than income in developing countries (although this is not always the case). This is not to say that consumption is easy to measure accurately. For one thing, consumption does not equal expenditures. Unlike expenditures, consumption includes the value of own-produced goods. The value of these goods is not easily assessed, since it has not been transacted in a market. Distinguishing consumption expenditures from investment expenditures is very difficult, but failure to do so properly can lead to double-counting in the consumption measure. For instance, a \$1 expenditure on education or machinery should not be counted as current consumption if the returns and the utility of such expenditure will only accrue later in the form of higher future earnings.

Similarly, the value of the services provided by those durable goods owned by

individuals ought also to enter into a complete consumption indicator, but the cost of these durable goods should not enter entirely into the consumption aggregate of the time at which the good was purchased. An important example of this is owner-occupied housing. Again, estimating the service value of durable goods is not easily done. Further difficulties arise from the assessment of the value of various non-market goods and services – such as those provided freely by the government – and the value of intangible benefits such as the quality of the environment, the extent of security and peace, and so on.

8.3 Price variability

Whether it is income or consumption expenditures that are measured and compared, an important issue is how to account for the variability of prices across space and time. Conceptually, this also includes variability in quality and in quantity constraints. Failure to account for such variability can distort comparisons of well-being across time and space. In Ecuador, for instance (see Hentschel and Lanjouw (1996)), and in many other countries, some households have free access to water, and tend to consume relatively large quantities of it with zero water expenditure. Others (often peri-urban dwellers) need to purchase water from private vendors and consequently consume a lower quantity of it at necessarily higher total expenditures. Ranking of households according to water expenditures could wrongly suggest that those who need to buy water are richer and derive greater utility from water consumption (since they spend more on it).

Microeconomic theory suggests that we may wish to account for price variability by comparing real as opposed to nominal consumption expenditures (or income). This can first be done by estimating the parameters of the indirect utility function of the economy's consumers. These parameters identify the ordinal preferences of the consumer. Let these parameters be denoted by ϑ and the indirect utility function be defined by $V(y, q, \vartheta)$, where q is the price vector. Suppose that reference prices are given by q^R . Equivalent consumption expenditure is then given implicitly by y^R :

$$V(y^R, q^R, \vartheta) = V(y, q, \vartheta). \quad (107)$$

Inversion of the indirect utility function yields an equivalent expenditure function e , which indicates how much expenditure at reference prices is needed to be equivalent to (or to generate the same utility as) the expenditure observed at current prices q :

$$y^R = e(q^R, \vartheta, V(y, q, \vartheta)). \quad (108)$$

Poverty analysis then proceeds by comparing real income to a poverty line defined in terms of the reference prices q^R .

An alternative procedure deflates by a cost-of-living index the level of nominal consumption expenditures. One way of defining such a cost-of-living index is to ask what expenditure is needed just to attain a poverty level of utility v_z at prices q . This is given by $e(q, \vartheta, v_z)$. A similar computation is carried out for the expenditure needed to attain v_z at prices q^R : this is $e(q^R, \vartheta, v_z)$. The ratio

$$e(q, \vartheta, v_z) / e(q^R, \vartheta, v_z) \quad (109)$$

is then a cost-of-living index. Dividing y by (109) yields real consumption expenditure, which can then be compared to the expenditure poverty line defined in reference to q^R .

In practice, cost-of-living indices are often taken to be those consumer price indices routinely computed by national statistical agencies. These consumer price indices vary usually across regions and time, but not across levels of income (*e.g.*, across the poor and the non-poor). In some circumstances (*i.e.*, for homothetic utility functions and when consumer preferences are identical), all of the above procedures are equivalent. In general, however, they are not the same.

The fact that utility functions are not generally homothetic, and that preferences are highly heterogeneous, has important implications for poverty measurement and public policy. First, the true cost-of-living index would normally be different across the poor and the rich. Using the same price index for the two groups may distort comparisons of well-being. An example is the effect of an increase in the price of food on economic well-being. Since the share of food in total consumption is habitually higher for the poor than for the rich, this increase should hurt disproportionately more the more deprived. Deflating nominal consumption by the same index for the entire population will, however, suggest that the impact of the food price increase is shared proportionately by all.

Spatial disaggregation is also important if consumption preferences and price changes vary systematically across regions. In few developing countries, however, are consumer price indices available or sufficiently disaggregated spatially. The alternative is then to produce different poverty lines for different regions (based on the same or different consumption baskets, but using different prices) or construct

food price indices. In both cases, the analyst would usually be using regional price information derived from LSMS-style survey data. The resulting indices would then be interpreted as cost-of-living indices, and would help correct for spatial price variation and regional heterogeneity in preferences.

To see why these adjustments are necessarily in part arbitrary, and to see why they can matter in practice, consider the case of Figure 16. It shows 3 indifference curves U_1 , U_2 and U_3 , for three consumers, 1, 2 and 3. Two of these consumers have relatively strong preferences for meat as opposed to fish, and the third (represented by U_3) has strong preferences for fish. Also shown are two budget constraints, one using relative prices q^c (c for *coastal* area), where the price of fish is relatively low, and the other with q^m (m for *mountainous* area), where the price of fish is high compared to the price of meat.

How is the standard of living for individuals 1,2 or 3 to be compared? One way to answer this question is to "cost" the consumption of the three individuals. For this, we may use either q^c or q^m . If we use the mountains' relative price, then the consumption bundles chosen by individuals 1 and 3 are equivalent in terms of value: they lie on the same budget constraint of value B in terms of meat (the numéraire). Individual 2 is clearly then the worst off of all three. If instead we use the coastal area's relative price, then the consumption bundles chosen by individuals 2 and 3 are equivalent, with a common value of A in terms of meat – and individual 1 is the best off.

Hence, choosing reference prices to assess and compare living standards can matter significantly. If we knew *a priori* that individuals 1 and 3 had equivalent living standards, then reference prices q^m would be the right one (conversely: q^c would be the correct reference prices if 2 and 3 could be assumed equally well off). But such information is generally not available. In some circumstances, such as in comparing 1 and 2, we can be fairly certain that one individual is better off than another, whatever the choice of reference prices, but even the extent of the quantitative difference in well-being will generally depend on the choice of reference prices.

The choice of reference prices and reference preferences will also matter for estimating the impact of price changes on well-being and poverty. Consider again Figure 16. Suppose that we wish to measure the impact on consumers' well-being of an increase in the price of fish. Assume for simplicity that this change in relative prices is captured by a move from q^m to q^c . If we were to choose as a reference bundle the bundle of meat and fish chosen by individuals 1 and 2 to capture the impact of this change, then the price impact would be estimated to be fairly low. For individual 1, for instance, this price increase would generate a

fall in real consumption (again with meat as the numéraire) from B to D . Using instead the preferences of individual 3 as reference tastes, real consumption would be estimated to fall from B to A , a much larger fall.

Furthermore, even if 3 had been deemed better off than 1 before the increase in the price of fish, it could well be that 3's strong preferences for fish would make him less well off than 1 after the price change. Hence, when consumer preferences are heterogeneous, price changes can reverse rankings in terms of well-being and poverty. Indeed, in Figure 16, the increase in the price of fish is visibly much more costly for fish eaters than for meat ones. And, if 3 had been deemed poorer than 1 initially, using 1's preferences to capture the change of 3's well-being would not be appropriate since in this case the preferences of the richer are significantly different from those of the poorer. This warns against the use of a common price index across all regions, and as well as across all socio-economic groups – rich and poor.

Exercises

Suppose the following direct utility function over the two goods x_1 and x_2 ,

$$U(x_1, x_2) = x_1^\nu x_2^{1-\nu},$$

with $\nu = 1/3$, and let prices q_1 and q_2 be set to 1.

1. For a function $\varrho(q)$ independent of y , prove that the indirect utility function $V(y, q)$ can be expressed as follows:

$$V(y, q) = \varrho(q) \cdot y.$$

2. If reference prices are $q_1^R = 1$ and $q_2^R = 1$, show that equivalent consumption is given by:

$$y^R = \frac{y}{\Upsilon(q)}$$

for some function $\Upsilon(q)$ independent of y .

3. What is the expenditure needed to attain a poverty level of utility of $v_z = 158.74$ at the reference prices? (Call this z_R .)
4. What are the quantities of goods 1 and 2 that are consumed at the poverty level of utility?

5. Suppose that the price of good 2 is increased from 1 to 3. What is the new cost of the poverty level of utility? (Call this z .)
6. Using definitions (107) and (108), prove the following :

$$y^R/z_R = y/z.$$

What does it imply?

7. Suppose now that a poverty analyst does not believe that consumption of goods 1 and 2 will adjust following good 2's price increase. What is the poverty line z that he would then obtain? (Hint: compute the cost of the initial commodity basket using the new prices.)
8. Using indifference curves and budget constraints, show the difference that taking account of changes in behavior can make for the computation of price indices and the assessment of poverty.

8.4 Household heterogeneity

8.4.1 Equivalence scales

A fundamental problem arises when comparing the well-being of individuals who live in households of differing sizes and composition. Differences in household size and composition can indeed be expected to create differences in household "needs". It is essential to take these needs into account when comparing the well-being of individuals living in differing households. This is typically done using equivalence scales. With these scales, the needs of a household of a particular size and composition are said to be comparable to a household of a particular number of "reference" or "equivalent" adults.

Estimation difficulties

Strategies for the estimation of equivalence scales are all contingent on the choice of comparable indicators of well-being. All such indicators are, however, intrinsically arbitrary. A popular example is food share in total consumption: at equal household food shares, individuals of various household types are deemed to be equally well-off. But, at equal well-being, one household type can well choose a food share that differs from that of the other household types. This would be the case, for instance, for households of smaller sizes for which it would make "sense" to spend more on food than on those goods for which economies of scale are arguably larger, such as housing.

Another difficulty arises when household size and composition are the result of a deliberate free choice. It may be argued, for instance, that a couple which elects freely to have a child cannot perceive this increase in household size to be utility-decreasing. This would be so even if the household's total consumption remained unchanged after the birth of the child (or even if it fell), despite the fact that most poverty analysts would judge this birth to increase household "needs". Another difficulty lies in the fact that the intra-household decision-making process can influence adversely the allocation of resources across household members, and thereby lead to wrong inferences of comparative needs. This is the case, for instance, when more is spent on boys than girls, not because of differential needs, but because of differential gender preferences on the part of the household decision-maker. Such observations may lead analysts to overestimate the real needs of boys relative to those of girls. Using these observed preferences for boys to estimate equivalence scales would then underestimate on average the level of deprivation experienced by girls and their households, since it would be wrongly assumed that girls are less "needy". An analogous analytical difficulty arises when the household decision-maker is a man, and the consumption of his spouse is observed to be smaller than his own. Is this due to gender-biased household decision-making, or to gender-differentiated needs?

To illustrate these issues, consider Figure 17, which graphs consumption of a reference good $x_r(y, q)$ against *household* income y . The predicted consumption of the reference good is plotted for two households, the first composed of only one man, and the other made of a couple (*i.e.*, a man and a woman). A common feature of the literature on the estimation of equivalence scales is the estimation of the total household income at which a reference consumption of a reference good is equal. The basic argument is that when the consumption of that reference good is the same across households, the well-being of household members should also be the same across households. Reference goods are often goods consumed exclusively by some members of the household, such as male or female clothing, perfume for women, a night out at the cinema for adults, *etc.*.

For Figure 17, take for instance the case of men's clothing for $x_r(y, q)$. Suppose that the reference level of that good is given by x_0 . Leaving aside issues of consumption heterogeneity within homogeneous households at a given income level, one would estimate that the one-member household would need an income y_c in order to consume x_0 (at point c), and that the two-member household would require total household income y_d to reach that same reference consumption. Hence, following this line of argument, the second household would need y_d/y_c as much income as the first one to be "as well off" in terms of consumption

of men's clothing. Said differently, the second household's needs would be y_d/y_c that of the single man household. The number of "equivalent adults" in the second household would then be said to be y_d/y_c . This procedure applied to different household types provides a full equivalence scale, which expresses the needs and the composition of households as a function of those of a reference (generally a one-adult) household.

This procedure faces many problems, most of which are very difficult to resolve. First, there is the choice of the reference level of $x_r(y, q)$. If a reference level of x_1 instead of x_0 were chosen in Figure 17, the number of adult equivalents in the second household would fall from y_d/y_c to y_f/y_e . There is little that can be done in general to determine which of these two scales is the right one. In such cases, one cannot use a welfare-independent equivalence scale – the equivalence scale ratios depend on the levels of the households' well-being.

Equivalence scale estimates also generally depend on the choice of the reference good that is used to compare well-being across heterogeneous households. For instance, the choice of adult clothing *versus* that of tobacco, alcohol or other commodities consumed strictly by adults will generally matter in trying to compare the needs of households with and without children. This is in part because preferences for these goods are not independent of – and do not depend in the same manner on – household composition. One additional problem is the issue of the price dependence of equivalence scale estimates. Choosing a different q in Figure 17, for instance, would generally lead to the estimation of different equivalence scale ratios.

Sensitivity analysis

In view of these difficulties, recent work has emphasized that the choice of a particular scale inevitably introduces important value judgements on how needs of individuals differing in non-income characteristics are assessed, and that it might therefore be appropriate to recognize the lack of agreement in this choice when measuring and comparing inequality and poverty levels.

Allowing the assessment of needs to vary turns out to be especially relevant in cross-country comparative analyses, particularly when countries compared differ significantly in their socio-economic fabric. There is in this case the added issue that not only can the appropriate scale rates be uncertain in a given country, but they may also be different between countries. Testing the sensitivity of inequality and poverty results to changes in the incorporation of needs is then a matter of crucial importance particularly for those international comparisons whose results can influence redistributive policies, *e.g.*, through the transfer of resources from some countries or regions to others, or in the assessment of transnational or alternative

anti-poverty policies.

To see how to carry out such sensitivity analysis, define an equivalence scale E as an index of household needs. This index will typically depend on the characteristics of the M different household members, such as their sex and age, and on household characteristics, such as location and size. Because E is normalized by the needs of a single adult, it can be interpreted as a number of "equivalent adults", *viz.*, household needs as a proportion of the needs of a single adult. A "parametric" class of equivalence scales is usually defined as a function of one or of a few relevant household characteristics, with parameters indicating how needs are modified as these characteristics change.

A survey of Buhmann *et al.* (1988) reported 34 different scales from 10 countries, which they summarized as

$$E(M) = M^s \quad (110)$$

with s being a single parameter summarizing the sensitivity of E to household size M . This needs elasticity, s , can be expected to vary between 0 and 1. For $s = 0$, no account is taken of household size. For $s = 1$, adult-equivalent income is equal to *per capita* household income. The larger the value of s , the smaller are the economies of scale in the production of well-being implicitly assumed by the equivalence scale, and the greater is the impact of household size upon household needs.

An obvious limitation of such a simple function such as (110) is its dependence solely on household size and not on household composition or other relevant characteristics. Most equivalence scales do indeed distinguish strongly between the presence of adults and that of children, and some like that of McClements (1977) even discriminate finely between children of different ages. An example of a class of equivalence scales that is more flexible than the above was suggested by Cutler and Katz (1992) – it takes separately into account the importance of the M_A adults and the $M - M_C$ children:

$$E(M, M_A) = (M + c(M - M_A))^s \quad (111)$$

where c is a constant reflecting the resource cost of a child relative to that of an adult, and s is now an indicator of the degree of overall economies of scale within the household. When $c = 1$, children count as adults (which is the assumption made in (110)); otherwise, adults and children have different needs.

8.4.2 Household decision-making and within-household inequality

Finally, and as elsewhere in distributive analysis, there is also the practically insoluble difficulty of having to make interpersonal comparisons of well-being across individuals – compounded by the fact that individuals here are heterogeneous in their household composition. On the basis of which observable variable can we really make interpersonal comparisons of well-being? Again, note that the assumption that well-being for the man is the same as well-being for the couple when $x_r(y, q)$ is equalized in Figure 17 is a very strong one. Furthermore, apart from influencing preferences and commodity consumption, household formation is as indicated above itself a matter of choice and is presumably the source of utility in its own right. Preferences for household composition are themselves heterogeneous, and so is the utility derived from a certain household status. All of this makes comparisons of well-being across heterogeneous individuals and the use of equivalence scales subject to arbitrariness and significant measurement errors.

An additional problem in measuring individual living standards using survey data comes from the presence of intrahousehold inequality. The final unit of observation in surveys is customarily the household. Little information is typically generated on the intrahousehold allocation of well-being (*e.g.*, of the individual benefits stemming from total household consumption). Because of this, the usual procedure is to assume that the adult-equivalent consumption (once computed) is enjoyed identically by all household members.

This, however, is at best an approximation of the true distribution of economic well-being in a household. If the nature of intrahousehold decision-making leads to important disparities in well-being across individuals, assuming equal sharing will significantly underestimate inequality and aggregate poverty. Not being able to account for intra-household inequities will also have important implications for profiling the poor, and also for the design of public policy. For instance, a poverty assessment that correctly showed the deprivation effects of unequal sharing within households could indicate that it would be relatively inefficient to target support at the level of the entire household – without taking into account how the targeted resources would subsequently be allocated within the household. Instead, it might be better to design public policy such as to self-select the least privileged individuals within the households, in the form of specific in-kind transfers or specially designed incentive schemes.

A final and related difficulty concerns who we are counting in aggregating poverty: is it individuals or households? Although this distinction is fundamental,

it is often surprisingly hidden in applied poverty profile and poverty measurement papers. The distinction matters since there is habitually a strong positive correlation between household size and a household's poverty status. Said differently, household poverty is usually found disproportionately among the larger households. Because of this, counting households instead of individuals will typically underestimate significantly the true proportion of individuals in poverty.

Part III

**Ethical robustness of poverty and
equity comparisons**

9 Poverty dominance

The main reason for carrying out analyses of poverty dominance is that comparisons of poverty across time, regions, socio-demographic groups or fiscal regimes (for instance) may be sensitive to the choice of the poverty line and to the choice of the poverty index. This is problematic since a different choice of poverty index or poverty line could reverse an earlier conclusion that poverty is greater in a region A than in a region B , or that poverty will decrease following the introduction of a particular fiscal policy or macroeconomic adjustment programme. Such sensitivity must be checked for one to have some confidence that a poverty ordering is robust to the choice of a poverty line or a poverty index. Another reason is that unknown errors in measuring well-being will necessarily affect cardinal poverty estimates; under some assumptions (admittedly restrictive), such errors will not contaminate ordinal poverty comparisons.

To see this better, consider the hypothetical example of Table ???. The second, third and fourth lines in the table show the standards of livings of three individuals in two hypothetical distributions, A and B . Thus, distribution A contains three standards of living of 4, 11 and 20 respectively. The bottom 3 lines of the table show the value of the two most popular indices of poverty, the headcount and the average poverty gap indices, at two alternative poverty lines, 5 and 10. Recall from section 5.2.2 that the poverty headcount gives the proportion of individuals in a population whose standard of living falls underneath a poverty line. At a poverty line of 5, there is only one such person in poverty in distribution A , and the headcount is thus equal to $1/3=0.33$. The average poverty gap index is the sum of the distance of the poor's standards of living from the poverty line, divided by the number of people in the population. For instance, at a poverty line of 10, there are 2 people in poverty in B , and the sum of their distance from the poverty line is $(10-6)+(10-9)=5$. Divided by 3, this gives 1.66 as the average poverty gap in B for a poverty line of 10.

The last column of Table ??? gives the poverty ranking of the two distributions according to the different choices of poverty lines and poverty indices. At a poverty line of 5, the headcount in A is clearly greater than in B , but this ranking is spectacularly reversed if we consider instead the same headcount but at a poverty line of 10. The ranking changes again if we use the same poverty line of 10 but now focus on the average poverty gap $\mu^g(z) : \mu_A^g(10) = 2 < 1.66 = \mu_B^g(z)$. Clearly, ranking A and B in terms of poverty can be quite sensitive to the precise choice of measurement assumptions.

Ordinal comparisons, on the other hand, do not attempt to put a precise nu-

merical value on the extent of poverty. They only try to rank poverty across two distributions, indicating whether it is higher or lower in the first than in the second. Ordinal comparisons of poverty do not, therefore, provide precise numerical data to compare with metric indicators of other aspects or effects of government policy, such as its administrative or efficiency cost. This is their main defect. They can, however, be highly robust to the choice of measurement assumptions, since they will sometimes be valid for wide ranges of such assumptions. When the problem is simply of resolving which of two policies will better alleviate poverty, or determining which of two distributions has the most poverty, ordinal comparisons can be sufficiently informative, that is, cardinal estimates will not be needed. In that case, ordinal comparisons will also be sufficiently convincing. For instance, we will see later in Section 9.1 that we can order robustly distributions A and B in Table ?? for all "distribution-sensitive" poverty indices and for any choice of poverty line.

A focus on ordinal comparisons has two major advantages. First, it saves most of the considerable energy and time often spent on choosing poverty lines and poverty indices. This includes avoiding the difficult debate on the choice of appropriate theoretical and econometric methods for estimating poverty lines. It also enables the poverty analyst to escape arguing on the relative merits and properties of the many poverty indices that have been proposed in the scientific literature. This is because ordinal poverty comparisons do not require that the precision of numerical poverty estimates be validated; it is simply their ordinal ranking across policies or distributions of well-being that is important, and for this, the poverty estimates do not require to be precisely known.

In essence, testing poverty dominance allows one to secure poverty comparisons that necessarily hold for groups (or classes) of poverty indices, as well as for ranges of poverty lines. These classes are defined for specific orders s of stochastic dominance. The first-order class of poverty indices regroups all poverty indices that weakly decrease when the living standard of someone in the population increases. By "weakly decrease", we mean that the poverty index will never increase following a rise in someone's living standard, and will sometimes decrease if the person involved initially had a living standard below the poverty line. These poverty indices have properties that are analogous to those of Paretian social welfare functions: *ceteris paribus*, the larger the levels of the individual living standards, the better off is society (the lower is poverty).

The second-order class of poverty indices contains those indices (among the first-order class of indices) that have a greater ethical preference for the poorer among the poor. Mathematically, these indices are convex in income: all other

things the same, the more equal the distribution of income among the poor, the lower the level of poverty. The indices thus display a preference for equality of income. If a transfer from a poor to a poorer person takes place without reversing the ranks of the two individuals, the indices in the class of second-order indices will never increase, and will sometimes fall. This equality-preferring property is analogous to the Pigou-Dalton principle of transfer for social welfare functions (social welfare increases when an equalizing transfer of income takes place). They are therefore "distribution-sensitive". All of the indices that belong to the second-order class of indices also belong to the first-order class.

To understand the third-order class of poverty indices, imagine four levels of living standard, for individuals 1, 2, 3, and 4, such that $y_2 - y_1 = y_4 - y_3 > 0$ and $y_1 < y_3$. Let a marginal transfer of \$1 of income be made from individual 2 to individual 1 (an equalizing transfer) at the same time as an identical \$1 is transferred from individual 3 to individual 4 (a disequalizing transfer). This is called in the literature a "favorable composite transfer". Note that the equalizing transfer is made lower down in the distribution of income than the disequalizing transfer. This can be seen by the fact the recipient of the first transfer, 1, has a lower standard of living than the donor of the second transfer, 3, since $y_3 > y_1$.

There are often sound ethical reasons to be socially more sensitive to what happens towards the bottom of the distribution of income than higher up in it. We may thus be less concerned about the "bad" disequalizing transfer higher up in the distribution of income than we are pleased about the "good" equalizing transfer lower down. Second-order poverty indices which exhibit this property by decreasing when a favorable composite transfer is effected are said to belong also to the third-order class of poverty indices, and to obey the "transfer-sensitivity" principle.

We can, if we wish, define subsequent classes of poverty indices in an analogous manner. As the order s of the class of poverty indices increases, the indices become more and more sensitive to the distribution of income among the poorest. At the limit, as s becomes very large, only the living standard of the poorest individual matters in comparing poverty across two distributions.

To see this analytically, we will focus for simplicity on classes of additive poverty indices denoted as $\Pi^s(z)$. The additive poverty indices $P(z)$ that are members of that class can be expressed as

$$P(z) = \int_0^1 \pi(Q(p); z) dp \quad (112)$$

where z is a poverty line. For expositional simplicity, we assume that $\pi(Q(p); z)$

is continuously differentiable in $Q(p)$ between 0 and z up to the appropriate order. We denote the i^{th} -order derivative of $\pi(Q(p); z)$ with respect to $Q(p)$ as $\pi^{(i)}(Q(p); z)$. We can think of the function $\pi(Q(p); z)$ as the contribution of an individual with living standard $Q(p)$ to overall poverty $P(z)$. Hence, we can also assume that $\pi(Q(p); z) = 0$ if $Q(p) > z$.

As mentioned above, the first class of poverty indices (denoted by $\Pi^1(z)$) re-groups all poverty indices that weakly decrease when the living standard of someone in the population increases. This implies that indices within $\Pi^1(z)$ are such that:

$$\Pi^1(z) = \left\{ P(z) \mid \pi^{(1)}(Q(p); z) \leq 0 \text{ when } Q(p) \leq z \right\}. \quad (113)$$

The second class of poverty indices, $\Pi^2(z)$, contains those indices that have a greater ethical preference for the poorer among the poor, and are therefore convex in income. This class $\Pi^2(z)$ is then:

$$\Pi^2(z) = \left\{ P(z) \mid \begin{array}{l} \pi^{(2)}(Q(p); z) \geq 0 \text{ when } Q(p) \leq z \\ \text{and } \pi(z; z) = 0 \end{array} \right\} \quad (114)$$

where, for expositional simplicity, we have added the continuity condition $\pi(z, z) = 0$ (to be discussed further below). Since $\pi(z; z) = 0$, $\pi(Q(p); z) \geq 0$ and $\pi^{(2)}(Q(p); z) \geq 0$ for $Q(p) \leq z$ jointly imply that $\pi^{(1)}(Q(p); z) \leq 0$ for $Q(p) \leq z$, we therefore have $\Pi^2(z) \subset \Pi^1(z)$. All of the indices belonging to $\Pi^2(z)$ also belong to $\Pi^1(z)$.

Mathematically, obeying the "transfer-sensitivity" principle requires the second derivative $\pi^{(2)}(Q(p); z)$ to be decreasing in $Q(p)$. Poverty indices belonging to the third-order class of poverty indices $\Pi^3(z)$ are then defined as:

$$\Pi^3(z) = \left\{ P(z) \mid \begin{array}{l} \pi^{(3)}(Q(p); z) \leq 0 \text{ when } Q(p) \leq z, \\ \pi(z, z) = 0 \text{ and } \pi^{(1)}(z, z) = 0 \end{array} \right\}. \quad (115)$$

As before, $\Pi^3(z) \subset \Pi^2(z)$.

We can, if we wish, define subsequent classes of poverty indices in an analogous manner. Generally speaking, poverty indices $P(z)$ will be members of class $\Pi^s(z)$ if $(-1)^s \pi^{(s)}(Q(p); z) \leq 0$ and if $\pi^{(i)}(z, z) = 0$ for $i = 0, 1, 2, \dots, s - 2$. As the order s of the class of poverty indices increases, the indices become more and more sensitive to the distribution of income among the poorest. At the limit, only the living standard of the poorest individual matters in comparing poverty across two distributions. Increasing the order s makes us focus on smaller subsets of poverty indices, in the sense that $\Pi^s(z) \subset \Pi^{s-1}(z)$.

A number of well-known poverty indices fit into some of the classes defined above. The headcount, for which $\pi(Q(p); z) = I[Q(p) \leq z]$, belongs only to $\Pi^1(z)$. We will also see below that it plays a crucial role in tests of first-order dominance. The average poverty gap, for which $\pi(Q(p), z) = g(p; z)$, belongs to $\Pi^1(z)$ and to $\Pi^2(z)$. The square of the poverty gaps index belongs to $\Pi^1(z)$, $\Pi^2(z)$ and $\Pi^3(z)$. More generally, the FGT indices, for which $\pi(Q(p); z) = g(p; z)^\alpha$, belong to $\Pi^s(z)$ when $\alpha \geq s - 1$. The Watts index, for which $\pi(p; z) = \ln(z) - \ln(Q^*(p))$, belongs to $\Pi^1(z)$ and $\Pi^2(z)$. A transformation of the Watts index, by which $\pi(Q(p); z) = g(p; z) [\ln(z) - \ln(Q^*(p))]$, would belong to $\Pi^3(z)$. The Chakravarty and Clark *et al.* indices belong to $\Pi^1(z)$ and $\Pi^2(z)$. The S-Gini indices of poverty also belong to $\Pi^1(z)$ and $\Pi^2(z)$.

To check whether poverty in A is greater than in B for all indices that are members of any one of these classes, there exist two approaches: a primal approach and a dual approach. We look at them in turn.

9.1 Primal approach

We are interested in whether we may assert confidently that poverty in a distribution A , as measured by $P_A(z)$, is larger than poverty in a distribution B , $P_B(z)$, for *all* of the poverty indices $P(z)$ belonging to one of the classes of poverty indices defined above and for a range of possible poverty lines. We are therefore interested in checking whether the following difference in poverty indices $\Delta P(z) = P_A(z) - P_B(z)$ is positive:

$$\begin{aligned} \Delta P(z) &= \int_0^1 \pi(Q_A(p); z) - \pi(Q_B(p); z) dp \\ &= \int_0^z \pi(y; z) (f_A(y) - f_B(y)) dy \\ &= \int_0^z \pi(y; z) \Delta f(y) dy \end{aligned} \tag{116}$$

where on the second line a change of variable has been effected and where $\Delta f(y)$ is the difference in the densities of income. To see how to test this, we will make repetitive use of integration by parts of equation (116). This process will make use of the stochastic dominance curves $D^s(z)$, for orders of dominance $s = 1, 2, 3, \dots$. $D^1(z)$ is simply the *cdf*, $F(z)$, namely, the proportion of individuals underneath varying values for the poverty line z . The higher order curves are iteratively defined as

$$D^s(z) = \int_0^z D^{s-1}(y) dy \tag{117}$$

where $c = 1/(s - 1)!$ is a constant that can be safely ignored in the use of dominance curves. Thus, $D^2(z)$ is simply the integral of the area under the *cdf* curve

until $F(z)$. This is illustrated in Figure 25. The curve shows the *cdf* $F(y)$ at different values of y . The grey-shaded area underneath that curve up to z thus gives $D^2(z)$.

Defined as in (117), dominance curves may seem complicated to calculate. Fortunately, there is a very useful link between the dominance curves and the popular FGT indices, a link that greatly facilitates the computation of $D^s(z)$. Indeed, we can show that

$$\begin{aligned} D^s(z) &= c \cdot \int_0^z (z - y)^{s-1} dy \\ &= c \cdot P(z; \alpha = s - 1) \end{aligned} \quad (118)$$

Therefore, to compute the dominance curve of order s , we only need to compute the FGT index at $\alpha = s - 1$, which is $P(z; \alpha = s - 1)$. Recall that $P(z; \alpha = 1)$ is the average poverty gap in the population when a poverty line z is used. Hence, the dominance curve of order 1 is simply the average poverty gap for different poverty lines. This can also be seen on Figure 25. The distance between z and y gives (when it is positive) the poverty gap at a given value of income y . For $y = y'$, for instance, Figure 25 shows that distance $z - y'$. $dF(y')$ – as measured on the vertical axis – gives the density of individuals at that level of income. The rectangular area given by the product of $(z - y')dF(y')$ then shows the contribution of those with income y' to the population average poverty gap. Integrating all such positive distances between y and z across the population thus amounts to calculating the average poverty gap – again, it is the sum of individual rectangles of lengths $(z - y)$ and heights $dF(y)$ or simply the grey-shaded area of Figure 25.

Let us now integrate by parts equation (116). This gives:

$$\Delta P(z) = \pi(z; z) \Delta D^1(z) - \int_0^z \pi^{(1)}(y; z) \Delta D^1(y) dy \quad (119)$$

where $\Delta D^s(y)$ is defined as $D_A^s(y) - D_B^s(y)$. If we wish to ensure that $\Delta P(z)$ will be positive for all of the indices that belong to $\Pi^s(z)$, we will need to ensure that (119) is positive for all poverty indices that satisfy (113), whatever the values of the first-order derivative $\pi^{(1)}(y; z)$, so long as the derivative is everywhere non-positive between 0 and z . For this to hold, we need that (recall that $D^1(y) = F(y)$):

$$F_A(y) > F_B(y), \text{ for all } y \in [0, z]. \quad (120)$$

We refer to this as first-order poverty dominance of B over A . Since the dominance in condition (120) is conditional on the poverty line being between 0 and z , we can also label it as “first-order conditional stochastic dominance”. It is a relatively stringent condition: it requires the headcount index in A to be always larger

than the headcount in B , for all of the poverty lines between 0 and z . If, however, condition (120) is found to hold in practice, a very robust poverty ordering is obtained: we can then unambiguously say that poverty is higher in A than in B for all of the poverty indices with non-positive first-order derivatives (including the headcount index), that is for all members of $\Pi^1(z)$. These indices include all those which are weakly decreasing in income. Since almost all of the poverty indices that have been proposed obey this restriction, this is a very powerful conclusion indeed. Furthermore, this same ordering is robustly valid over all of the classes of poverty indices $\Pi^1(\zeta)$ such that $0 \leq \zeta \leq z$, which adds a very useful element of robustness over the space of poverty lines. This result can be summarized as follows:

First-order poverty dominance:

$$\begin{aligned} P_A(\zeta) - P_B(\zeta) &> 0 \text{ for all } P(\zeta) \in \Pi^1(\zeta) \text{ and for all } \zeta \in [0, z] \\ \text{iff } D_A^1(\zeta) &> D_B^1(\zeta) \text{ for all } \zeta \in [0, z] \end{aligned}$$

We now move on to second-order poverty dominance. For this, we integrate once more by parts equation (119), and find the following:

$$\Delta P(z) = \pi(z; z) \Delta D^1(z) - \pi^{(1)}(z; z) \Delta D^2(z) + \int_0^z \pi^{(2)}(y; z) \Delta D^2(y) dy. \quad (121)$$

Recall that the indices that are members of $\Pi^2(z)$ are such that $\pi^{(2)}(Q(p); z) \geq 0$ when $Q(p) \leq z$ and $\pi(z, z) = 0$. Hence, if we wish $\Delta P(z)$ to be positive for all of the indices that belong to $\Pi^2(z)$, we must have:

$$\Delta D^2(y) > 0 \text{ for all } y \in [0, z]. \quad (122)$$

This is second-order poverty dominance of B over A . It is a less stringent condition than first-order poverty dominance, since by (117), when first-order dominance over $[0, z]$ holds, then second-order dominance over $[0, z]$ must also hold, but not necessarily the converse. Second-order poverty dominance requires that the average poverty gap in A be always larger than the average poverty gap in B , for all of the poverty lines between 0 and z . If condition (122) is found to hold in practice, then a rather robust poverty ordering is obtained: we can unambiguously say that poverty is higher in A than in B for all of the poverty indices that are continuous at the poverty line and with positive second-order derivatives (also including the average poverty gap itself). Most of the indices found in the literature fall into that category, a major exception being the headcount (it is discontinuous

at the poverty line) and the Sen indices. The same ordering in terms of indices that are members of $\Pi^2(z)$ is also robustly valid for any choice of a poverty line between 0 and z . This can be summarized as follows:

Second-order poverty dominance (primal):

$$\begin{aligned} P_A(\zeta) - P_B(\zeta) &> 0 \text{ for all } P(\zeta) \in \Pi^2(z) \text{ and for all } \zeta \in [0, z] \\ \text{iff } D_A^2(\zeta) &> D_B^2(\zeta) \text{ for all } \zeta \in [0, z] \end{aligned}$$

In fact, a comparison of distributions A and B in Table ?? shows that this condition is obeyed for any choice of z . Hence, saying that A has more poverty than B in that table is quite a robust statement, since it is valid for all distribution-sensitive poverty indices (the headcount is not distribution-sensitive, hence it does not always indicate more poverty in A) and for any choice of poverty line. As mentioned, second-order poverty dominance is a less stringent criterion than first-order dominance to check in practice. The price of this, however, is that the set of indices over which poverty dominance is checked is smaller for second-order dominance than for first-order dominance.

We can repeat this process for any arbitrary order of dominance, by successive integration by parts and by determining the conditions under which all of the poverty indices that are members of a class $\Pi^s(\zeta)$ will indicate more poverty in A than in B , and this for all of the poverty lines ζ between 0 and z . This gives the general formulation of s^{mboxth} order poverty dominance:

s^{th} -order poverty dominance:

$$\begin{aligned} P_A(\zeta) - P_B(\zeta) &> 0 \text{ for all } P(\zeta) \in \Pi^s(z) \text{ and for all } \zeta \in [0, z] \\ \text{iff } D_A^s(\zeta) &> D_B^s(\zeta) \text{ for all } \zeta \in [0, z] \end{aligned}$$

This is illustrated in Figure 12 for general s -order dominance, where dominance holds until z_{max} , but would not hold if z_{max} exceeded z_s . Checking poverty dominance is thus conceptually straightforward. For first-order dominance, we use what has been termed “the poverty incidence curve”, which is the headcount index as a function of the range of poverty lines $[0, z]$. For second-order dominance, we use the “poverty deficit curve”, which is the area underneath the poverty incidence curve or more simply the average poverty gap, again as a function of the range of poverty lines $[0, z]$. Third-order dominance makes use of the area underneath the poverty deficit curve, or the square of poverty gaps index (also called the poverty severity curve) for poverty lines between 0 and z . Dominance curves for greater orders of dominance simply aggregate greater powers of poverty gaps, graphed against the same range of poverty lines $[0, z]$.

9.2 Dual approach

There exists a dual approach to testing first-order and second-order poverty dominance, which is sometimes called a p , percentile or quantile approach. Whereas the primal approach makes use of curves that censor the population's income at varying poverty lines, the dual approach makes use of curves that truncate the population at varying percentiles. The dual approach has interesting graphical properties, which makes it useful and informative in checking poverty dominance.

To illustrate this second approach, we focus in this section on indices that aggregate poverty gaps using weights that are functions of p :

$$\Gamma(z) = \int_0^1 g(p; z) \omega(p) dp. \quad (123)$$

Using aggregates of poverty gaps as in (123) is more restrictive than using functions $\pi(Q(p); z)$ defined separately over $Q(p)$ and z , as in (112). When the poverty lines are identical across distributions (as was implicitly assumed above for the primal approach), the dominance rankings are, however, the same, as we will see below.

Membership in the dual first-order class $\dot{\Pi}^1(z)$ of poverty indices only requires that the weights $\omega(p)$ be non-negative functions of p :

$$\dot{\Pi}^1(z) = \{\Gamma(z) \mid \omega(p) \geq 0\}. \quad (124)$$

If we want that $\Delta\Gamma(z) = \Gamma_A(z) - \Gamma_B(z)$ be positive for all of the indices that belong to $\dot{\Pi}^1(z)$, we need to check that $g_A(p; z) - g_B(p; z)$ is positive whenever $g_B(p; z) > 0$. In fact, if this condition is verified, $\Delta\Gamma(\zeta)$ will be positive for all $\zeta \in [0, z]$. This yields the dual first-order poverty dominance condition:

First-order poverty dominance (dual approach):

$$\begin{aligned} & \Gamma_A(\zeta) - \Gamma_B(\zeta) > 0 \text{ for all } \Gamma(\zeta) \in \dot{\Pi}^1(\zeta) \text{ and for all } \zeta \in [0, z] \\ & \text{iff } g_A(p; z) > g_B(p; z) \text{ whenever } g_B(p; z) > 0 \end{aligned}$$

This requires poverty gaps to be nowhere lower in A than in B , whatever the percentiles p considered. We can show that this is also equivalent to the primal first-order poverty dominance condition. Technically, if and only if $\Gamma_A(\zeta) - \Gamma_B(\zeta) > 0$ for all $\Gamma(\zeta) \in \dot{\Pi}^1(\zeta)$, then $P_A(\zeta) - P_B(\zeta) > 0$ for all $P(\zeta) \in \Pi^1(\zeta)$, and for all $\zeta \in [0, z]$. If we have robustness over $\Pi^1(z)$, we also have robustness over $\dot{\Pi}^1(z)$. In fact, first-order poverty dominance (primal or dual) implies robustness over all poverty indices (additive or otherwise) that are weakly decreasing in

income. To check such wide degree of robustness, we can equivalently use the primal or the dual first-order poverty dominance condition.

Membership in the dual second-order class $\dot{\Pi}^2(z)$ of poverty indices requires that the weights $\omega(p)$ be positive and non-increasing functions of p :

$$\dot{\Pi}^2(z) = \left\{ \Gamma(z) \mid \omega^1(p) \leq 0 \text{ and } \omega(p=1) \geq 0 \right\}. \quad (125)$$

By integration by parts of (125), we can show that if we want $\Delta\Gamma(\zeta)$ to be positive for all of the indices that belong to $\dot{\Pi}^2(\zeta)$ and for all $\zeta \in [0, z]$, we need to ensure the following condition:

Second-order poverty dominance (dual approach):

$$\begin{aligned} & \Gamma_A(\zeta) - \Gamma_B(\zeta) > 0 \text{ for all } \Gamma(\zeta) \in \dot{\Pi}^2(\zeta) \text{ and for all } \zeta \in [0, z] \\ & \text{iff } G_A(p; z) > G_B(p; z) \text{ for all } p \in [0, 1] \end{aligned}$$

Again, we can show that this is equivalent to the primal second-order poverty dominance condition. Hence, if and only if $\Gamma_A(\zeta) - \Gamma_B(\zeta) > 0$ for all $\Gamma(\zeta) \in \dot{\Pi}^2(\zeta)$, for all $\zeta \in [0, z]$, then $P_A(\zeta) - P_B(\zeta) > 0$ for all $P(\zeta) \in \Pi^2(\zeta)$, for all $\zeta \in [0, z]$. If we have robustness over $\Pi^2(z)$, we also have robustness over $\dot{\Pi}^2(z)$. Again, second-order poverty dominance (primal or dual) implies robustness over all poverty indices that are continuous at z and that are weakly convex in income. To check robustness over all such poverty indices and for all poverty lines between 0 and z , we can use either the primal or the dual second-order poverty dominance condition.

9.3 Assessing the limits to dominance

We saw above that it is often necessary to specify lower and upper bounds for verifying dominance conditions over ranges of ζ and p . These bounds can be obtained from previous empirical or ethical work on plausible ranges of poverty lines and/or percentiles over which to compare poverty and distributions of income. Alternatively, they can be specified *a priori* and arbitrarily by the researcher. An often better research strategy is, however, to use the available sample information and estimate directly from it the lower and upper bounds between which a distributive comparison can be inferred to be robust. We can then interpret these lower and upper bounds as “critical” bounds. Critical bounds for poverty lines will define the range of poverty lines which must not be exceeded for a robust ordering of poverty between A and B to be possible. A similar interpretation is valid for

“critical” p values in the dual approach. We will come back later to the sampling distribution of these estimated critical bounds.

For now, we will assume that these bounds exist in the two populations of income being compared. To describe them, we start with our primal or additive class of unrestricted poverty indices. We assume that B initially dominates A , but that their dominance curves eventually cross and that their ranking is reversed. Hence, for a positive η , assume that $D_A^s(\zeta) > D_B^s(\zeta)$ for all $\zeta \in [0, \eta]$. Let $\zeta^+(s)$ then be the first crossing point of the curves, with $D_A^s(\zeta^+(s)) = D_B^s(\zeta^+(s))$. Distribution B then dominates distribution A up to a critical poverty line $\zeta^+(s)$. Hence, instead of specifying a more or less arbitrary value of z in inferring unrestricted poverty dominance, we may simply say that B dominates A at order s for all poverty lines between 0 and $\zeta^+(s)$. We can repeat this process for any desired order of poverty dominance. Indeed, we can show that $\zeta^+(s)$ is increasing in s , with $\zeta^+(s+1) > \zeta^+(s)$. Thus, if we find at order s the robust range of poverty lines to be too limited, we can extend it by moving to a higher order of dominance.

It is also worth adding that it is empirically unlikely that we will be able to infer dominance over low values of ζ and p . This leads to a search for restricted dominance. Hence, there will generally be a need and a gain in estimating also the *lower* critical bound of the range of poverty lines over which robustness can be inferred. Let $[\zeta^-(s), \zeta^+(s)]$ be a range of ζ over which $D_A^s(\zeta) > D_B^s(\zeta)$. This leads to restricted poverty dominance over the range $[\zeta^-(s), \zeta^+(s)]$. Sample estimates of $\zeta^-(s)$ and $\zeta^+(s)$ can be estimated for any desired order of dominance. Unlike the case of unrestricted dominance above, neither $\zeta^-(s)$ nor $\zeta^+(s)$ is necessarily increasing or decreasing in s .

A similar exercise can be done for the dual approach to restricted dominance, assuming an upper bound z^+ for z . We define $\psi^-(s)$ and $\psi^+(s)$ as the population bounds for the intervals of percentiles p over which it is possible to ascertain first- or second-order dual restricted dominance. For first-order dominance, this will give the range of p , $[\psi^-(1), \psi^+(1)]$ over which $g_A(p; z^+) > g_B(p; z^+)$. For second-order dominance, this will give the range of p , $[\psi^-(2), \psi^+(2)]$, over which $G_A(p; z^+) > G_B(p; z^+)$.

10 Inequality dominance

As for poverty and welfare dominance, we can define classes of relative inequality indices over which to check the robustness of the inequality orderings of two distributions of income. As we will see, these classes of inequality indices have properties which are analogous to those of the classes of social welfare indices. They react in a given manner to changes in or reallocations of income. Unlike social welfare functions, however, relative inequality indices also need to be homogeneous of degree 0 in all income. This means that an equi-proportionate change in all income will not affect the value of these relative inequality indices.

We will at first ignore the class of inequality indices of the first-order. Besides being homogeneous of degree 0 in income, indices that belong to the class Υ^2 of second-order inequality indices weakly decrease after a mean-preserving equalizing transfer. Like the class of social welfare indices Ω^2 , they thus obey the Pigou-Dalton principle of transfers⁹. These inequality indices are also said to be Schur-convex. Almost all of the frequently used inequality indices (including the Atkinson, S-Gini and generalized entropy indices, with the notable exception of the variance of logarithms) are members of Υ^2 . Inequality indices that belong to the class Υ^3 of third-order inequality indices belong to Υ^2 , and weakly decrease after a favorable composite transfer. This includes the Atkinson indices and some of the generalized entropy indices, but not the S-Gini indices. Classes of higher order inequality indices can be similarly defined. For instance, to be members of the class of fourth-order inequality indices, inequality indices must be members of Υ^3 and must be more sensitive to favorable composite transfers when they take place lower down in the distribution of income.

Comparing the definitions of the classes Ω^s and Υ^s , we find that when the means of the distributions are equal, the social welfare ranking is the same as the inequality ranking, in the sense that if $I_A > I_B$, then $W_A < W_B$, and vice-versa. Thus, in such cases, checking for inequality dominance can be done by checking for welfare dominance. When the means are not equal, we can simply normalize all incomes by their mean, and then use the welfare dominance results described in section 11 for Ω^s to check for dominance over a class Υ^s of relative inequality indices. Hence, to check for inequality dominance, we can check for welfare dominance once incomes have been normalized by their mean. When B dominates A at order s , we can say that I_B is lower than I_A for all of the inequality

⁹Note, however, that an equalizing transfer increases the value of the social welfare indices that are members of Ω^2 , whereas it decreases the value of the inequality indices that are members of Υ^2 .

indices that belong to Υ^s .

Further, for each of the relative inequality indices considered above, we can find a group of homothetic social welfare indices whose common EDE living standard is defined in such a way that $\xi = \mu \cdot (1 - I)$. In such cases, assessing social welfare can be done in two steps: first, compute the mean income, and second, assess the level of inequality. If a distribution B has a greater mean living standard and a lower level of inequality than a distribution A , it necessarily has a greater level of social welfare. If, besides having a larger mean, distribution B has a robustly lower level of inequality over the class Υ^s of inequality indices, then it also necessarily has a greater level of social welfare over the class Ω^s of social welfare indices that are homothetic.

In what follows, we summarize the inequality dominance results for each of the primal and the dual approaches.

10.1 Primal approach

As indicated above, to check for inequality dominance, we may simply use the welfare dominance curves after a normalization of incomes by their mean. For the primal dominance curves, this gives $\bar{D}^s(l\mu)$:

$$\bar{D}^s(l\mu) = \frac{D^s(l \cdot \mu)}{(l\mu)^{(s-1)}}. \quad (126)$$

The curve $\bar{D}^s(l\mu)$ has a nice equivalence in terms of the normalized FGT indices $\bar{P}(z; \alpha)$:

$$\bar{D}^s(l\mu) = c\bar{P}(z = l\mu; \alpha = s - 1). \quad (127)$$

where c is again a constant that can be safely neglected. Thus, estimating the normalized dominance curve at $l\mu$ and order s is equivalent to computing the normalized FGT index for a poverty line equal to $l\mu$ and for α equal to $s - 1$. The primal second-order inequality dominance condition is then as follows:

Second-order inequality dominance (primal):

$$\begin{aligned} I_A - I_B > 0 \text{ for all } I \in \Upsilon^2 \\ \text{iff } \bar{D}_A^2(\lambda\mu) > \bar{D}_B^2(\lambda\mu) \text{ for all } \lambda \in [0, \infty[\end{aligned} \quad (128)$$

This requires the normalized FGT index to be greater in A than in B for all of the poverty lines set to non-negative proportions of the mean. The general s -order inequality dominance condition is then:

***s*-order inequality dominance (primal):**

$$\begin{aligned} I_A - I_B &> 0 \text{ for all } I \in \Upsilon^s \\ \text{iff } \bar{D}_A^s(\lambda\mu) &> \bar{D}_B^s(\lambda\mu) \text{ for all } \lambda \in [0, \infty[\end{aligned} \quad (129)$$

10.2 Dual approach

The dual approach to (unconditional) inequality dominance is most convenient for second-order dominance, and has been extensively applied in the literature. Dual conditions for third-order dominance have also been proposed in the literature, but they are not as easily checked as the primal conditions.

Second-order inequality dominance (dual):

$$\begin{aligned} I_A - I_B &> 0 \text{ for all } I \in \Upsilon^2 \\ \text{iff } L_A(p) &< L_B(p) \text{ for all } p \in]0, 1[. \end{aligned} \quad (130)$$

10.3 Inequality and progressivity

In the absence of reranking, if a tax and/or a transfer is *TR*- or *IR*-progressive, then all inequality indices that are members of Υ^2 are necessarily reduced by the progressive tax and/or transfer. Further, again in the absence of reranking, if a tax and/or transfer T_1 is more *IR*-progressive than a tax and/or transfer T_2 , T_1 necessarily reduces inequality more than T_2 when inequality is measured by any of the inequality indices that belong to Υ^2 .

This can be seen by the movement of the Lorenz curve, using the concentration curve and in the absence of reranking (see equation (95)). A concentration curve for net income that lies above the Lorenz curve of gross income pushes the Lorenz curve of net income above, which decreases inequality for most (all "second-order") measures of inequality.

An alternative way to see this effect is to display the change in the cumulative share of the poorer individuals – that is, simply the change in the Lorenz curve. This change is positive if the tax and benefit system is progressive. We may also be concerned about the impact of a tax and benefit system on the class of first-order inequality indices, *viz.*, on indices that are monotonic in income shares, but not always in cumulative income shares. To check whether this impact reduces first-order inequality indices, we must check whether $\bar{T}(X)/X$ or $\bar{T}(p)/X(p)$ is always lower than μ_T/μ_X for all X or p that are below some censoring point $l^+\mu$ or p^+ . This supposes, however, that the tax does not induce reranking. When it

does, one way to account for the reranking effect is to compute "income growth" curves, given by $(N(p) - X(p))/X(p)$. When these curves exceed the growth in average income $-(\mu_N - \mu_X)/\mu_X$ for all $p \leq p^+$ or $p \leq F_X(l^+ \mu_X)$, then all first-order inequality indices fall.

11 Welfare dominance

As for the measurement of poverty, we may wish to determine if the ranking of two distributions of income in terms of social welfare is robust to the choice of indices. One way to check such robustness would be to verify the ranking of the two distributions for a large number of the many social welfare indices that have been proposed in the literature. This, however, would be a relatively tedious task. Another way, simpler and generally more powerful, is through tests of stochastic dominance over the whole distributions of income (as opposed to the censored or truncated distributions for the comparisons of poverty). As for the measurement of poverty, there are two approaches, a primal and a dual one. The primal approach has the advantage of being applicable to any desired (however large) order of dominance, and uses curves of the well-known FGT indices for an infinite range of “poverty lines” or censoring points. The second approach is convenient only for first and second order dominance, but uses curves that are graphically informative and appealing and which have been used and documented extensively in the literature. For first and second order dominance, if a robust ranking is obtained using one approach, the same ranking will be obtained using the other approach: in other words, the two approaches are equivalent in terms of their ability to rank distributions robustly over classes of first- and second-order social welfare indices.

As for poverty dominance, for both of these approaches we will make use of classes of social welfare indices defined by the reactions of their indices to changes in or reallocations of income. The class of first-order social welfare indices, Ω^1 , regroups all anonymous and continuous social welfare indices that are weakly increasing in income (again, by “weakly” increasing, we mean that the indices never fall when the living standard of any one individual increases, and that they sometimes increase following such a change). They thus obey the Pareto principle.

The class Ω^2 of social welfare indices regroups all of the Ω^1 social welfare indices that are weakly increasing in mean-preserving equalising transfers. Such transfers redistribute one dollar of living standard from a richer to a poorer person. The indices that are members of Ω^2 thus obey the Pigou-Dalton principle of transfers.

The class Ω^3 of social welfare indices include all indices that are members of Ω^2 and that obey the principle of “transfer sensitivity”. This principle requires that equalising transfers have a greater impact on social welfare when they occur lower down in the distribution of income than when they take place across

richer individuals, higher up in the distribution of income. The importance of the Pigou-Dalton principle of transfers is thus decreasing in the income of the recipient of a transfer. For additive social welfare indices, this requires the concavity of individual utility functions to be decreasing in income. To understand further this property, define a “favourable composite transfer” as a mean- and variance preserving dual transfer, a first one which equalises and another one which disequalises income. The equalising transfer takes one dollar from a richer individual 2 to a poorer individual 1, and the disequalising transfer takes one dollar from a poorer individual 3 and give it to a richer individual 4. Since the disequalising transfer takes place higher up in the distribution of income, this means that:

$$y_2 - y_1 = y_4 - y_3, \text{ and } y_1 < y_3.. \quad (131)$$

The class Ω^3 of social welfare indices regroups all of the Ω^2 indices that weakly increase following this favourable composite transfer, and thus obey the principle of transfer sensitivity.

Higher orders of classes can be defined analogously. Ω^4 , for instance, would be defined using two favourable composite transfers, and would require the favourable composite transfer that takes place lower down in the distribution of income to have more impact on the social welfare indices. Generally speaking, membership in a higher order class of indices requires these indices to be more more sensitive to the income of the very poor. Ω^s implies membership in Ω^{s-1} , and for additive social welfare functions of the type of (32), this implies that $(-1)^{(i)}U^{(i)}(Q(p)) \leq 0$ for $i = 1, \dots, s$.

To check whether social welfare in B is greater than in A , there exist two approaches : an income-censoring (or primal) approach, and a p -truncating approach (or dual approach). We look at them in turn.

11.1 Primal approach

For social welfare in B to be greater than in A for all of the Ω^1 indices, we need to have:

First-order welfare dominance (primal):

$$\begin{aligned} W_A - W_B < 0 \text{ for all } W \in \Omega^1 \\ \text{iff } D_A^1(\zeta) > D_B^1(\zeta) \text{ for all } \zeta \in [0, \infty[. \end{aligned} \quad (132)$$

This is the same as requiring that $P_A(\zeta; \alpha = 0) > P_B(\zeta; \alpha = 0), \forall \zeta \in [0, \infty[$, that is, as requiring that the headcount index be higher for A than for B for all non-

negative poverty lines ζ . Similar relations apply to second- and s -order welfare dominance:

Second-order welfare dominance (primal):

$$\begin{aligned} W_A - W_B &< 0 \text{ for all } W \in \Omega^2 \\ \text{iff } D_A^2(\zeta) &> D_B^2(\zeta) \text{ for all } \zeta \in [0, \infty[\end{aligned} \quad (133)$$

s -order welfare dominance (primal):

$$\begin{aligned} W_A - W_B &< 0 \text{ for all } W \in \Omega^s \\ \text{iff } D_A^s(\zeta) &> D_B^s(\zeta) \text{ for all } \zeta \in [0, \infty[\end{aligned} \quad (134)$$

Hence, for the primal approach, checking for s -order welfare dominance simply requires computing the FGT indices for $\alpha = s - 1$ and for the entire range of non-negative poverty lines.

11.2 Dual approach

The dual condition for first-order welfare dominance is as follows:

First-order welfare dominance (dual):

$$\begin{aligned} W_A - W_B &< 0 \text{ for all } W \in \Omega^1 \\ \text{iff } Q_A(p) &< Q_B(p) \text{ for all } p \in]0, 1] \end{aligned} \quad (135)$$

This requires checking that “Pen’s parade of dwarfs and giants” be everywhere higher in B than in A , whatever the p -quantiles being compared. The dual second-order condition makes use of the Generalised Lorenz curve, defined as:

$$GL(p) = \int_0^p Q(q) dq. \quad (136)$$

The generalised Lorenz curve cumulates income up to percentile p . It is simply the non-normalised Lorenz curve. $GL(p)$ is the per capita living standard that would be available if society could rely only on the income of the bottom p proportion of the population. Said differently, it is the absolute contribution to a society’s *per capita* living standard of the bottom p proportion of the population.

Second-order welfare dominance (dual):

$$\begin{aligned} W_A - W_B &< 0 \text{ for all } W \in \Omega^2 \\ \text{iff } GL_A(p) &< GL_B(p) \text{ for all } p \in [0, 1] \end{aligned} \quad (137)$$

Figure 24 shows four cases of comparisons of average income and inequality across two distributions A and B . In Case 1, A inequality-dominates B , and it also has a higher average living standard. Hence, there is generalized-Lorenz-dominance of A over B . In Case 2, A also inequality-dominates B according to the Lorenz criterion, but $\mu_A < \mu_B$; $GL_A(p)$ crosses $GL_B(p)$ and there can be no unambiguous social welfare ranking. Case 3 depicts an ambiguous ranking of inequality across A and B . However, because μ_A is well above μ_B , the generalized Lorenz curve for A dominates that for B . Finally, Case 4 shows a circumstance in which inequality and social welfare rankings clash. A has unambiguously less inequality than B according to the Lorenz criterion, but μ_A being significantly below μ_B , A has unambiguously less social welfare than B according to the generalized-Lorenz criterion.

Part IV

**Poverty and equity: policy design
and assessment**

12 Poverty alleviation: policy and growth

12.1 Measuring the benefits of public spending

A major question in analyzing the impact of policy in poverty and welfare concerns the equity of the distribution of public expenditures.

Say that the expected benefit at rank p is given by $\bar{B}(p)$. (This can be estimated non-parametrically.) The cumulative effect of the benefit up to rank p is given by:

$$\text{GC}_B(p) = \int_0^p B(q) dq. \quad (138)$$

This shows the absolute contribution of a bottom proportion p of the population to *per capita* benefits. When the mean of the benefit is given by μ_B , the concentration curve of the benefit up to rank p can be defined as:

$$C_B(p) = \frac{\text{GC}(p)}{\mu_B}. \quad (139)$$

The concentration curve $C_B(p)$ at p gives the percentage of the total benefits that accrue to those with rank p or lower. Recall that the Lorenz curve of income is given by:

$$L(p) = \frac{1}{\mu} \int_0^p Q(q) dq. \quad (140)$$

12.2 Checking the distributive effect of public expenditures

- For social welfare: we may compare the distributive curves $B(p)$ and $\text{GC}_B(p)$ across types of expenditures, $\forall p \in [0, 1]$.
- For poverty: we may compare the distributive curves $B(p)$ and $\text{GC}_B(p)$ across types of expenditures, $\forall p \in [0, F(z)]$.
- For assessing equity and relative poverty, we may compare $B(p)/\mu_B$ with $X(p)/\mu_X$, and $C_B(p)$ with $L(p)$, $\forall p \in [0, 1]$ if we are concerned about the whole population, otherwise for $\forall p \in [0, F(z)]$ if we are only concerned about the poor.

We may also use "income growth curves", $(N(p) - X(p))/X(p)$, to detect whether net incomes are higher than gross (pre-benefit) incomes. This is equivalent to graphing changes in quantile curves for net incomes. If growth is everywhere positive, absolute poverty must have been decreased. If the growth curve for the poor is above mean growth, then the system is "pro-poor", in the sense that it increases the incomes of the poor faster than those of the rest of the population. An exactly equivalent test can be done by comparing the normalized quantiles for gross and net incomes – recall that normalized quantiles $Q(p)/\mu$ are just incomes as a proportion of mean income. If normalized quantiles for the poor are increased by the benefit system, then the system is "pro-poor".

Income growth curves may also be used to consider the impact of state benefits and public expenditures upon *relative* poverty. The procedure is similar to that of checking whether the benefit system is pro-poor – we compare income growth for the poor to the growth of central tendency of the distribution of income for the whole population. One difference with the measurement of pro-poor growth is that the central tendency of interest may be median income (or any other quantile) if the relative poverty line is set as a proportion of that median income.

Measuring the benefits of public spending can also be done by assessing by how much the cumulative income of bottom proportions p of the poor is affected by that public spending. If that cumulative income increases for all p lower than the initial headcount $F(z)$, then poverty is necessarily reduced by the benefits of public spending for all distribution-sensitive poverty indices. Further, if the growth in the cumulative incomes of the poor is larger than the growth in mean income, then public spending reduces inequality and enhances for the benefit of the poor the vertical equity of the distribution of living standards. It is equivalent to checking whether the Lorenz curve is pushed up by public spending for all possible cumulative proportions of the poor.

12.3 The impact of targeting and public expenditure reforms on poverty

For policy purposes, it is often more useful to assess the impact of *reforms* to a benefit or public expenditure system than to evaluate the effect of *existing* systems. An essential question to answer in the assessment of such reforms is: how will the benefits of such reforms be spread? There is clearly an infinite number of ways in which reforms to the pattern of public expenditures can be made.

We consider five such "reforms" in this section. The first one channels public

expenditure benefits to members of specific and easily observable socio-economic groups. The main issue then is, "in which socio-economic group is additional public money best spent to reduce aggregate poverty?" The second type of reform consists in an increase in public expenditures that raises all incomes in some socio-economic groups by some proportional amount. Again, an important question is for which socio-economic group this increase of public expenditures would reduce aggregate poverty the fastest. This second type of reform can also be thought as a process that increase the quality of infrastructure and the quantity of economic activity in a particular group or region – in a way that affects proportionally all incomes and that is thus distributionally neutral in the sense of not affecting inequality within the groups affected.

The third type of reform considers a change in the price of some commodities, either through some macroeconomic or external shocks, or through a change in commodity taxes or subsidies. How is the distribution of well-being, and poverty in particular, affected by such a price change? The following question we ask is: what type of reform to a system of commodity taxes and subsidies could we implement, with no change in overall government revenues, but with a fall in poverty? That is, which commodities should be prime targets for a reduction in their tax rate or for an increase in their rate of subsidy? The fifth and last type of reform affects proportionally all incomes of a certain type – such as some type of farm income, the labor income of some type of workers, *etc...*

For all such reforms, we measure their poverty impact by the change in the FGT poverty indices that they cause. Recall also that the use of the FGT poverty indices is closely connected to checks for stochastic dominance and ethical robustness of poverty changes. Hence, we can use the methods below to determine how the reforms affect poverty as measured not only by the FGT poverty indices, but by all poverty indices which obey some ethical conditions. For instance, if we find that some form of targeting decreases a FGT index of some α value for a range $[0, z^+]$ of poverty lines, then we know that the reform will also decrease all poverty indices of ethical order $\alpha + 1$, whatever the choice of poverty line within $[0, z^+]$.

12.3.1 Group-targeting a constant amount

We consider first the effect of a transfer of a constant amount of income to every one in a group k . For this, recall that the FGT index can be decomposed as:

$$P(z; \alpha) = \sum_{k=1}^K \phi(k) P(k; z; \alpha). \quad (141)$$

The *per capita* cost to the government of granting an *equal* amount $\eta(k)$ to each member of a group k is equal to:

$$R = \sum_{k=1}^K \phi(k) \eta(k). \quad (142)$$

Aggregate poverty after such transfers is equal to:

$$P(k; z; \alpha) = \int_0^1 [z - Q(k; p; z) - \eta(k)]_+^\alpha dp. \quad (143)$$

To determine which group k should be of greatest priority for the targeting of government expenditures, we need to determine for which group k government expenditures (in the form of η_k) reduce aggregate poverty the most. In other words, we need to compare across k the aggregate poverty reduction benefits of targeting \$1 to group k .

When $\alpha \neq 0$, we can show that the marginal reduction of aggregate poverty per dollar of *per capita* government expenditures is given by:

$$\frac{\partial P(z; \alpha)}{\partial \eta(k)} / \frac{\partial R}{\partial \eta(k)} = -\alpha P(k; z; \alpha - 1) \leq 0. \quad (144)$$

To reduce $P(z; \alpha)$ the most, we must therefore target groups for which $P(k; z; \alpha - 1)$ is the greatest. The greater the value of α , the greater the chance that we will favor those groups where extreme poverty is highest.

When $\alpha = 0$, the per dollar reduction of aggregate poverty is given by $f(k; z)$, group k 's density of income at the poverty line:

$$\frac{\partial P(z; \alpha = 0)}{\partial \eta(k)} / \frac{\partial R}{\partial \eta(k)} = -f(k; z) \leq 0. \quad (145)$$

We must therefore target those groups with the greatest proportion of people just around the poverty line, regardless of how much poverty there is below that poverty line – another consequence of the distributive insensitivity of the head-count index.

12.3.2 Inequality-neutral targeting

Consider now a transfer that is proportional to the income $Q(k; p; z)$ of each member of a group k . Let this proportional increase be $\lambda(k) - 1$. The FGT index for group k after such a transfer is then:

$$P(k; z; \alpha) = \int_0^1 [z - Q(k; p; z) \cdot \lambda(k)]_+^\alpha dp. \quad (146)$$

The marginal impact of a change in $\lambda(k)$ is given by

$$\frac{\partial P(z; \alpha)}{\partial \lambda(k)} = \alpha \phi(k) [P(k; z; \alpha) - zP(k; z; \alpha - 1)] \leq 0. \quad (147)$$

How (147) varies across values of k depends on two factors. First, there is the factor $[P(k; z; \alpha) - zP(k; z; \alpha - 1)]$. Groups in which there is a significant presence of extreme poverty will tend to see their $P(k; z; \alpha)$ poverty indices fall significantly with α , thus leading to a large value of $[P(k; z; \alpha) - zP(k; z; \alpha - 1)]$. We may thus expect that these groups should be a target for government targeting. However, those groups with considerable incidence of extreme poverty are also those for which a proportional increase in income has the least impact on the average income of the poor – since there is then little income on which growth may take effect. Hence, whether those groups with higher incidence of extreme poverty will exhibit a higher value of $[P(k; z; \alpha) - zP(k; z; \alpha - 1)]$ is ambiguous.

The second factor that enters into (147) is population share $\phi(k)$. *Ceteris paribus*, targeting government expenditures (in the form of an increase in $\lambda(k)$) to groups with a high population share will naturally tend to decrease overall poverty fastest. But this fails to take into account that a given increase in $\lambda(k)$ will generally be more costly for the government to attain for groups with a large share of the population. Because of this, we may instead wish to compare across groups the ratio of the benefit in poverty reduction to the group *per capita* increase in income. Assume that the cost of this group *per capita* income increase is entirely borne by the government. The *per capita* revenue impact of such a transfer on the government budget equals $\partial R / \partial \lambda(k)$, where:

$$R = \phi(k)\mu(k)\lambda(k). \quad (148)$$

When $\alpha \neq 0$, the reduction of aggregate poverty per dollar *per capita* spent is then:

$$\frac{\partial P(z; \alpha)}{\partial \lambda(k)} / \frac{\partial R}{\partial \lambda(k)} = \frac{\alpha [P(k; z; \alpha) - zP(k; z; \alpha - 1)]}{\mu(k)} \leq 0. \quad (149)$$

To reduce $P(z; \alpha)$ the fastest, the government should therefore target those groups for which the term on the right is the greatest in absolute value. Compared to (147), (149) does not feature population shares as a factor, since it is cancelled by the revenue impact of the government transfer. There now appears, however, the term $\mu(k)$ in the denominator. Indeed, if it must bear the entire cost of the income increase, the government will have to pay more to achieve a given increase in $\lambda(k)$ for those groups with a high average income than for those with a lower average income level. Finally, and for the same reasons as above, whether those groups with higher incidence of extreme poverty will exhibit a higher value of $[P(k; z; \alpha) - zP(k; z; \alpha - 1)]$ is ambiguous.

When $\alpha = 0$, the per dollar reduction of aggregate poverty from a proportional-to-income transfer is given by

$$\frac{\partial P(z; \alpha = 0)}{\partial \lambda(k)} / \frac{\partial R}{\partial \lambda(k)} = -\frac{z \cdot f(k; z)}{\mu(k)} \leq 0. \quad (150)$$

Those groups with a high density of income at the poverty line, and whose average income is small, are then a prime target for poverty-efficient proportional-to-income transfer scheme.

12.3.3 Price changes

The level of relative prices is an important determinant of the distribution of living standards, and can therefore matter significantly for poverty analysis. Governments can affect their levels directly or indirectly. Maintaining high import tariffs or failing to implement regulations to foster competition may protect national producers by maintaining high domestic producer prices, but it will also lead to high consumption prices, which will hurt consumers. The use of sales and indirect

taxes to raise tax revenues also affects relative prices, and thus consumer and producer well-being. Price subsidies on food, education, energy and transportation are yet another example of government policies which affect relative prices and thus poverty.

To see how changes in relative prices (and therefore how price-changing reforms) can impact poverty, we denote by y a household-specific level of exogenous income, and express consumers' preferences as ϑ . The indirect utility function is given by $V(y, q; \vartheta)$, where q is a vector of consumer and producer prices. We define a vector of reference prices as q^R – this is necessary to assess consumers' well-being at constant prices. Denote the real income in the post-reform situation by y^R , where y^R is measured on the basis of the reference prices q^R . y^R is implicitly defined by $v(y^R, q^R; \vartheta) = v(y, q; \vartheta)$, and explicitly by the real income function $y^R = R(y, q, q^R; \vartheta)$, where

$$V(R(y, q, q^R; \vartheta), q^R; \vartheta) \equiv V(y, q; \vartheta). \quad (151)$$

By definition, y^R gives the level of income that provides under q^R the same real living standard as y yields under q .

We then wish to determine how real living standards are affected by a marginal change in prices. Let $x_c(y, q; \vartheta)$ be the net consumption of good c (which can be negative if the individual or household is a net producer of good c) of a consumer/producer with income y , preferences ϑ and facing prices q . Let q_c be the price of good c . Using Roy's identity and setting reference prices to pre-reform prices, we find:

$$\begin{aligned} \left. \frac{\partial R(y, q, q^R; \vartheta)}{\partial q_c} \right|_{q=q^R} &= \left. \frac{\partial V(y, q; \vartheta) / \partial q_c}{\partial V(y^R, q^R; \vartheta) / \partial y^R} \right|_{q=q^R} \\ &= - \left. \frac{\partial V(y, q; \vartheta) / \partial y}{\partial V(y^R, q^R; \vartheta) / \partial y^R} \right|_{q=q^R} \cdot x_c(y, q; \vartheta) \\ &= -x_c(y, q^R; \vartheta). \end{aligned} \quad (152)$$

Equation (152) says that observed pre-reform net consumption of good c is a sufficient statistic to know the impact on the real living standard of a marginal change in the price of good c . This simple relationship is in fact also valid for rationed goods. Equation (152) gives a "first-order approximation" to the true change in real income that occurs from a change in the price of good c . The

approximation is *exact* when the price change is marginal. It becomes less exact if the price change is non-marginal and if the compensated demand for good c varies significantly with q_c .

Assume that preferences ϑ and exogenous income y are jointly distributed according to the distribution function $F(y, \vartheta)$. The conditional distribution of ϑ given y is denoted by $F(\vartheta | y)$, and the marginal distribution of income y is given by $F(y)$. Let preferences belong to the set Θ , and assume income to be distributed over $[0, a]$. Expected consumption of good c at income y is given by $x_c(y, q)$, such that

$$x_c(y, q) = E_{\vartheta} [x_c(y, q; \vartheta)] = \int_{\Omega} x_c(y, q; \vartheta) dF(\vartheta | y). \quad (153)$$

By (152), $-x_c(y, q)$ is also proportional to the expected fall in real living standards of those with income y that arises from an increase in q_c .

Let $x_c(q)$ then be the *per capita* consumption of the c^{th} good, defined as $x_c(q) = \int_0^a x_c(y, q) dF(y)$. By (152), $x_c(q)$ is also the average welfare cost of an increase in the price of good c . As a proportion of *per capita* consumption, consumption of good c at income y is expressed as $\bar{x}_c(y, q) = x_c(y, q)/x_c(q)$.

It is now at last useful to see how the FGT indices $\bar{P}(z; \alpha)$ are affected by a change in the price of good k ¹⁰. Using (152), we can show that:

$$\left. \frac{\partial \bar{P}(z; \alpha)}{\partial q_c} \right|_{q=q^R} = \begin{cases} x_c(z, q^R) f(z), & \text{if } \alpha = 0 \\ \alpha z^{-\alpha} \int_0^z x_c(y, q^R) (z - y)^{\alpha-1} dF(y) & \text{if } \alpha = 1, 2, \dots \end{cases} \quad (154)$$

where $f(z)$ is again the density of income at z . When the effect on the FGT indices of an increase in the price of a good c is graphed over a range of poverty lines z , this generates the “consumption dominance” $CD_c(z; \alpha)$ curve of good c :

$$CD_c(z; \alpha) = \frac{\partial \bar{P}(z; \alpha)}{\partial q_c}, \alpha = 0, 1, 2, \dots \quad (155)$$

Note that the impact on poverty depends on α and z . By (154), $CD_c(z; \alpha = 0)$ only takes into account the consumption pattern of those precisely at z . The impact of an increase in the price of good c on the headcount index will be large if there are many around the poverty line ($f(z)$ is large) and/or if they consume much of good c ($x_c(z, q^R)$ is large). The $CD_c(z; \alpha = 1)$ curve gives the absolute contribution to total consumption of c of those individuals with income less than

¹⁰For normalized FGT indices, we simply multiply the results by z^α .

z . It is therefore an informative statistics on the distribution of consumption expenditures, similar in content to the generalized concentration curve $GC_{x_c}(p)$ for good c – which gives the absolute contribution to total x_c consumption of those below a certain rank p . For $\alpha = 2, 3, \dots$, progressively greater weight is given for the computation $CD_c(z; \alpha)$ to the shares of those with higher poverty gaps.

12.3.4 Tax/subsidy policy reform

The above section gave us the tools needed to assess the impact of marginal price changes on poverty. We may now use these tools to assess whether a revenue-neutral tax and subsidy reform could be implemented that would reduce aggregate poverty.

For this, we need to consider the government budget constraint, and more particularly the net revenues that the government raises from its policy of commodity taxes and subsidies. Setting producer prices to 1 and assuming them for simplicity to be constant and invariant to changes in t , the vector of tax rates on the C goods, we then have $q = 1 + t$ and $dq_c = dt_c$, where q_c and t_c denote respectively the consumer price of and the tax rate on good c . Let *per capita* net commodity tax revenues be denoted as $R(q)$. They are equal to $R(q) = \sum_{c=1}^C t_c x_c(q)$. Without loss of generality, assume that the government’s tax reform increases the tax rate on the j^{th} commodity and uses the proceeds to decrease the tax rate (or to increase the subsidy) on the l^{th} commodity. Revenue neutrality (a conventional assumption in the optimal taxation literature) of the tax reform requires that

$$dR(q) = \left[x_j(q) + \sum_{c=1}^C t_c \frac{\partial x_c(q)}{\partial q_j} \right] dq_j + \left[x_l(q) + \sum_{c=1}^C t_c \frac{\partial x_c(q)}{\partial q_l} \right] dq_l = 0. \quad (156)$$

Define γ as

$$\gamma = \frac{x_l(q) + \sum_{c=1}^C t_c \frac{\partial x_c(q)}{\partial q_l}}{x_l(q)} \bigg/ \frac{x_j(q) + \sum_{c=1}^C t_c \frac{\partial x_c(q)}{\partial q_j}}{x_j(q)}, \quad (157)$$

The numerator in (157) gives the marginal tax revenue of a marginal increase in the price of good l , per unit of the average welfare cost that this price increase imposes on consumers. Equivalently, this is 1 minus the deadweight loss of taxing good l , or the inverse of the marginal cost of public funds (MCPF) from taxing l (see Wildasin (1984)). The denominator gives exactly the same measures for an increase in the price of good j . γ is thus the economic (or “average”) efficiency of taxing good l relative to taxing good j . We may thus interpret γ as the efficiency

cost of taxing j relative to that of taxing l (the MCPF for j over that for l). The higher the value of γ , the less economically efficient is taxing good j .

By simple algebraic manipulation, we can then rewrite equation (156) as

$$dq_j = -\gamma \left(\frac{x_l(q)}{x_j(q)} \right) dq_l, \quad (158)$$

which fixes dq_j as a revenue-neutral proportion of dq_l . This last relationship gives us a nice synthetic expression for the impact on a FGT index $P(z; \alpha)$ of a revenue-neutral tax reform that increases the tax on a good l for a benefit of a fall in the tax on a good j :

$$\left. \frac{\partial P(z; \alpha)}{\partial t_l} \right|_{\text{revenue neutrality}} = CD_l(z; \alpha) - \gamma x_l(q) CD_j(z; \alpha) / x_j(q), \quad (159)$$

As mentioned above, we may wish to check whether such a tax reform would lead to an "ethically robust" fall in poverty, namely, that poverty would fall as measured by any one of the poverty indices of some ethical order and for a range of poverty lines. To do this, it is useful to define and use normalized CD curves, denoted as $\overline{CD}_c(z; \alpha)$. Normalized CD curves are just the above-defined CD curves for good c normalized by the average consumption of that good, $x_c(q)$:

$$\overline{CD}_c(z; \alpha) = \frac{CD_c(z; \alpha)}{x_c(q)}. \quad (160)$$

\overline{CD} curves are thus the ethically weighted (or social) cost of taxing c as a proportion of the average welfare cost. Comparing normalized $\overline{CD}_c(z; \alpha)$ curves allows us to compare the distributive benefits of decreasing tax rates (or increasing subsidies) across commodities. If $\overline{CD}_l(z; \alpha) > \overline{CD}_j(z; \alpha) \geq 0$, then investing 1 dollar of public expenditures to decrease the burden of taxation on l will have a greater distributive benefit (will lead to greater poverty reduction) than for commodity j .

For overall social efficiency, we must also take into the parameter of economic efficiency, γ . We may check whether a tax reform is "poverty efficient" by verifying whether the following condition holds:

$$\overline{CD}_l(z; \alpha) - \gamma \overline{CD}_j(z; \alpha) \geq 0, \forall z \in [0, z^+]. \quad (161)$$

If condition (161) holds, then it is s -order poverty efficient (where $s = \alpha + 1$) to decrease the tax on good l for a revenue-neutral benefit of an increase in the tax on good j . Poverty will decrease following this tax reform for all poverty indices of the s order and for all poverty lines within $[0, z^+]$.

12.3.5 Income-component and sectoral growth

It is just a matter of notational change to use the tools developed above to assess the poverty impact of growth in some income component or in some sector of economic activity. We may assess, for instance, by how much would aggregate poverty fall per percentage of growth in the industrialized sector, or per dollar of growth in agricultural income (an income component that enters into aggregate income).

Assume that total income X is the sum of C income components, with $X = \sum_{c=1}^C \lambda_c X_c$, and where λ_c is a factor that multiplies income component c and that can be subject to growth. The derivative of the normalized FGT with respect to λ_c is given by

$$\left. \frac{\partial \bar{P}_X(z; \alpha)}{\partial \lambda_c} \right|_{\lambda_c=1, c=1, \dots, C} = CD_c(z; \alpha), \quad (162)$$

where this $CD_c(z; \alpha)$ curve is now a "component dominance" curve for income component or for sectoral income X_c . Multiplied by a proportional change $d\lambda_c$, $CD_c(z; \alpha)$ will give us the marginal change in the FGT indices that we can expect from growth in sector or in component c .

We can reasonably expect, however, that a given percentage change in a large sector or for a large income component will have a larger impact on poverty than otherwise. To take this factor into account, we may wish instead to compute the change in the FGT indices *per dollar of per capita* growth. This is given by the normalized CD curves for component c :

$$\left. \frac{\partial \bar{P}(z; \alpha) / \partial \lambda_c}{\partial \mu_X / \partial \lambda_c} \right|_{\lambda_c=1, c=1, \dots, C} = \overline{CD}_c(z; \alpha). \quad (163)$$

Finally, we may wish to compute the *elasticity* of poverty with respect to overall growth, where again that overall growth comes strictly from growth in a component X_c . From (163), this is given by:

$$\frac{\left. \frac{\partial \bar{P}(z; \alpha) / \partial \lambda_c}{\partial \mu_X / \partial \lambda_c} \right|_{\lambda_c=1, c=1, \dots, C}}{\bar{P}(z; \alpha) / \mu_X} = \frac{\overline{CD}_c(z; \alpha)}{\bar{P}(z; \alpha)} \cdot \mu_X. \quad (164)$$

12.4 Overall growth elasticity of poverty

How fast can inequality-neutral growth in the economy be expected to reduce poverty? From which group inequality-neutral growth can be expected to reduce

aggregate poverty the fastest? And in which group would poverty fall the fastest due to such growth?

Using (147) above, it can be shown that the elasticity of total poverty $P(z; \alpha)$ with respect to total income – when growth in total income comes exclusively from inequality-neutral growth in group k – equals

$$\varepsilon^y(k; z; \alpha) = \frac{\alpha [P(k; z; \alpha) - zP(k; z; \alpha - 1)]}{P(z; \alpha)} \quad (165)$$

for $\alpha \neq 0$. As discussed above, it is ambiguous whether growth in groups with significant incidence of extreme poverty would be those whose growth would be most beneficial for overall poverty reduction. Replacing $P(k; z; \alpha)$ and $P(k; z; \alpha - 1)$ by $P(z; \alpha)$ and $P(z; \alpha - 1)$ respectively gives as a special case the elasticity of poverty with respect to inequality-neutral growth in the overall economy, $\varepsilon^y(z; \alpha)$. Replacing $P(z; \alpha)$ by $P(k; z; \alpha)$ provides as another special case the elasticity of poverty in group k with respect to inequality-neutral growth in that same group.

When $\alpha = 0$, (165) becomes:

$$\varepsilon^y(k; z; \alpha = 0) = \frac{-zf(k; z)}{F(z)}. \quad (166)$$

This expression has a nice graphical interpretation. To see this, consider Figure 30 where the density $f(y)$ of income at different y is shown. Recall that the area underneath the $f(y)$ curve up to $y = z$ gives the headcount $F(z)$. The term $z \cdot f(k; z)$ in (166) is the area in Figure 30 of the rectangle with width z and height $f(z)$. Hence, the elasticity (in absolute value) of the headcount with respect to inequality-neutral growth is given in Figure 30 by the ratio of the rectangular area $z \cdot f(k; z)$ over the shaded area $F(z)$.

It is clear, then, that this elasticity is larger than one whenever the poverty line z is lower than the (first) mode of the distribution. In fact, it will be above one in Figure 30 for any poverty line up to approximately z' . For poverty lines larger than z' , the growth elasticity will in absolute value fall below 1. This can have important policy consequences. For societies in which the poverty line is deemed to be lower than the mode (which is usually not far from the median), then the headcount in these societies will fall at a proportional rate that is faster than the growth rate in average living standards. But for societies in which the headcount is initially high (larger than 0.5, say), we can expect the growth elasticity of the

headcount to be lower than 1. This means that inequality-neutral growth can be expected to have a proportionately smaller impact on the number of the poor in poorer societies than in richer ones.

12.5 The Gini elasticity of poverty

It may also be of interest to predict how changes in inequality will affect poverty. The difficulty is that, unlike for the case of growth in mean income, it is not obvious which pattern of changing inequality we should consider. As discussed above, a natural reference case for analyzing the impact of growth is the case of inequality-neutral growth – all incomes vary proportionately by the same growth rate in mean income. For inequality changes, which inequality measure should we use? And even if we were to agree on the choice of one single summary inequality index, there are many different ways in which a given change in such an index may be obtained, even keeping mean income constant. And these different ways could also have a dramatically different impact on poverty.

So trying to predict the effect on poverty of a process of changing inequality, through the use of a single indicator of inequality, is really to ask too much of summary indices of inequality. There cannot exist any stable structural relationship between inequality indices and poverty, even keeping mean income constant. This casts severe doubt on the structural validity of the many studies that regress changes in poverty indices upon changes in inequality indices.

What can be done, however, is to *illustrate* how some peculiar and simplistic pattern of changing inequality can affect poverty. Such an illustration can be made using the single-parameter (λ) process of bi-polarization shown by equation (17). How does poverty change when inequality changes due to this bi-polarization? For this, we use the most popular indices of poverty and inequality, the FGT and the Gini indices (the result is exactly the same for the broader class of S-Gini indices). The elasticity of poverty with respect to the Gini index – when changing inequality comes from a λ that moves marginally away from 1– is given when $\alpha > 0$ by

$$\varepsilon^G(z; \alpha) = \alpha \left(1 + \frac{P(z; \alpha - 1)}{P(z; \alpha)} (\mu - z) \right) \quad (167)$$

for the un-normalized FGT indices, and by

$$\bar{\varepsilon}^G(z; \alpha) = \left(1 + \frac{P(z; \alpha - 1)}{P(z; \alpha)} \left(\frac{\mu}{z} - z \right) \right) \quad (168)$$

for the normalized FGT. When the headcount is used, we obtain

$$\varepsilon^G(z; \alpha = 0) = f(z) (\mu - z). \quad (169)$$

Note that even with this highly simplified process of changing inequality, the impact of changing inequality on poverty is ambiguous. It depends in part on the sign of $(\mu - z)$. When mean income is below the poverty line, an increase in the Gini index can – and, for the headcount index, *will* – mean a *fall* in poverty.

13 The impact of policy and growth on inequality

13.1 Growth, tax and transfer policy, and price shocks

We may now turn to the impact of policy and growth on inequality. The approach we will use will enable us to consider the impact on inequality of several ways in which income changes may occur. One is growth that takes place within a particular socio-economic group. Another is growth that affects the value of some income sources – such as agricultural income or informal urban labor income. Another is the impact of price changes, which affect real income and its distribution. One more is the impact of changes in some tax or benefit policies, such as changing the subsidy rates on some production or consumption activity, or increasing the amount of monetary transfers made to some socio-economic groups.

For each such income-changing phenomenon, we may be interested in the absolute amount by which inequality will change, or in the absolute amount by which inequality will change for each percentage change in mean real income, or in the elasticity of inequality with respect to mean income.

Assume that we have as above that total standard of living X is the sum of C components, $X = \sum_{c=1}^C \lambda_c X_c$, to which we apply again a factor λ_c . We then have that $\mu_X = \sum_{c=1}^C \lambda_c \mu_{X_c}$. If we are interested in total consumption, then we may think of the X_c as different types of consumption expenditures. If we are thinking of tax and benefit policy, then some of the X_c may be transfers or taxes. If we are alternatively concerned with the impact of sectoral growth on income inequality, then we may think of the X_c as different sources of income, or of the income of different socio-economic groups. By how much, then, is inequality affected by variations in λ_c ?

We will consider two ways of measuring inequality, the Lorenz curve and the S-Gini inequality indices – of which the traditional Gini is again a special case. The derivative of the Lorenz curve of X with respect to λ_c is given by:

$$\left. \frac{\partial L_X(p)}{\partial \lambda_c} \right|_{\lambda_c=1, c=1, \dots, C} = \frac{\mu_{X_c}}{\mu_X} (C_{X_c}(p) - L_X(p)). \quad (170)$$

(170) therefore gives the change in the Lorenz curve per 100% proportional change in the value of X_c . Say that we predict that income component X_c will increase by approximately 10% over the next year. Then we can predict that the Lorenz curve $L_X(p)$ will move by approximately $10\% \cdot \mu_{X_c}/\mu_X (C_{X_c}(p) - L_X(p))$ over that same period. How big an impact this will be on inequality will depend on the size of the proportional change, on the importance of the component (μ_{X_c}), and on the concentration of the component relative to that of total living standards (the difference $C_{X_c}(p) - L_X(p)$).

A similar result is obtained for the Gini indices:

$$\left. \frac{\partial I_X(\rho)}{\partial \lambda_c} \right|_{\lambda_c=1, c=1, \dots, C} = \frac{\mu_{X_c}}{\mu_X} (IC_{X_c}(\rho) - I_X(\rho)). \quad (171)$$

Thus, if for instance the removal of a subsidy or the advent of an external shock is foreseen to increase by 10% the price of a good X_c , the Gini index can be predicted to move by approximately $- [10\% \cdot \mu_{X_c}/\mu_X (C_{X_c}(p) - L_X(p))]$ (the negative sign comes from the fact that an increase in the price of a consumption good leads to a fall in the real value of the expenditures made on that good).

We may also wish to assess the impact on inequality per 100% of mean income change. This is given by

$$\mu_X \left. \frac{\partial L_X(p)/\partial \lambda_c}{\partial \mu_X/\partial \lambda_c} \right|_{\lambda_c=1, c=1, \dots, C} = C_{X_c}(p) - L_X(p) \quad (172)$$

for the Lorenz curve and by

$$\mu_X \left. \frac{\partial I_X(\rho)/\partial \lambda_c}{\partial \mu_X/\partial \lambda_c} \right|_{\lambda_c=1, c=1, \dots, C} = IC_{X_c}(\rho) - I_X(\rho) \quad (173)$$

for the Gini indices. Multiplying the above two expressions by the proportional impact that some change in X_c is predicted to have on total *per capita* income will give the predicted absolute change in inequality. For instance, if we predict that growth in rural areas will lift mean income in a country by 5%, then the Lorenz curve of total income at p will shift by approximately $0.05 (C_{X_c}(p) - L_X(p))$, where X_c is rural income.

Finally, it may be preferred to know the elasticity of inequality with respect to μ_X , when growth comes entirely from X_c . It is given by

$$\frac{C_{X_c}(p)}{L_X}(p) - 1 \quad (174)$$

for the Lorenz curve and by

$$\frac{IC_{X_c}(\rho)}{I_X(\rho)} - 1 \quad (175)$$

for the Gini indices. Thus, an increase in taxes on formal labor income that reduce total mean income (net of taxes) by 1 percent will change the Gini index by $1 - IC_{X_c}(\rho)/I_X(\rho)$ percent.

13.2 Tax and subsidy reform

As in the case of poverty, it is useful to assess the impact of a price reform (through consumption and production taxation) on inequality. Assume that we are interested in the effects of a revenue-neutral tax reform that increases the tax on a good l for a benefit of a fall in the tax on a good j . Recall that γ is the MCPF for j over that for l – the larger the value of γ , the larger the fall in t_j that we can generate for a given revenue-neutral increase in t_l . A 100% increase in the price of good l then has the following impact

$$\left. \frac{\partial L_X(p)}{\partial t_l} \right|_{\text{revenue neutrality}} = \frac{\mu_{x_l(q)}}{\mu_X} \left[\gamma (C_{X_j}(p) - L_X(p)) - (C_{X_l}(p) - L_X(p)) \right] \quad (176)$$

on the Lorenz curve and

$$\left. \frac{\partial I_X(\rho)}{\partial t_l} \right|_{\text{revenue neutrality}} = \frac{\mu_{x_l(q)}}{\mu_X} \left[\gamma (IC_{X_j}(\rho) - I_X(\rho)) - (IC_{X_l}(\rho) - I_X(\rho)) \right] \quad (177)$$

for the Gini indices. When $\gamma = 1$, *viz.*, when the relative economic efficiency of taxing l or j is not of concern, then expressions (176) and (177) reduce to a multiple of the difference between the concentration curves and the concentration indices for the two goods. It is then better for inequality reduction to tax more the good less concentrated among the poor for the benefit of a reduction in the tax rate on the other good, less concentrated among the rich. Changes in the value of γ

may sometimes change the good whose tax rate should be increased for inequality reduction. We may also wish to express the above changes in inequality per 100% change in the value of *per capita* real income. This is then given by

$$\gamma \left(C_{X_j}(p) - L_X(p) \right) - (C_{X_l}(p) - L_X(p)) \quad (178)$$

on the Lorenz curve and

$$\gamma \left(IC_{X_j}(\rho) - I_X(\rho) \right) - (IC_{X_l}(\rho) - I_X(\rho)) \quad (179)$$

for the Gini indices.

Part V

**Estimation and inference for
distributive analysis**

14 Non parametric estimation for distributive analysis

14.1 Density estimation

14.1.1 Univariate density estimation

The use of *non-parametric* density estimation procedures for distributive analysis has grown significantly in the last fifteen years or so¹¹. Before then, a significant literature had developed on the estimation of *parametric* models of the distributions of income. These models assumed that the distribution of income followed a particular functional form with unknown parameters, such as the log-normal, the Pareto, or variants of the beta or gamma distributions. A major aim was then to estimate the unknown parameters of that functional form, and to test whether a given functional form appeared to estimate better the density of income than another functional form.

Non-parametric density estimation does not start by positing a functional form for the distribution of income. Instead, it lets the "data speak for themselves". The method is most easily understood by starting with a review of the density estimation provided by traditional histograms. Histograms provide an estimate of the density of a variable y by counting how many observations fall into bins, and by dividing that number by the width of the bin times the number of observations in the sample. To see this more precisely, denote the origin of the bins by y_0 and the bins of the histogram by $[y_0 + mh, y_0 + (m + 1)h]$ for positive or negative integers m . For instance, if we take $m = 0$, then the bin is described by the interval ranging from the origin to the origin plus h . The value of the histogram over each of these bins is then defined by

$$\hat{f}(y) = \frac{(\# \text{ of } Y_i \text{ in same bin as } y)}{n \cdot (\text{width of bin containing } y)}. \quad (180)$$

Such a histogram is shown on Figure 26 by the rectangles of varying heights over identical widths, starting with origin y_0 . For bins defined by $[y_0 + mh, y_0 + (m + 1)h]$, the bin width is a constant set to h , but we can also allow the widths to vary across the bins of the histogram. The choice of h controls the amount of *smoothing* performed by the histogram. A small bin width will lead to significant fluctuations in the value of the histogram across the h , and a very large width

¹¹This section draws significantly from Silverman (1986), to which readers are referred for more details and in-depth analysis.

will set the histogram to the constant h^{-1} . Choosing an appropriate value for such a smoothing parameter is in fact a pervasive preoccupation in non-parametric estimation procedures, as we will discuss later. The choice of the origin can also be important, especially when n is not very large. There can be, however, little guidance on that latter choice, except perhaps when the nature of the data suggest a natural value for y_0 . One way to avoid choosing such a y_0 is by constructing what will appear soon to be a "naive" kernel density estimator, that is, one in which the point y in $\hat{f}(y)$ is always at the center of the bin:

$$\hat{f}(y) = (2hn)^{-1} (\# \text{ of } Y_i \text{ falling in } [y - h, y + h]). \quad (181)$$

This naive estimator can also be obtained from the use of a weight function $w(u)$, defined as:

$$w(u) = \begin{cases} 0.5 & \text{if } |u| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (182)$$

and by defining

$$\hat{f}(y) = (nh)^{-1} \sum_{i=1}^n w\left(\frac{y - Y_i}{h}\right). \quad (183)$$

This frees the density estimation from the choice of y_0 . For the estimation of the naive density estimator, each observation Y_i provides a "box" of width $2h$ and height 0.5 that is centered at Y_i . $\hat{f}(y)$ over a range of y is just the sum of these n boxes over that range. Because such boxes are not continuous, their sum $\hat{f}(y)$ will also be discontinuous, just like the traditional histogram. There will be "jumps" in $\hat{f}(y)$ at the points $Y_i \pm h$, making this density estimator expositionally unattractive. The naive estimator can also be improved statistically by choosing weighting functions that are smoother than $w(u)$. For this, we can think of replacing the weight function $w(u)$ by a general "kernel function" $K(u)$, such that

$$\hat{f}(y) = (nh)^{-1} \sum_{i=1}^n K\left(\frac{y - Y_i}{h}\right). \quad (184)$$

A smooth kernel estimate of the density function that generated the histogram is shown on Figure 26.

In general, we would wish $\int_{-\infty}^{\infty} K(u)du = 1$, since then $\int_{-\infty}^{\infty} \hat{f}(y)dy = 1$. For $\hat{f}(y)$ to qualify fully as a probability density function, we would also require $K(u) \geq 0$ since we would then be guaranteed that $\hat{f}(y) \geq 0$, although there are sometimes reasons to allow for negativity of the kernel function. h is usually

referred to as the window width, the bandwidth or the smoothing parameter of kernel estimation procedures. There are also arguments to adjust the window width that applies for observation Y_i to the number of observations that surround Y_i , making h larger for areas where there are fewer observations¹². As for the naive density estimator, each observation will provide a box or a "bump" to the density estimation of $f(y)$, and that bump will have a shape and a width determined by the shape of $K(u)$ and the size of h respectively.

The definition of $\hat{f}(y)$ in (184) makes it inherit the continuity and differentiability properties of $K(u)$. It is often sound and convenient to choose a kernel function that is symmetric around 0, with $\int K(u)du = 1$, $\int uK(u)du = 0$ and $\int u^2K(u)du = \sigma_K^2 > 0$. One such kernel function that has nice continuity and differentiability properties is the Gaussian kernel, defined by

$$K(u) = (2\pi)^{-0.5} \exp^{-0.5u^2}. \quad (185)$$

The "bumps" provided by the Gaussian kernel have the familiar bell shapes, are smoothly differentiable up to any desired level, and are such that $\sigma_K^2 = 1$.

14.1.2 Statistical properties of kernel density estimation

The efficiency of non-parametric estimation procedures can be measured naturally by the *mean square error* (MSE) that there is in estimating a function at a point y . The MSE in estimating $f(y)$ is defined by

$$\text{MSE}_y(\hat{f}) = E \left[(\hat{f}(y) - f(y))^2 \right]. \quad (186)$$

The most common way of defining a measure of *global* accuracy simply sums the mean square error across values of y . This yields the *mean integrated square error* (or MISE), a measure of the accuracy of estimating $f(y)$ over the whole range of y :

$$\text{MISE}(\hat{f}) = \int \text{MSE}_y(\hat{f}) dy = E \left[\int (\hat{f}(y) - f(y))^2 dy \right]. \quad (187)$$

The relative efficiency of a particular choice of a kernel function $K(u)$ can then be assessed relative to that choice of the kernel function which would minimize the MISE. The Gaussian kernel function has very good efficiency properties, although they are not quite as good as some other (less smooth) kernel functions,

¹²This is done for instance by the nearest neighbor and the adaptive kernel methods.

such as the (efficiency-optimal) Epanechnikov, the biweight or the triangular kernels, which are described and discussed for instance in Silverman (1986) (see in particular Table 3.1).

14.1.3 Choosing a window width

Even, however, if we were to agree on a particular shape for an observation-centered kernel function, there would still remain the question of which window width to choose. Again, conditional on the choice of a particular form for $K(u)$, we can choose the window width that minimizes the MISE. To see what this implies, note first that we can decompose the MSE at y as a sum of the squared of the bias and of the variance that there is in estimating $\hat{f}(y)$:

$$\text{MSE}_y(\hat{f}) = \left[E(\hat{f}(y) - f(y))^2 \right] + \text{var}\hat{f}(y). \quad (188)$$

For symmetric kernel functions, the bias can be shown to be approximately equal to

$$0.5h^2\sigma_K^2 f^{(2)}(y), \quad (189)$$

and the variance to:

$$n^{-1}h^{-1}f(y)c_K. \quad (190)$$

where $c_K = \int K(u)^2 du$. Substituting (189) and (190) in (188) then gives:

$$\text{MSE}_y(\hat{f}) = 0.5h^2\sigma_K^2 f^{(2)}(y) + n^{-1}h^{-1}f(y)c_K. \quad (191)$$

Hence, considering (189), we find that the bias of $\hat{f}(y)$ will be low if the kernel function has a low variance, since it is then the observations that are "closer" to y that will count more, and it is those observations that provide the least biased estimate of the density at y . But the bias also depends on the curvature of $f(y)$: in the absence of such a curvature, the density function is linear and the bias provided by using observations on the left of y is just (locally) outweighed by the bias provided by using observations on the right of y . Looking at (190), we find that a flatter kernel (with a lower c_K) decreases the variance of $\hat{f}(y)$, since the estimator is then more equally dependent on the value of the sample observations. We also obtain the familiar result that the variance of the estimator decreases proportionately with the size of the sample.

An increase in h plays an offsetting role on the precision of $\hat{f}(y)$, as is shown by (191). When $f^{(2)}(y) \neq 0$, a large h increases the bias by making the estimators too smooth: too much use is made of those observations that are not so close to y . Conversely, a large h reduces the variance of $\hat{f}(y)$ by making it less variable and less dependent on the particular value of those observations that are close to y . Hence, in choosing h in an attempt to minimize $\text{MISE}(\hat{f}(y))$, a compromise needs to be struck between the competing virtues of bias and variance reductions. The precise nature of this compromise will depend on the shape of the kernel function as well as on the true population density function. For instance, if the Gaussian kernel is used and if the true density function is normal with variance σ^2 , then the choice of h that minimizes the MISE is given by (see for instance Silverman (1986) p.45):

$$h^* = 1.06\sigma n^{-0.2} \quad (192)$$

This value of h^* is conditional on both $K(u)$ and $f(y)$ being normal pdf. Silverman (1986) also argues for a more robust choice of h^* , given by

$$h^* = 0.9An^{-0.2}, \quad (193)$$

where $A = \min(\text{standard deviation}, \text{interquartile range}/1.34)$. This is because (193)

(...) will yield a mean integrated square error within 10% of the optimum for all the t -distributions considered, for the log-normal with skewness up to about 1.8, and for the normal mixture with separation up to 3 standard deviations. (...) For many purposes it will certainly be an adequate choice of window width, and for others it will be a good starting point for subsequent fine tuning. (Silverman (1986), p.48)

Further (asymptotic) results show that, under some mild assumptions — in particular, that the density function $f(y)$ is continuous at x , and that $h \rightarrow 0$ and $nh \rightarrow \infty$ as $n \rightarrow \infty$ — the kernel estimator $\hat{f}(y)$ converges to $f(y)$ as $n \rightarrow \infty$. When h is chosen optimally, it is of the order of $n^{-1/5}$, and by (191) MISE is then of the order of $n^{-4/5}$. This is slightly lower than the usual rate of convergence of parametric estimators, which is n^{-1} .

14.1.4 Multivariate density estimation

Kernel estimation can also be used for multivariate density estimation. Let \mathbf{u} , \mathbf{y} and \mathbf{Y}_i be d -dimensional vectors. We can estimate a d -dimensional density function as:

$$\hat{f}(\mathbf{y}) = (nh^d)^{-1} \sum_{i=1}^n K\left(\frac{\mathbf{y} - \mathbf{Y}_i}{h}\right). \quad (194)$$

The multivariate Gaussian kernel is given by $K(\mathbf{u}) = (2\pi)^{-d/2} \exp(-0.5\mathbf{u}^T\mathbf{u})$. The issues of kernel function and window width selections are similar to those discussed above for univariate density estimation. The approximately optimal window converges at the rate $n^{-1/(d+4)}$, and the optimal window width for the Gaussian kernel and a multivariate normal density $f(\mathbf{y})$ is given by $\frac{4}{n(2d+1)}^{1/(d+4)}$. Again, further details can be found in Silverman (1986).

14.1.5 Simulating from a nonparametric density estimate

Simulations from an estimated density are sometimes needed to compute estimates of functionals of the unknown true density function. This is the case, for instance, for the estimation of indices of classical horizontal inequity. The estimation of such indices requires information on the net income distribution of those who have the same gross incomes, and such information cannot be gathered directly from sample observations of net and gross incomes since very few (if any) exact equals can be observed in the ordinary random samples of finite sizes. Another use of simulated distributions is when one wishes to compute bootstrap estimates of the sampling distribution of some estimators. The usual bootstrap procedure proceeds by conducting successive random sampling (with replacement) from the original sample $\{\mathbf{Y}_i\}_{i=1}^n$. This constrains the new samples to contain only those observations \mathbf{Y}_i that were contained in the original sample. Those new samples could instead be generated from the density estimate $\hat{f}(\{\mathbf{y}\})$, which would yield a bootstrap estimate that would be smoother and less dependent on the precise values that the observations \mathbf{Y}_i took in the original sample.

Consider first the case of generating J independent realizations, $\{Y_j^*\}_{j=1}^J$, in a univariate case, supposing that a non-negative kernel with window width h is to be used to estimate $f(y)$. Assume that observation i has sampling weight w_i , and also assume for simplicity that these observations were drawn independently from each other. The following simple algorithm is adapted slightly from Silverman (1986, p.143). For $j = 1, \dots, J$, we do:

Step 1 Choose i with replacement from $\{k\}_{k=1}^n$ with probability $\{w_k\}_{k=1}^n$

Step 2 Choose ε to have probability density function K

Step 3 Set $Y_j^* = Y_i + h\varepsilon$.

Note that this algorithm does not even require computing directly $\hat{f}(y)$.

For the multivariate case, the above algorithm becomes just slightly more complicated. For instance, for the estimation of classical HI at gross income x , we need to generate a random sample of net incomes, $\{Y_j^*\}_{j=1}^J$, that follows the estimated kernel conditional density $\hat{f}(y|x)$. For this, we use the original sample $\{x_i, Y_i, w_i\}_{i=1}^n$ with sampling weights w_i . For $j = 1, \dots, J$, we then do:

Step 1 Choose i with replacement from $\{k\}_{k=1}^n$ with probability

$$\left\{ \frac{w_k K\left(\frac{x_k - x}{h}\right)}{\sum_{l=1}^n K\left(\frac{x_l - x}{h}\right)} \right\}_{k=1}^n$$

Step 2 Choose ε to have probability density function K

Step 3 Set $Y_j^* = Y_i + h\varepsilon$.

This gives a simulated sample of net incomes $\{Y_j^*\}_{j=1}^J$, conditional upon gross income being exactly equal to x . A local index of classical HI at x can then be computed using this simulated sample, and global indices of classical HI can be estimated simply by repeating this procedure at all of the observed values of gross incomes, $\{X_i\}_{i=1}^n$.

Because they follow an estimated density function that is on average smoother than the true one, the simulated samples generated by the above algorithms will have a variance that is generally larger than both the variance observed in the sample and the true population variance. In the univariate case, for instance, the variance of the simulated Y_j^* will equal $\sigma_Y^2 + h^2\sigma_Y^2$, where σ_Y^2 is the sample variance of the Y_i . This can be a problem if, as is the case for the measurement of indices of classical HI , the quantity of interest is intimately linked to the dispersion of income. There may also be a wish to constrain the simulated samples of net incomes to have precisely the same sample mean, \bar{Y} , as the original sample. Constraining the simulated samples to have the same mean and variance as in the original sample can be done by translating and re-scaling the simulated samples, replacing **Step 3** by

Step 3' Set $Y_j^* = \bar{Y} + \frac{Y_i - \bar{Y} + h\varepsilon}{(1+h^2\sigma_K^2/\sigma_Y^2)^{0.5}}$

in the univariate case. For the bivariate case, we also use **Step 3'**, but replace \bar{Y} by $E[y|x]$ and σ_Y^2 by $\sigma_{y|x}^2$, which can be respectively estimated as:

$$E[y|x] = \left(\sum_{i=1}^n w_i K \left(\frac{X_i - x}{h} \right) \right)^{-1} \sum_{i=1}^n w_i K \left(\frac{X_i - x}{h} \right) Y_i \quad (195)$$

and

$$\sigma_{y|x}^2 = \left(\sum_{i=1}^n w_i K \left(\frac{X_i - x}{h} \right) \right)^{-1} \sum_{i=1}^n w_i K \left(\frac{X_i - x}{h} \right) (Y_i - E[y|x])^2. \quad (196)$$

Equation (195) is in fact an example of a kernel regression of y on x , a procedure to which we now turn.

14.2 Non-parametric regression

The estimation of an expected relationship between variables is the second most important sphere of recent applications of kernel estimation techniques. Basically, one is interested in estimating the predicted response, $\mu y|x$, of a variable y at a given value of a (possibly multivariate) variable x , that is,

$$\mu y|x = E[y|x], \quad (197)$$

Alternatively, if the joint density exists and if $f(x) > 0$, $\mu y|x$ can also be defined as:

$$\mu y|x = \frac{\int y f(x, y) dy}{f(x)}. \quad (198)$$

The difficulty in estimating the relationship $\mu y|x$ is that we typically do not observe in a sample a response of y at that particular value of x . Furthermore, even when we do, there is rarely another observation with exactly the same value of x that would allow us to compute the average or expected response in which we are really interested.

Let then $\{X_i, Y_i\}_{i=1}^n$ be a sample of n observed realizations jointly of x and y . The response information that is provided by the sample can be expressed as:

$$Y_i = m(X_i) + \varepsilon_i, \text{ where } E[\varepsilon_i] = 0. \quad (199)$$

To estimate $\mu y|x$, kernel regression techniques use a local averaging procedure that involve weights $K(u)$ that are analogous to those used in section 14.1 for density estimation. Recalling (184) and (198), this leads to the following Nadaraya-Watson non-parametric estimator of $\mu y|x$:

$$\begin{aligned} \widehat{m}(x) &= (nh\widehat{f}(x))^{-1} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) Y_i \\ &= \frac{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) Y_i}{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)}. \end{aligned} \quad (200)$$

To reduce the bias of using neighboring Y_i , the kernel weights $K\left(\frac{x-X_i}{h}\right)$ are typically inversely proportional to the distance between x and X_i . They also depend on the window width h .

As for the kernel density estimators, the kernel smoother $\widehat{\mu y|x}$ can be shown to be consistent under relatively weak conditions, including that $\mu y|x$ and $f(x)$ are continuous functions of x , and that $h \rightarrow 0$ and $nh \rightarrow \infty$ as $n \rightarrow \infty$; see for instance Härdle (1990), Proposition 3.1.1. Again, the variance of $\widehat{\mu y|x}$ alone does not fully capture the convergence of $\widehat{\mu y|x}$ to $\mu y|x$ since we must also take into account the bias of $\widehat{\mu y|x}$, which comes from the smoothing of the Y_i in (200). Under suitable regularity conditions, including that $h \sim n^{-0.2}$, the asymptotic distribution of the kernel estimator $(nh)^{0.5}(\widehat{\mu y|x} - \mu y|x)$ can be shown to be normal, with its center shifted by its asymptotic bias¹³. This asymptotic bias is a function of the form of the kernel $K(u)$ and of the derivatives of $\mu y|x$ and $f(x)$. It is given by:

$$\sigma_K^2 \left(m^{(2)}(x) + 2m^{(1)}(x)f^{(1)}(x)/f(x) \right). \quad (201)$$

This asymptotic bias can be estimated consistently using estimates of $m^{(2)}(x)$, $m^{(1)}(x)$, $f^{(1)}(x)$ and $f(x)$. This, however, complicates significantly the computation of the sampling distribution of $\widehat{\mu y|x}$, and it can be avoided if we can expect (or make) the bias to be small compared to the variance. This will be the case if $\mu y|x$ is relatively constant, or if we make h fall just a bit faster than its optimal speed of $n^{-0.2}$; see the discussion of this in Härdle (1990), pp.100–102.

¹³See Härdle (1990), Theorem 4.2.1, for a demonstration.

The variance of $(nh)^{0.5}(\hat{\mu}_{y|x} - \mu_{y|x})$ is given by:

$$\sigma_{y|x}^2 c_K / f(x). \quad (202)$$

The conditional variance $\sigma_{y|x}^2$ can be estimated consistently as in (196). As for kernel density estimation, note again that the smoothing process makes the rate of convergence of kernel estimator $\hat{\mu}_{y|x}$ to be $n^{-0.4}$, instead of the usual slightly faster parametric convergence rate of $n^{-0.5}$.

15 Symbols

- y : living standards or income
- y^R : "real" income
- y_i : living standard of i^{th} ordered observation
- n : number of observations
- M : household size
- M_A : number of adults in household
- $E(\cdot)$: equivalence scale
- s : equivalence scale parameter
- c : equivalence scale parameter
- w_i : sampling weight of observation i
- p : percentile
- p_i : sample percentile of i^{th} ordered observation
- $F(y)$: distribution function
- $f(y)$: density function
- $Q(p)$: p -quantile
- y_{max} : maximum income
- μ : mean living standard
- z : poverty line
- $Q^*(p; z)$: p -quantile censored at z
- $g(p; z)$: poverty gap
- $\mu^*(z)$: mean of censored income
- $\mu^g(z)$: average poverty gap

- $L(p)$: Lorenz curve
- $GL(p)$: generalised Lorenz curve
- $G(p; z)$: Cumulative Poverty Gap (CPG) curve
- $\kappa(p)$: weight used in linear indices
- $\kappa(p; \rho)$: weight used in S-Gini indices
- $\omega(p)$: weight on income used in linear indices
- $\omega(p; \rho)$: weight on income used in S-Gini indices
- $I(\rho)$: S-Gini inequality indices
- $\delta(p, q)$: relative deprivation of $Q(p)$ in relation to $Q(q)$
- $\bar{\delta}(p)$: expected relative deprivation at p
- $\bar{\delta}^*(p; z)$: expected relative deprivation at p , in censored income
- $AD(z)$: absolute deprivation
- $RD(z; \rho)$: relative deprivation in censored income
- $U(Q(p))$: utility function of income
- ξ : equally distributed equivalent (EDE) living standard
- $\xi^*(z)$: EDE living standard of censored distribution
- $\xi^g(z)$: EDE poverty gap
- $\bar{\xi}^g(z; \alpha)$: EDE poverty gap for the normalized FGT indices
- $\xi^g(z; \alpha)$: EDE poverty gap for the un-normalized FGT indices
- $\Xi^*(z)$: cost of inequality in censored income
- $\Xi^g(z)$: cost of inequality in poverty gaps
- $I^*(z)$: index of inequality in censored income
- $W(\rho, \epsilon)$: Atkinson-Gini social welfare functions

- $I(\rho, \epsilon)$: Atkinson-Gini inequality indices
- $I(\theta)$: Generalised entropy inequality indices
- K : number of mutually exclusive population subgroups
- $\phi(k)$: share of the population found in subgroup k
- $\mu(k)$: mean living standard in subgroup k
- $I(k; \theta)$: inequality within subgroup k
- $\bar{I}(\theta)$: between-group inequality
- s : order of stochastic dominance
- $\pi(Q(p); z)$: contribution of $Q(p)$ to additive poverty index
- $P(z)$: additive poverty index
- $\Delta P(z)$: difference in poverty indices $P_A(z) - P_B(z)$
- $D^s(y)$: stochastic dominance curve of order s at y
- $\bar{D}^s(l\mu)$: normalised stochastic dominance curve of order s at $l\mu$
- $\Pi^s(z)$: class of primal (additive) poverty indices of the s -order
- $\dot{\Pi}^s(z)$: class of dual (linear) poverty indices of the s -order
- $\check{\Pi}^s(z)$: class of restricted primal poverty indices of the s -order
- $\overset{\circ}{\Pi}^s(z)$: class of restricted dual poverty indices of the s -order
- $\tilde{\Pi}^s(z)$: class of poverty indices of the s -order without continuity assumptions
- $\Pi^s(z_1, \dots, z_K)$: class of additive poverty indices of the s -order for a population with heterogeneous subgroups
- $\tilde{\Pi}^s(z_1, \dots, z_K)$: class of additive poverty indices of the s -order for a population with heterogeneous subgroups and without continuity assumptions

- $\ddot{\Pi}^{s_x, s_y}(z_x, z_y)$: class of bidimensional additive poverty indices of the s_x and s_y orders
- ζ : poverty line at which dominance must be checked
- λ : proportion of the mean at which dominance must be checked; also, scaling of a distribution.
- z^- : lower bound of range of poverty lines over which dominance must be checked
- $\zeta^-(s)$: lower bound of range of poverty lines over which dominance holds at order s
- z^+ : upper bound of range of poverty lines over which dominance must be checked
- $\zeta^+(s)$: lower bound of range of poverty lines over which dominance holds at order s
- l : proportion of mean as relative poverty line
- l^- : lower bound of range of proportions over which dominance must be checked
- $\lambda^-(s)$: lower bound of range of proportions over which dominance holds at order s
- l^+ : upper bound of range of proportions over which dominance must be checked
- $\lambda^+(s)$: upper bound of range of proportions over which dominance holds at order s
- p^- : lower bound of range of percentiles over which dominance must be checked
- $\psi^-(s)$: lower bound of range of percentiles over which dominance holds at order s
- p^+ : upper bound of range of percentiles over which dominance must be checked

- $\psi^+(s)$: upper bound of range of percentiles over which dominance holds at order s
- X : gross income
- $X(p)$: p -quantile of gross income
- N : net income
- $N(p)$: p -quantile of net income
- $T(X)$: deterministic portion of tax at X
- ν : stochastic tax determinant
- $F_{N|X=x}$: *cdf* of N conditional on $X = x$
- $N(q|p)$: q -quantile of N conditional on p -quantile of X
- $\bar{N}(p)$: expected net income of those at rank p in the distribution of gross incomes
- $\bar{T}(p)$: expected net tax of those at rank p in the distribution of gross incomes
- $C_N(p)$: concentration curve of N (ordered in terms of X)
- $GC_N(p)$: generalized concentration curve of N (ordered in terms of X)
- $IC_N(\rho)$: S-Gini concentration index of N (ordered in terms of X)
- μ_T : overall average tax
- t : overall average tax as a proportion of average gross income
- $t(X)$: expected tax at X as a proportion of X
- $LP(X)$: Liability Progression at X
- $RP(X)$: Residual Progression at X
- $P(z; \rho, \epsilon)$: Atkinson-Gini poverty index
- $P(z; \epsilon)$: Clark, Hemming and Ulph (CHU) poverty index
- $P(z; \rho)$: S-Gini poverty index

- $PW(z)$: Watts poverty index
- $PC(z; \epsilon)$: Chakravarty poverty index
- $P(z; \alpha)$: Foster-Greer-Thorbecke (FGT) poverty index
- $P(k; z; \alpha)$: FGT poverty index of subgroup k
- $\bar{P}(z; \alpha)$: normalised FGT poverty index
- $\varepsilon^G(z; \alpha)$: Gini elasticity of FGT indices
- $\varepsilon^y(z; \alpha)$: Income elasticity of FGT indices
- $x_c(y, q; \vartheta)$: consumption of good c at income y and with preferences ϑ , when facing prices q
- C : number of goods
- $x_c(y, q)$: expected consumption of good c for those at income y when facing prices q
- $x_c(q)$: average consumption of good c with prices q
- ϑ : taste parameter
- Θ : set of possible taste parameters
- $V(y, q, \vartheta)$: indirect utility function
- $e(q, \vartheta, v)$: expenditure function
- $R(y, q, q^R; \vartheta)$: real income function
- v_z : poverty level of utility
- q : price vector
- t : vector of tax rates
- q_c : price of good c
- t_c : commodity tax rate on good c

- q^R : reference prices
- ν : parameter of Cobb-Douglas utility function
- $\sigma_{y|x}^2$: variance of y conditional on some value x
- $\mu_{y|x}$: expected value of y conditional on some value x
- h : window width in non-parametric estimation
- $K(y)$: kernel function
- η : some positive value

16 References

References

- [1] Ahmad, E. and N.H. Stern (1984), "The Theory of Reform and Indian Indirect Taxes", *Journal of Public Economics*, **25**, 259–298.
- [2] Ahmad, E. and N.H. Stern (1991), *The Theory and Practice of Tax Reform in Developing Countries*, Cambridge University Press, Cambridge.
- [3] Araar, A. (1998), "Les mesures d'inégalité relative et les fonctions de bien-être social", ch. 3, in *Le bien-être des ménages et la transition économique en Pologne*, Ph.D. thesis, département d'économique, Université Laval, 49–68.
- [4] Araar, A. and J.-Y. Duclos (1999), "An Atkinson-Gini Family of Social Evaluation Functions" Working Paper 9826, CRÉFA, Département d'économique, Université Laval.
- [5] Anderson, G. (1996), "Nonparametric Tests of Stochastic Dominance In Income Distributions" *Econometrica*, **64 (5)**, 1183–1193.
- [6] Atkinson, A.B. (1995), *Incomes and the Welfare State*, Cambridge University Press, Cambridge.
- [7] Atkinson, A.B. (1991), "Measuring Poverty and Differences in Family Composition", *Economica*, **59 (233)**, 1–16.
- [8] Atkinson, A.B. (1987), "On the Measurement of Poverty", *Econometrica*, **55 (4)**, 749–764.
- [9] Atkinson, A.B. (1979), "Horizontal Equity and the Distribution of the Tax Burden", in H.J. Aaron and M.J. Boskin (eds), *The Economics of Taxation*, ch. 1, Brookings Institution, Washington DC, 3–18.
- [10] Atkinson, A.B. (1970), "On the Measurement of Inequality", *Journal of Economic Theory*, **2**, 244–263.
- [11] Atkinson, A.B. and F. Bourguignon (1987), "Income Distribution and Differences in Needs" in George R. Feiwel (ed), ch. 12 in *Arrow and the Foundations of the Theory of Economic Policy*, New York University Press, 350–370.

- [12] Atkinson, A.B. and F. Bourguignon (1982), “The Comparison of Multi-Dimensional Distributions of Economic Status”, ch. 2 in *Social Justice and Public Policy*, Harvester Wheatsheaf, London; and in *Review of Economic Studies*, 1983, **49**, (2), 183–201.
- [13] Bahadur, R.R. (1966), “A Note on Quantiles in Large Samples” *Annals of Mathematical Statistics*, **37**, 577–80.
- [14] Banks, J. and P. Johnson (1994), “Equivalence Scale Relativities Revisited”, *The Economic Journal*, **104** (425), 883–890.
- [15] Barrett, G.F., and K. Pendakur (1995), “The Asymptotic Distribution of the Generalized Gini Indices of Inequality”, *Canadian Journal of Economics*, **28** (4b), 1042–1055.
- [16] Beach, C. and R. Davidson, N. (1983), “Distribution-Free Statistical Inference with Lorenz Curves and Income Shares”, *Review of Economic Studies*, **50** (4), 723–35.
- [17] Beach, C. and S.F. Kaliski (1986), “Lorenz Curve Inference with Sample Weights: An Application to the Distribution of Unemployment Experience”, *Applied Statistics*, **35** (1), 38–45.
- [18] Beach, C. and J. Richmond (1985), “Joint Confidence Intervals for Income Shares and Lorenz ’”, *International Economic Review*, **26** (2), 439–50.
- [19] Besley, T. and R. Kanbur (1988), ”Food Subsidies and Poverty Alleviation”, *The Economic Journal*, **98**, 701–719.
- [20] Bishop, J.A., K.V. Chow and B. Zheng (1995), “Statistical Inference and Decomposable Poverty Measures” *Bulletin of Economic Research*, **47**, 329–340.
- [21] Bishop, J.A., J.P. Formby and B. Zheng (1997), “Statistical Inference and the Sen Index of Poverty” *International Economic Review*, **38**, 381–387.
- [22] Bishop, J.A., J.P. Formby and W.J. Smith (1991), “Lorenz Dominance and Welfare: Changes in the U.S. Distribution of Income, 1967-1986”, *Review of Economics and Statistics*, **73**, 134–140.

- [23] Bishop, J.A., J.P. Formby and P.D. Thistle (1992), “Convergence of the South and Non-South Income Distributions, 1969–1979”, *American Economic Review*, **82**, 262–72.
- [24] Bishop, J.A., S. Chakraborti and P.D. Thistle (1989), “Asymptotically Distribution-Free Statistical Inference for Generalized Lorenz Curves” *The Review of Economics and Statistics*, **LXXI**, 725–27.
- [25] Blackorby, C. and D. Donaldson (1978), “Measures of Relative Equality and Their Meaning in Terms of Social Welfare” *Journal of Economic Theory*, **18**, 59–80.
- [26] Blackorby, C. and D. Donaldson (1980), “Ethical Indices for the Measurement of Poverty” *Econometrica*, **48**, 1053–1062.
- [27] Bourguignon, F. and G. Fields (1997), “Discontinuous Losses from Poverty, Generalized Measures, and Optimal Transfers to the Poor” *Journal of Public Economics*, **63**, 155–175.
- [28] Buhmann, B., L. Rainwater, G. Schmaus and T. Smeeding (1988), “Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study Database”, *Review of Income and Wealth*, **34**, 15–142.
- [29] Chakravarty, S.R. (1983a), “A New Index of Poverty” *Mathematical Social Sciences*, **6**, 307–313.
- [30] Chakravarty, S.R. (1983b), “Ethically Flexible Measures of Poverty”, *Canadian Journal of Economics*, **XVI**, (1), 74–85.
- [31] Chakravarty, S.R. (1990), *Ethical Social Index Numbers*, New York, Springer-Verlag.
- [32] Chakravarty, S.R. (1997), “On Shorrocks’ Reinvestigation of the Sen Poverty Index” *Econometrica*, **65**, 1241–1242.
- [33] Clark, S., R. Hamming and D. Ulph (1981), “On Indices for the Measurement of Poverty”, *The Economic Journal*, **91**, 515–526.
- [34] Coulter, F.A.E., F.A. Cowell and S.P. Jenkins (1992), “Differences in Needs and Assessment of Income Distributions”, *Bulletin of Economic Research*, **44** (2), 77–124.

- [35] Coulter, F.A.E., F.A. Cowell and S.P. Jenkins (1992), “Equivalence Scale Relativities and the Measurement of Inequality and Poverty”, *The Economic Journal*, **102**, 1–16.
- [36] Cowell, Frank (1995), *Measuring Inequality*, LSE Handbook in Economics, Prentice Hall / Harvester Wheatsheaf.
- [37] Cowell, F.A. (1989), “Sampling Variance and Decomposable Inequality Measures”, *Journal of Econometrics*, **42**, 27–41.
- [38] Cowell, F.A., and M.P. Victoria-Feser (1996), “Robustness Properties of Inequality Measures”, *Econometrica*, **42**, 77–101.
- [39] Dalton, H. (1920), “The Measurement of the Inequality of Incomes”, *The Economic Journal*, **30 (119)**, 348–61.
- [40] Dasgupta, P., A. Sen and D. Starret (1973), “Notes on the Measurement of Inequality”, *Journal of Economic Theory*, **6 (2)** 180–187.
- [41] Datt, G. and M. Ravallion (1992), “Growth and Redistribution Components of Changes in Poverty Measures: a Decomposition with Applications to Brazil and India in the 1980’s”, *Journal of Development Economics*, **38**, 275–295.
- [42] Davidson, R. and J.Y. Duclos (1997), “Statistical Inference for the Measurement of the Incidence of Taxes and Transfers”, *Econometrica*, **65 (6)**, 1453–1465.
- [43] Davidson, R. and J.-Y. Duclos (2000), “Statistical Inference for Stochastic Dominance and the for the Measurement of Poverty and Inequality”, *Econometrica*, **68**, 1435–1465.
- [44] Davies, J. and M. Hoy (1995), “Making Inequality Comparisons When Lorenz Curves Intersect”, *American Economic Review*, **85 (4)**, 980–986.
- [45] Davies, J. and M. Hoy (1994), “The Normative Significance of Using Third-Degree Stochastic Dominance in Comparing Income Distributions”, *Journal of Economic Theory*, **64 (2)**, 520–530.
- [46] De Vos, Klaas, and M. Asghar Zaidi (1997), “Equivalence Scale Sensitivity of Poverty Statistics for the Member States of the European Community”, *Review of Income and Wealth*, **43 (3)**, 319–334.

- [47] Donaldson, D. and J.A. Weymark (1983), “Ethically Flexible Gini Indices for Income Distributions in the Continuum”, *Journal of Economic Theory*, **29 (2)**, 353–358.
- [48] Donaldson, D. and J.A. Weymark (1980), “A Single Parameter Generalization of the Gini Indices of Inequality”, *Journal of Economic Theory*, **22**, 67–86.
- [49] Duclos, J.-Y. (1999) “Gini Indices and the Redistribution of Income”, forthcoming in *International Tax and Public Finance*.
- [50] Duclos, J.-Y. (1997), “The Asymptotic Distribution of Linear Indices of Inequality, and Redistribution”, *Economics Letters*, **54 (1)**, 51–57.
- [51] Duclos, J.-Y. (1996), “Les tendances et les outils de la redistribution des revenus”, ch.9 in *La réinvention des institutions et le rôle de l'état*, actes du 21 ème congrès annuel de l'ASDEQ.
- [52] Duclos, J.-Y. (1995), “On Equity Aspects of Imperfect Poverty Relief”, *Review of Income and Wealth*, **41 (2)**, 177–190.
- [53] Duclos, J.-Y. (1993), “Progressivity, Redistribution and Equity, with Application to the British Tax and Benefit System”, *Public Finance /Finances Publiques*, **48 (3)** 350–365.
- [54] Duclos, J.-Y. and M. Mercader-Prats (1999), “Household Needs and Poverty: With Application to Spain and the UK”, *Review of Income and Wealth*, **45 (1)**, 77–98.
- [55] Duclos, J.-Y., and P.J. Lambert (1999), “A Normative and Statistical Approach to Measuring Classical Horizontal Equity”, forthcoming in *Canadian Journal of Economics*.
- [56] Duclos, J.-Y., and M. Tabi (1999), “Inégalité et redistribution du revenu, avec une application au Canada”, forthcoming in *Actualité Économique*.
- [57] Fields, G.S. (1994), “Data for Measuring Poverty and Inequality Changes in the Developing s”, *Journal of Development Economics*, **44**, 87–102.
- [58] Findlay, J., and R.E. Wright (1996), “Gender Poverty and the Intra-Household Distribution of Resources”, *Review of Income and Wealth*, **42 (3)**, 335–352.

- [59] Fishburn, P.C., and R.D. Willig (1984), "Transfer Principles in Income Redistribution" *Journal of Public Economics*, **25**, 323–328.
- [60] Formby, J.P., W.J. Smith and B. Zheng (1998), "Inequality Orderings, Stochastic Dominance and Statistical Inference" paper presented at the 1998 Chicago Meeting of the Econometric Society.
- [61] Foster, J.E. (1998), "What is Poverty and Who Are the Poor? Redefinition for the United States in the 1990's, *AEA Papers and Proceedings*, **88** (2) 335–341.
- [62] Foster, J.E., (1984), "On Economic Poverty: A Survey of Aggregate Measures" in R.L. Basmann and G.F. Rhodes, eds., *Advances in Econometrics*, **3**, Connecticut, JAI Press, 215–251.
- [63] Foster, J.E. and A. Sen (1997), "On Economic Inequality after a Quarter Century", in *On Economic Inequality* (expanded edition), Oxford, Clarendon Press.
- [64] Foster, J.E. and A.F. Shorrocks (1991), "Subgroup Consistent Poverty Indices", *Econometrica*, **59** (3), 687–709.
- [65] Foster, J.E. and A.F. Shorrocks (1988a), "Poverty Orderings", *Econometrica*, **56** (1), 173–177.
- [66] Foster, J.E. and A.F. Shorrocks (1988b), "Poverty Orderings and Welfare Dominance", *Social Choice Welfare*, **5**, 179–198.
- [67] Foster, J.E. and A.F. Shorrocks (1988c), "Inequality and Poverty Orderings", *European Economic Review*, **32**, 654–662.
- [68] Foster, J.E., J. Greer and E. Thorbecke (1984), "A Class of Decomposable Poverty Measures", *Econometrica*, **52** (3), 761–776.
- [69] Gottschalk, P. and T.M. Smeeding (1997), "Cross-National Comparisons of Earnings and Income Inequality", *Journal of Economic Literature*, **35**, June, 633–687.
- [70] Greer, J. and E. Thorbecke (1986), "A Methodology for Measuring Food Poverty Applied to Kenya", *Journal of Development Economics*, **24**, 59–74.

- [71] Gustafsson, B. and L. Nivorozhkina (1996), "Relative Poverty in Two Egalitarian Societies: A Comparison Between Taganrog, Russia During the Soviet Era and Sweden", *Review of Income and Wealth*, **42** (3), 321–334.
- [72] Haddad, L. and R. Kanbur (1990), "How Serious is the Neglect of Intra-Household Inequality?", *The Economic Journal*, **100**, 866–881.
- [73] Hagenaars, A. (1987), "A Class of Poverty Indices", *International Economic Review*, **28** (3), 583–607.
- [74] Hagenaars, A. and K. de Vos (1988), "The Definition and Measurement of Poverty", *The Journal of Human Resources*, **XXIII** (2), 211–221.
- [75] Hagenaars, A.J.M. and B.M.S. Van Praag (1985), "A Synthesis of Poverty Line Definitions", *The Review of Income and Wealth*, **31** (2), 139–153.
- [76] Härdle, W. (1990), *Applied Nonparametric Regression*, Cambridge University Press, Cambridge.
- [77] Hey, J.D. and P.J. Lambert (1980), "Relative Deprivation and the Gini Coefficient: Comment", *Quarterly Journal of Economics*, **95**, 567–573.
- [78] Howes, S. (1994), "Distributional Analysis Using Dominance Criteria: With Applications to Chinese Survey Data", STICERD, London School of Economics.
- [79] Howes, S. (1993), "Asymptotic Properties of Four Fundamental Curves of Distributional Analysis", STICERD, London School of Economics.
- [80] Howes, S., and J.O. Lanjouw (1998), "Does Sample Design Matter for Poverty Rate Comparisons?", *Review of Income and Wealth*, **44**, 99–109.
- [81] Jenkins, S.P. (1988), "Reranking and the Analysis of Income Redistribution", *Scottish Journal of Political Economy*, **35**, 65–76.
- [82] Jenkins, S.P. and F.A. Cowell (1994), "Parametric Equivalence Scales and Scale Relativities", *Economic Journal*, **104**, 891–900.
- [83] Jenkins, S.P. and P.J. Lambert (1998), "Three 'I's of Poverty Curves and Poverty Dominance: TIPs for Poverty Analysis", *Research on Economic Inequality*, **8**.

- [84] Jenkins, S.P. and P.J. Lambert (1997), "Three 'I's of Poverty Curves, With an Analysis of UK Poverty Trends", *Oxford Economic Papers*, **49**, 317–327.
- [85] Jenkins, S.P. and P.J. Lambert (1993), "Ranking Income Distributions when Needs Differ", *Review of Income and Wealth*, **39** (4), 337–356.
- [86] Kakwani, N.C. (1993), "Statistical Inference in the Measurement of Poverty", *Review of Economics and Statistics*, **75** (3), 632–639.
- [87] Kakwani, N.C. (1980), "On a Class of Poverty Measures" *Econometrica*, **48**, 437–446.
- [88] Kakwani, N.C. (1977), "Measurement of Tax Progressivity: An International Comparison", *Economic Journal*, **87**, 71–80.
- [89] King, M.A. (1983), "Welfare Analysis of Tax Reforms Using Household Data", *Journal of Public Economics*, **21**, 183–214.
- [90] Kodde, D.A. and F.C. Palm (1986), "Wald Criteria for Jointly Testing Equality and Inequality Restrictions" *Econometrica*, **54**, 1243–1248.
- [91] Kolm, S.C. (1970), "Unequal Inequalities, I" *Journal of Economic Theory*, **12**, 416–42.
- [92] Lambert, P. (1993), *The Distribution and Redistribution of Income: a Mathematical Analysis*, second edition, Manchester University Press, Manchester, 306p.
- [93] Lambert, P.J. and J.R. Aronson (1993), "Inequality Decomposition Analysis and the Gini Coefficient Revisited", *Economic Journal*, **103**, 1221–1227.
- [94] Lanjouw, P. and M. Ravallion (1995), "Poverty and Household Size", *The Economic Journal*, **105**, 1415–1434.
- [95] Mayshar, J. and S. Yitzhaki (1996), "Dalton-improving Tax Reform: When Households Differ in Ability and Needs", *Journal of Public Economics*, **62**, 399–412.
- [96] Mayshar, J. and S. Yitzhaki (1995), "Dalton-Improving Indirect Tax Reform", *American c Review*, September, **85** (4), 793–807.

- [97] Mehran, F. (1976), “Linear Measures of Income Inequality”, *Econometrica*, **44** (4), 805–809.
- [98] Mills, J.A. and S. Zandvakili (1997), “Statistical Inference Via Bootstrapping For Measures of Inequality”, *Journal of Applied Econometrics*, **12**, 133–150.
- [99] Muliere, P. and M. Scarsini (1989), “A Note on Stochastic Dominance and Inequality Measures”, *Journal of Economic Theory*, **49** (2) 314–323.
- [100] Newbery, D.M. (1995), “The Distributional Impact of Price Changes in Hungary and the United Kingdom”, *The Economic Journal*, **105**, 847–863.
- [101] O’Higgins, M., G. Schmaus and G. Stephenson (1989), “Income Distribution and Redistribution: A Microdata Analysis for Seven Countries”, *Review of Income and Wealth*, **35** (2), 107–131.
- [102] Okun, A.M. (1975), *Equality and Efficiency: The Big Tradeoff*, Brookings Institution, Washington.
- [103] Palme, M. (1996), “Income Distribution Effects of the Swedish 1991 Tax Reform: An Analysis of a Microsimulation Using Generalized Kakwani Decomposition”, *Journal of Policy Modelling*, **18**, 419–443.
- [104] Pfahler, W. (1987), “Redistributive Effects of Tax Progressivity: Evaluating a General Class of Aggregate Measures”, *Public Finance/ Finances Publiques*, **42** (3), 1–31.
- [105] Phipps, S.A. (1993), “Measuring Poverty Among Canadian Households, Sensitivity to Choice of Measure and Scale”, *Journal of Human Resources*, **28** (1), 162–184.
- [106] Plotnick, R. (1981), “A Measure of Horizontal Inequity”, *The Review of Economics and Statistics*, **LXII** (2), 283–288.
- [107] Preston, I. (1995), “Sampling Distributions of Relative Poverty Statistics” *Applied Statistics*, **44**, 91–99.
- [108] Rao, R.C. (1973), *Linear Statistical Inference and Its Applications*, John Wiley and Sons , New York, 625 pages.

- [109] Ravallion, M. (1996), “Issues in Measuring and Modelling Poverty”, *The Economic Journal*, **106**, 1328–1343.
- [110] Ravallion, M. (1994), *Poverty Comparisons*, Fundamentals of Pure and Applied Economics, Harwood Academic Publishers, Switzerland.
- [111] Ravallion, M. and B. Bidani (1994), “How Robust is a Poverty Profile?”, *The World Bank Economic Review*, **8** (1), 75–102.
- [112] Richmond, J. (1982), “A General Method for Constructing Simultaneous Confidence Intervals” *Journal of the American Statistical Association*, **77**, 455–460.
- [113] Rongve, I. (1997), “Statistical Inference for Poverty Indices with Fixed Poverty Lines”, *Applied Economics*, **29**, 387–392.
- [114] Rothschild, M. and J.E. Stiglitz (1973), “Some Further Results on the Measurement of Inequality”, *Journal of Economic Theory*, **6** (2), 188–204.
- [115] Runciman, W. G. (1966), *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth-Century England*, Berkeley and Los Angeles, University of California Press.
- [116] Sahn, D.E., Stephen Younger and Kenneth R. Simler (1996), “Dominance Testing of Transfers in Romania”, Cornell University, 1–16.
- [117] Sen, A.K. (1992), *Inequality Re-examined*, Clarendon Press, Oxford.
- [118] Sen, A.K. (1985), *Commodities and Capabilities*, North-Holland, Amsterdam.
- [119] Sen, A.K. (1984), “Poor, Relatively Speaking”, in *Resources, Values and Development*, Basil Blackwell, Oxford, 325–345.
- [120] Sen, A.K. (1981), *Poverty and Famine: An Essay on Entitlement and Deprivation*, Clarendon Press, Oxford University Press.
- [121] Sen, A.K. (1976), “Poverty: An Ordinal Approach to Measurement”, *Econometrica*, **44**, 219–231.
- [122] Sen, A.K. (1973), *On Economic Inequality*, Clarendon Press, Oxford.

- [123] Shorrocks, A.F. (1998), “Deprivation Profiles and Deprivation Indices” ch.11 in *The Distribution of Household Welfare and Household Production*, ed. S. Jenkins et al., Cambridge University Press.
- [124] Shorrocks, A.F. (1995), “Revisiting the Sen Poverty Index”, *Econometrica*, **63** (5), 1225–1230.
- [125] Shorrocks, A.F. (1987), “Transfer Sensitive Inequality Measures”, *Review of Economic Studies*, **LIV**, 485–497.
- [126] Shorrocks, A.F. (1984), “Inequality Decomposition by Population Subgroups”, *Econometrica*, **52** (6), 1369–1385.
- [127] Shorrocks, A.F. (1983), “Ranking Income Distributions”, *Economica*, **50**, 3–17.
- [128] Shorrocks, A.F. (1980), “The Class of Additively Decomposable Inequality Measures”, *Econometrica*, **48** (3), 613–625.
- [129] Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis*, London, Chapman and Hall.
- [130] Smeeding, T, L. Rainwater, and M. O’Higgins (1990). *Poverty, Inequality, and the Distribution of Income in an International Context: Initial Research from the Luxembourg Income Study (LIS)*, Wheatsheaf Books, London.
- [131] Spencer, B.D. and S. Fisher (1992), “On Comparing Distributions of Poverty Gaps”, *The Indian Journal of Statistics*, **54** (B), 114–126.
- [132] Summers, R. and A. Heston (1991), “The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988” *The Quarterly Journal of Economics*, **106**, 327–368.
- [133] Szulc, A. (1995), “Measurement of Poverty: Poland in the 1980’s”, *The Review of Income and Wealth*, **41** (2), 191–206.
- [134] Takayama, N. (1979), “Poverty, Income Inequality, and their Measures: Professor sen’s Axiomatic Approach Reconsidered” *Econometrica*, **47**, 747–759.

- [135] Thistle, P.D. (1990), “Large Sample Properties of two Inequality Indices”, *Econometrica*, **58** (2), March, 725–728.
- [136] Thon, D. (1979), “On Measuring Poverty”, *Review of Income and Wealth*, **25**, 429–439.
- [137] Van den Bosch, K., T. Callan, J. Estivill, P. Hausman, B. Jeandidier, R. Muffels, and J. Yfantopoulos (1993), “A Comparison of Poverty in Seven European Countries and Regions Using Subjective and Relative Measures” *Journal of Population Economics*, **6**, 235–259.
- [138] Watts, H.W. (1968), “An Economic Definition of Poverty”, in D.P. Moynihan (ed.), *On Understanding Poverty*, New York: Basic Books.
- [139] Wildasin, D. (1984), “On Public Good Provision With Distortionary Taxation”, *Economic Inquiry*, **22**, 227–243.
- [140] Wolak, F.A. (1989), “Testing Inequality Constraints in Linear Econometric Models” *Journal of Econometrics*, **41**, 205–235.
- [141] Xu, K. and L. Osberg (1998), “A Distribution-Free Test for Deprivation Dominance” Department of Economics, Dalhousie University, Halifax.
- [142] Yitzhaki, S., (1983), “On an Extension of the Gini Index”, *International Economic Review*, **24**, 617–628.
- [143] Yitzhaki, S. (1979), “Relative Deprivation and the Gini Coefficient”, *Quarterly Journal of Economics*, **93**, 321–324.
- [144] Yitzhaki, S. and J. Lewis (1996), “Guidelines on Searching for a Dalton-Improving Tax Reform: An Illustration with Data from Indonesia”, *The World Bank Economic Review*, **10** (3), 1–562.
- [145] Yitzhaki, S. and J. Slemrod (1991), “Welfare Dominance: An Application to Commodity Taxation”, *American Economic Review*, **LXXXI**, 480–496.
- [146] Yitzhaki, S. and W. Thirsk (1990), “Welfare Dominance and the Design of Excise Taxation in the Cote D’Ivoire”, *Journal of Development Economics*, **33**, 1–18.
- [147] Zheng, B. (1997a), “Aggregate Poverty Measures”, *Journal of Economic Surveys*, **11**, 123–63.

- [148] Zheng, B. (1997b), “Statistical Inferences for Poverty Measures with Relative Poverty Lines” mimeo, Department of Economics, University of Colorado.
- [149] Zheng, Buhong (1999), “On the Power of Poverty Orderings”, forthcoming in *Social Choice and Welfare*.

17 Graphs and tables

Figure 1: Capabilities, achievements and consumption

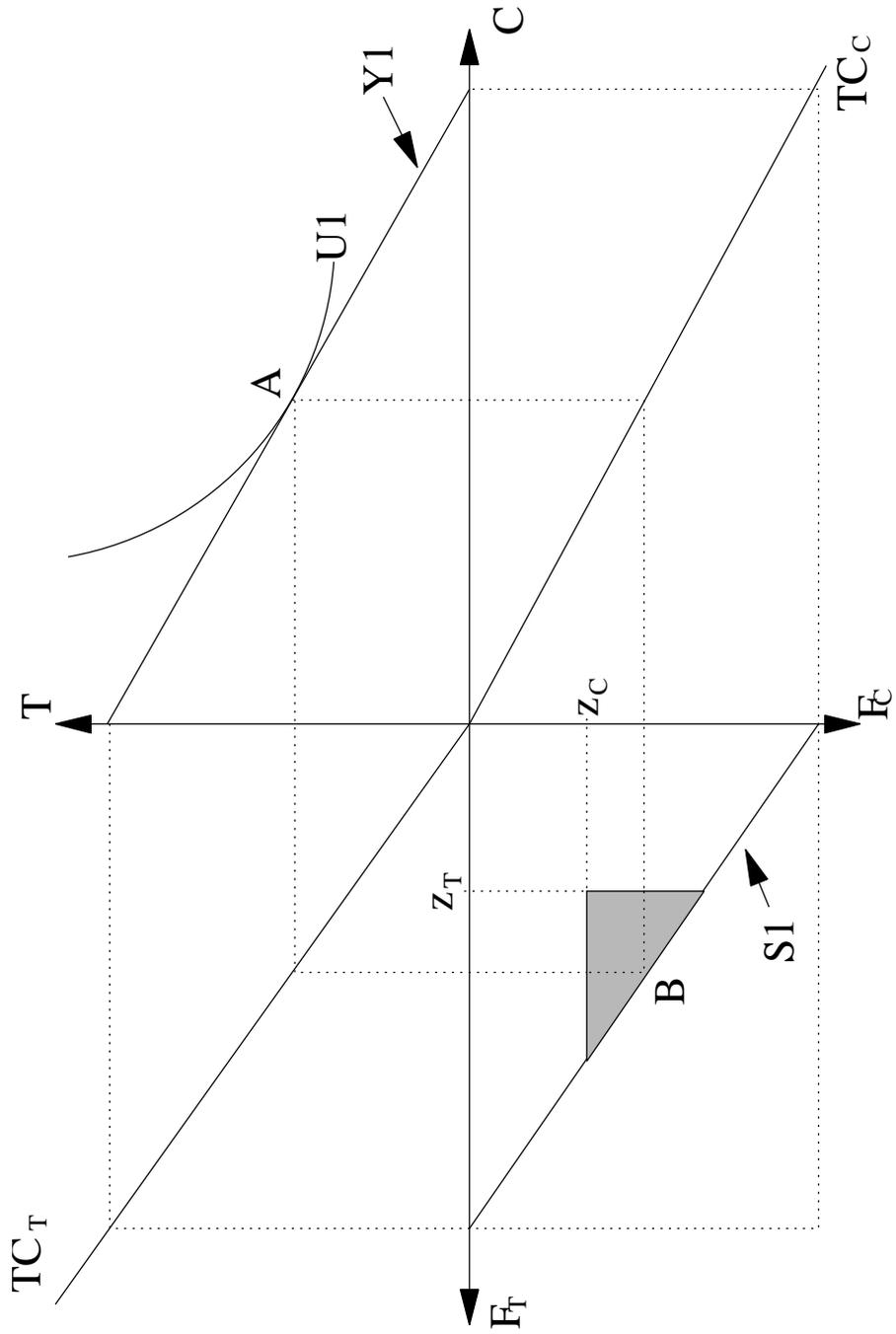


Figure 2: Capabilities and achievements under varying preferences

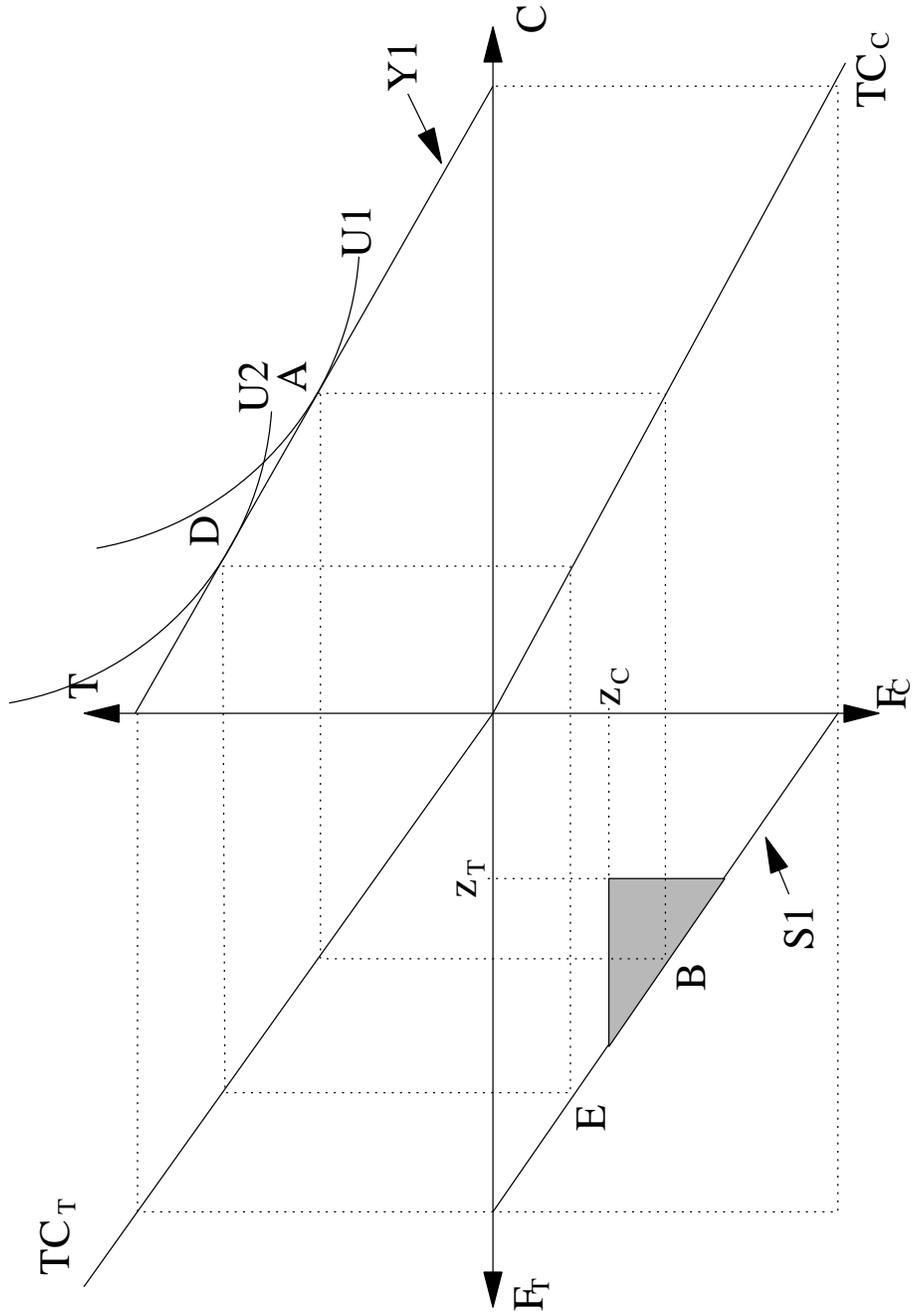


Figure 3: Capability set and achievement failure

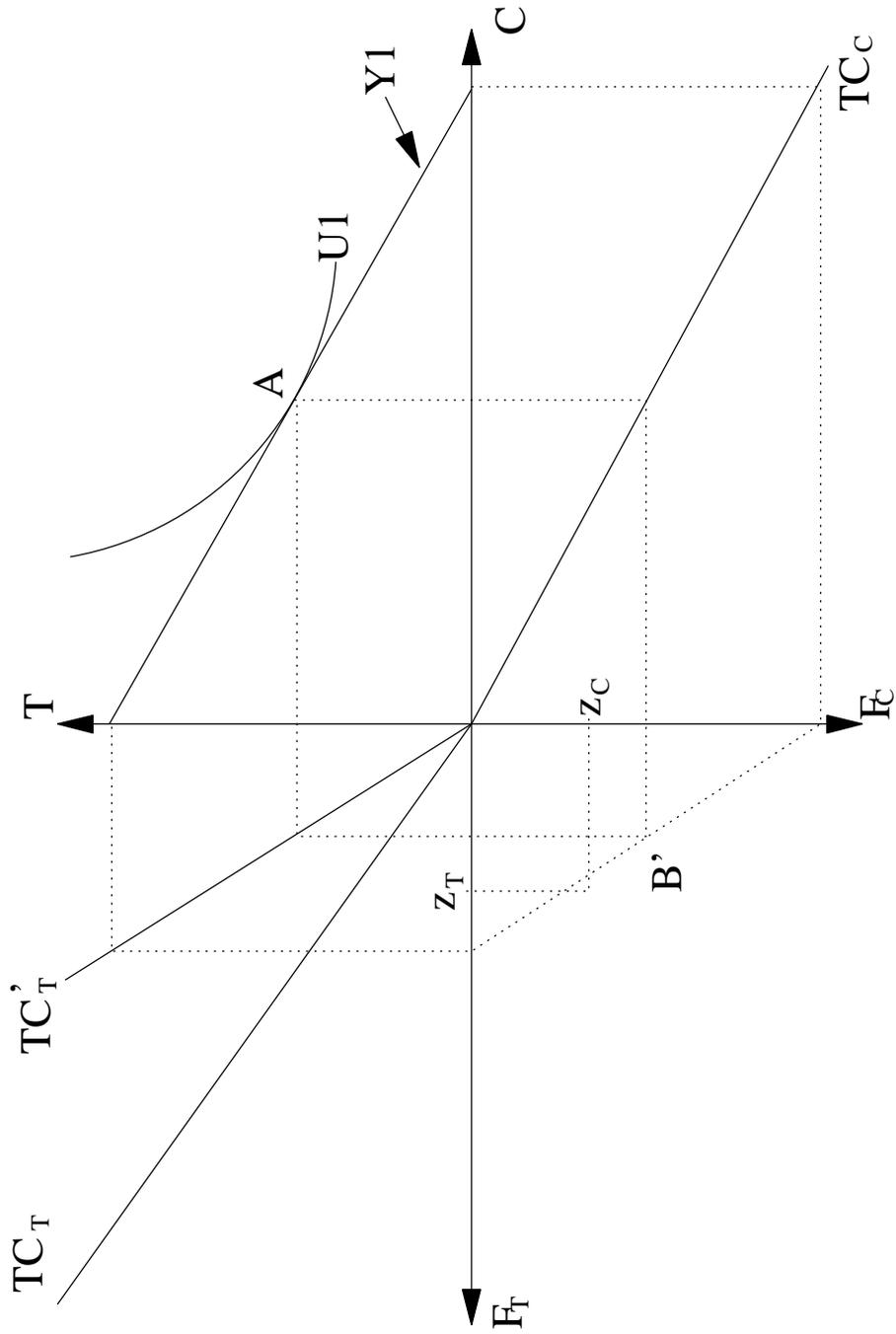


Figure 4: Minimum consumption needed to escape capability poverty

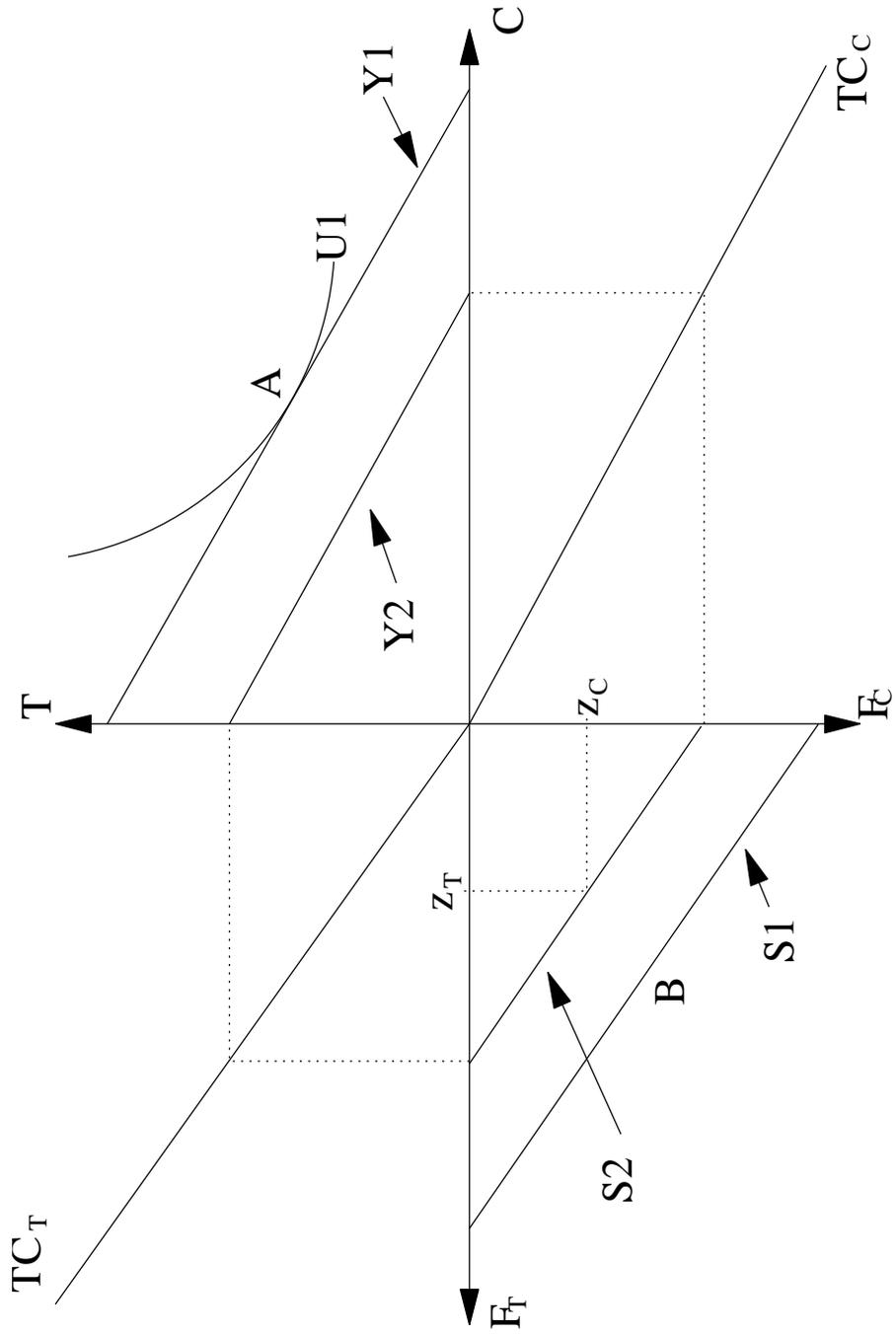


Figure 5: The cumulative poverty gap (CPG) curve

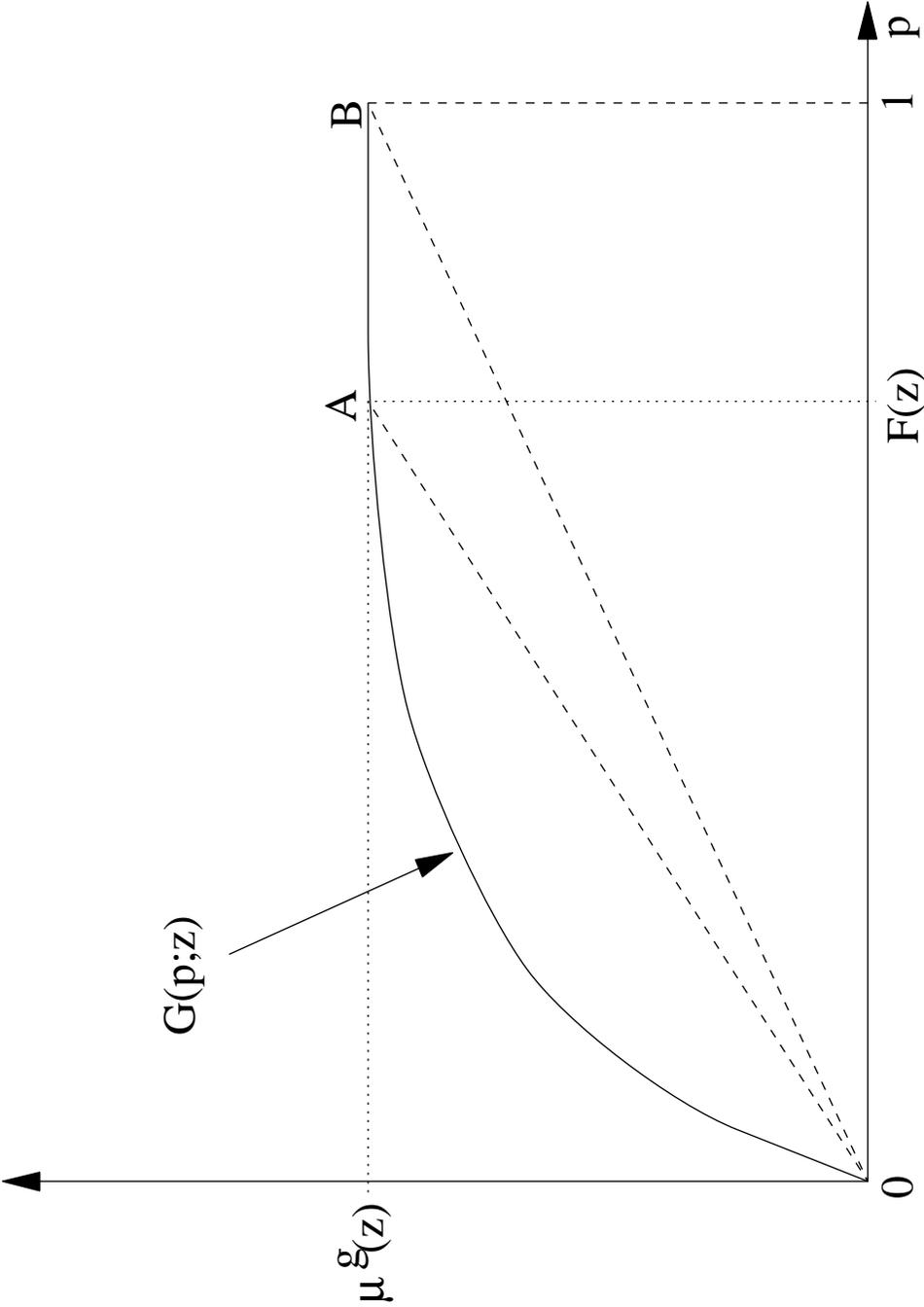


Figure 6: Engel curves and Cost-of-basic-needs baskets

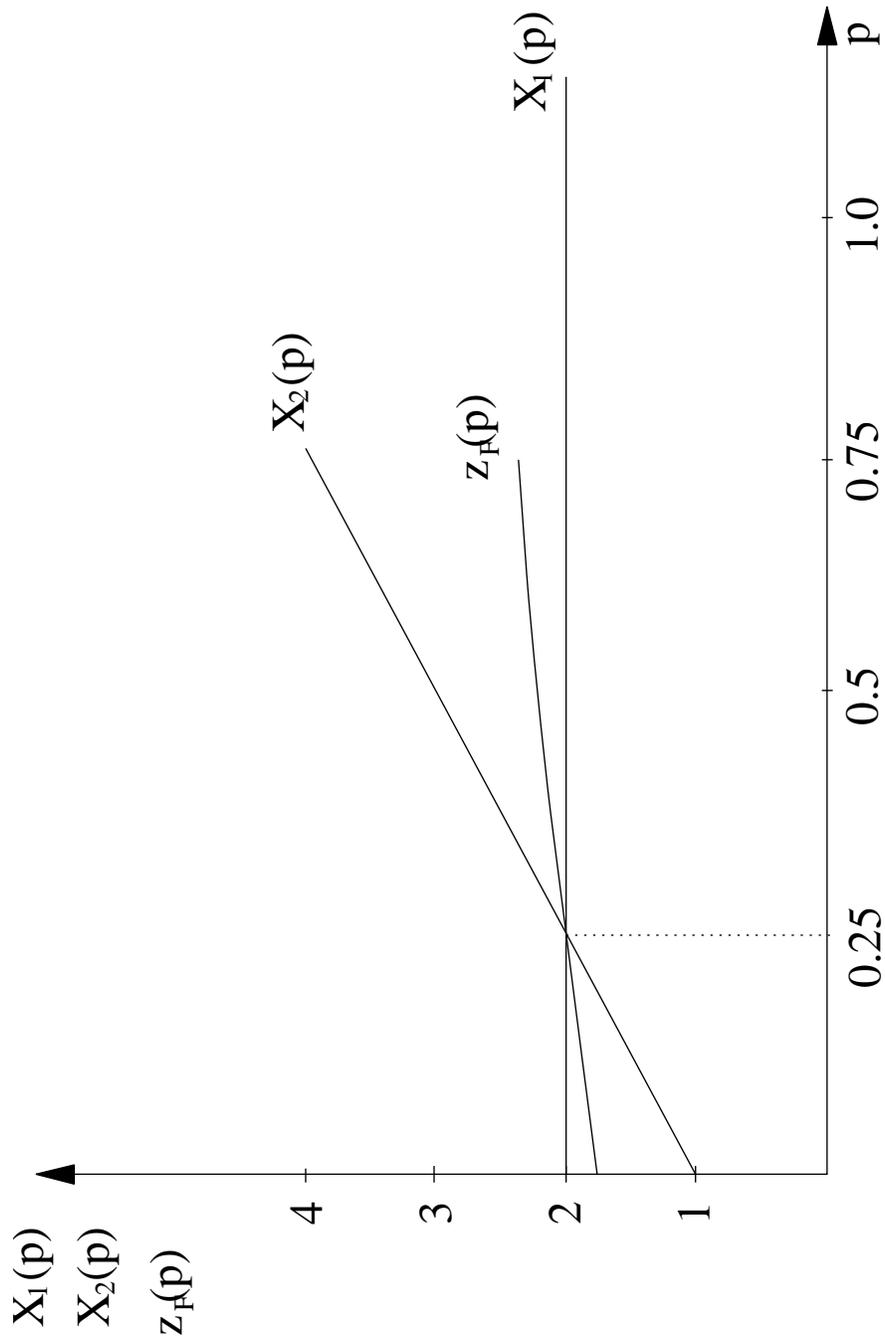


Figure 7: Food preferences and cost of minimum calorie intake

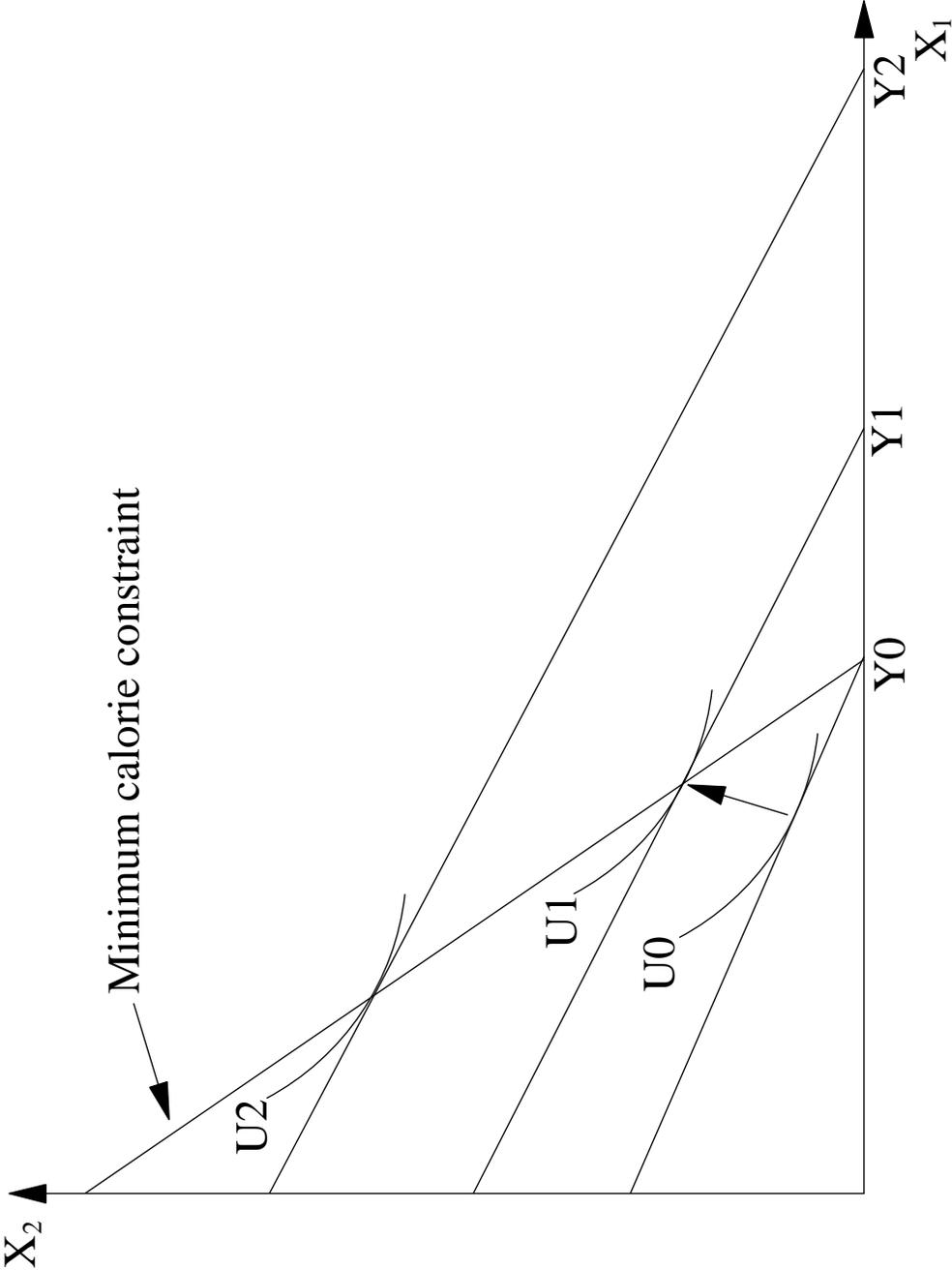


Figure 8: Non-food and total poverty lines

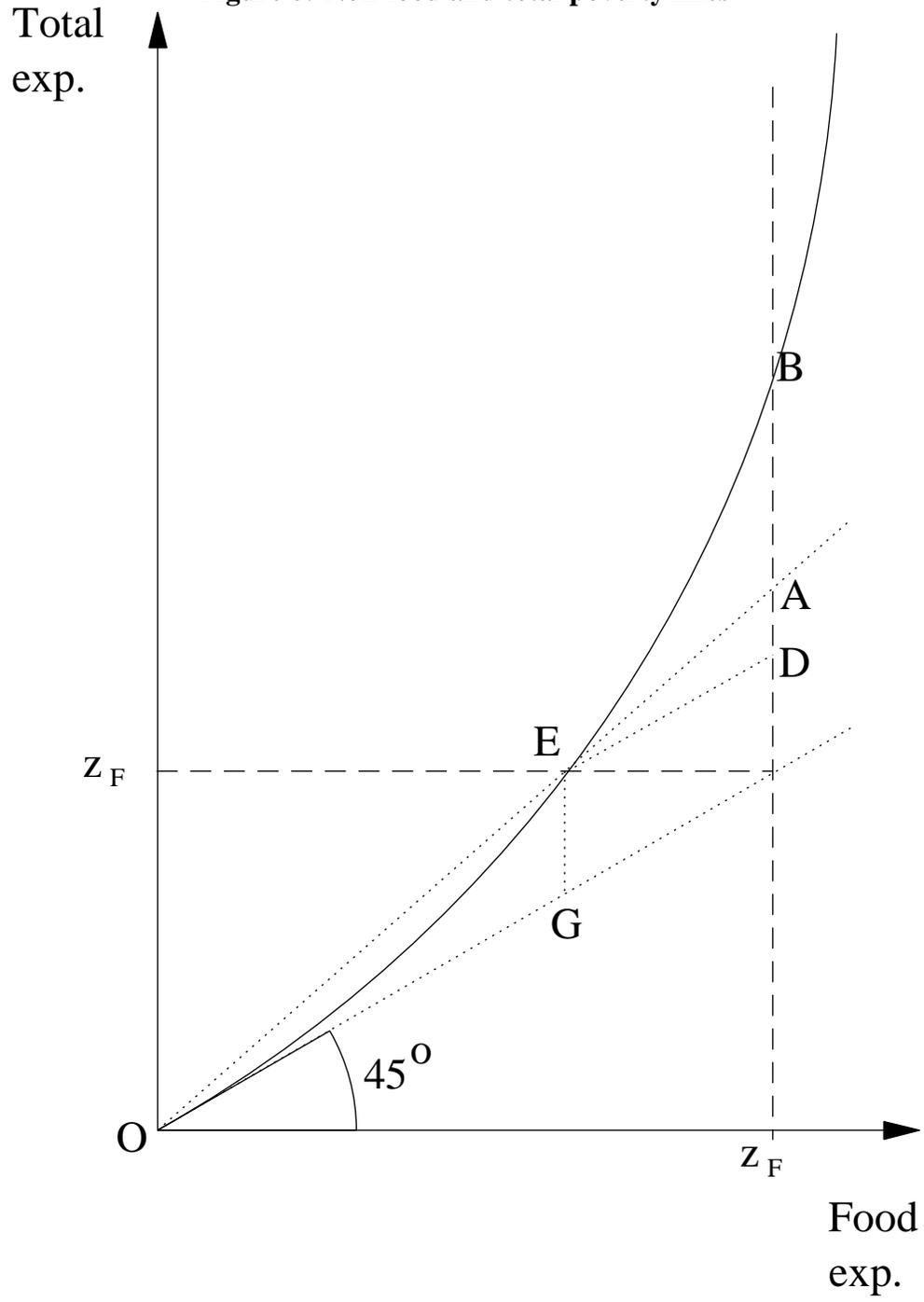


Figure 9: Expenditure and calorie intake

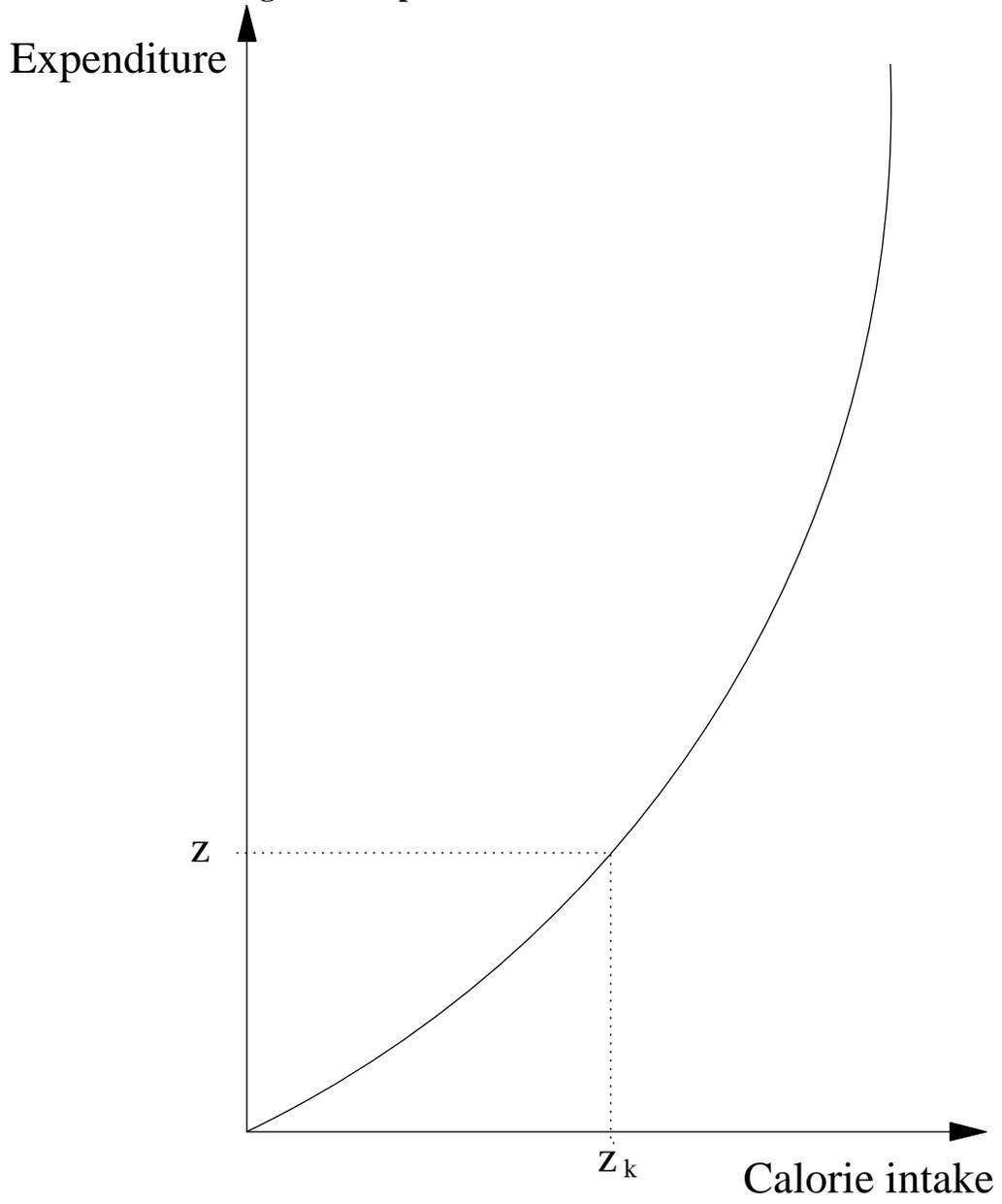


Figure 10: Subjective poverty lines

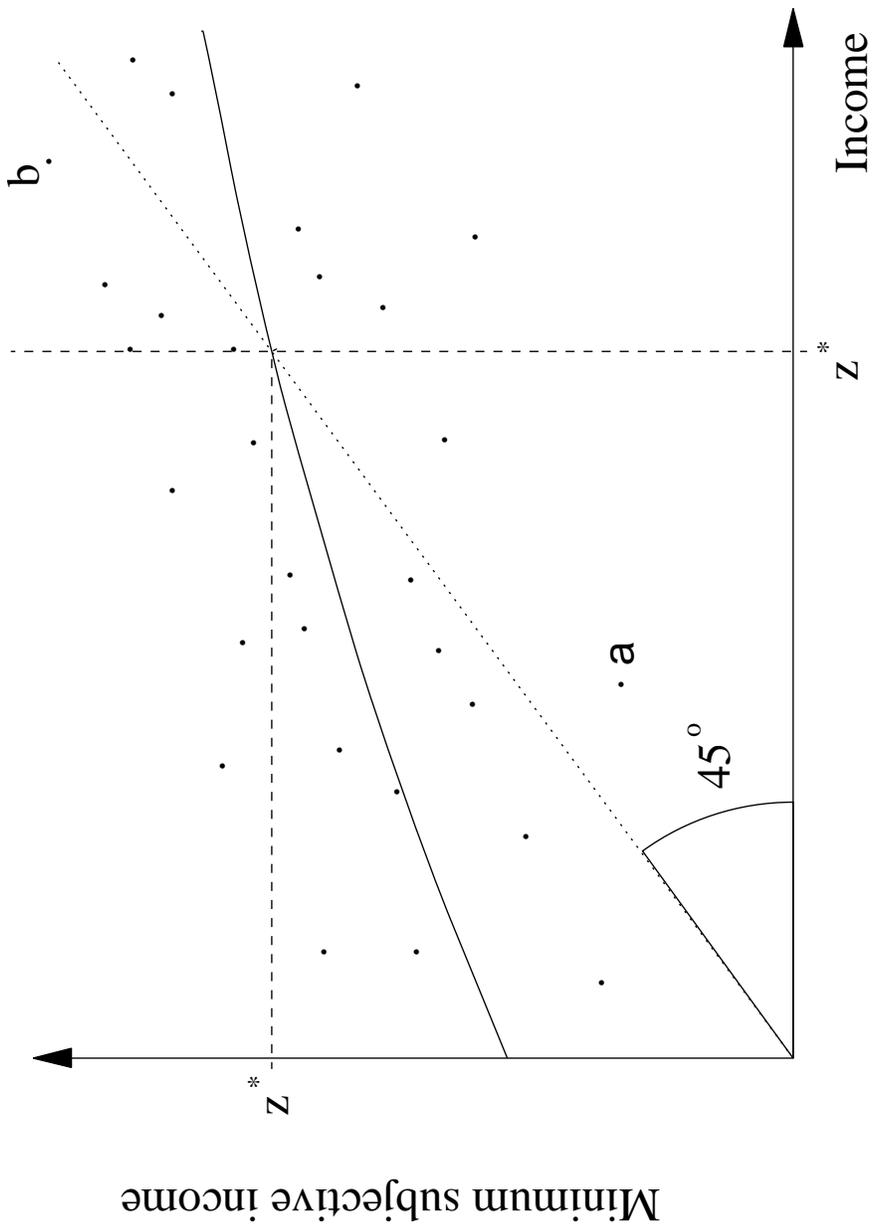


Figure 11: Estimating a subjective poverty line with discrete subjective information

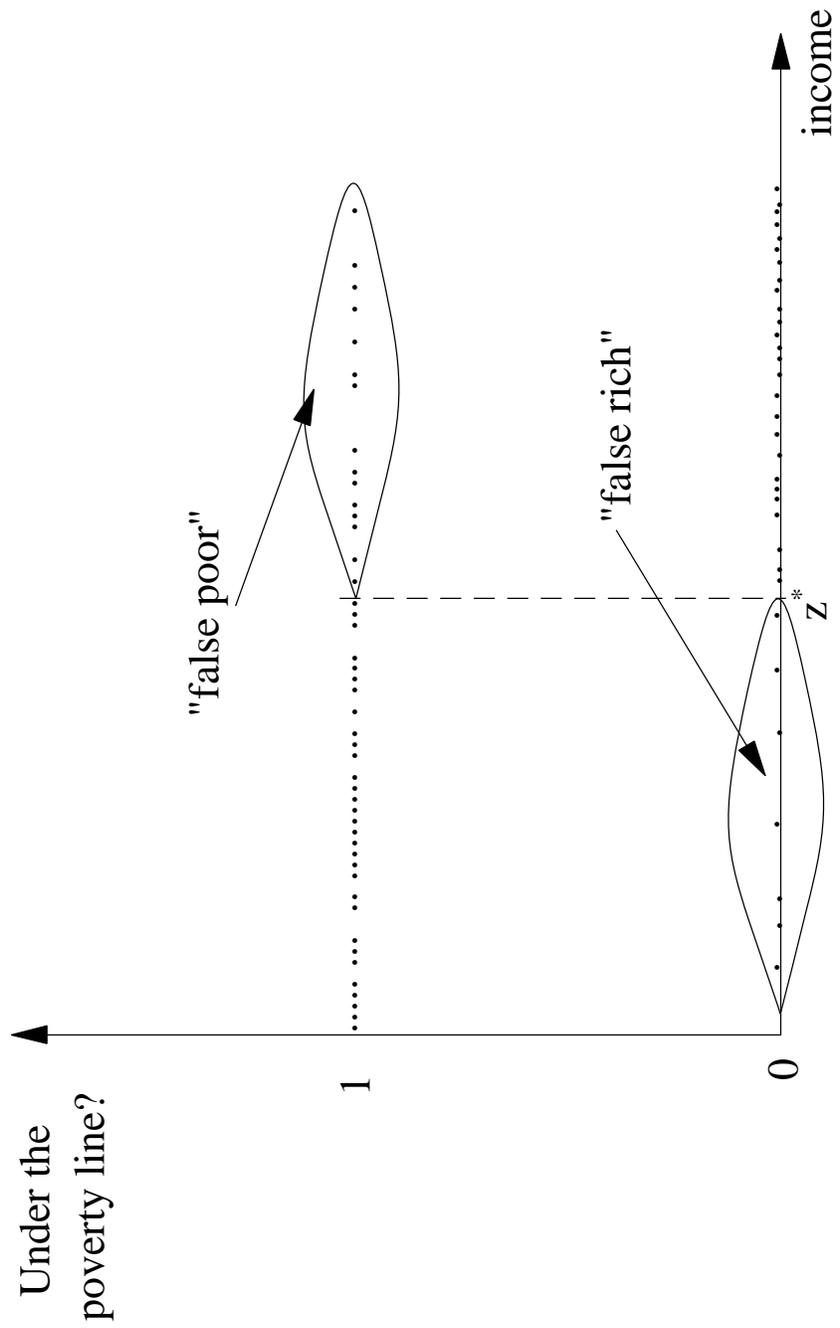


Figure 12: s -order poverty dominance

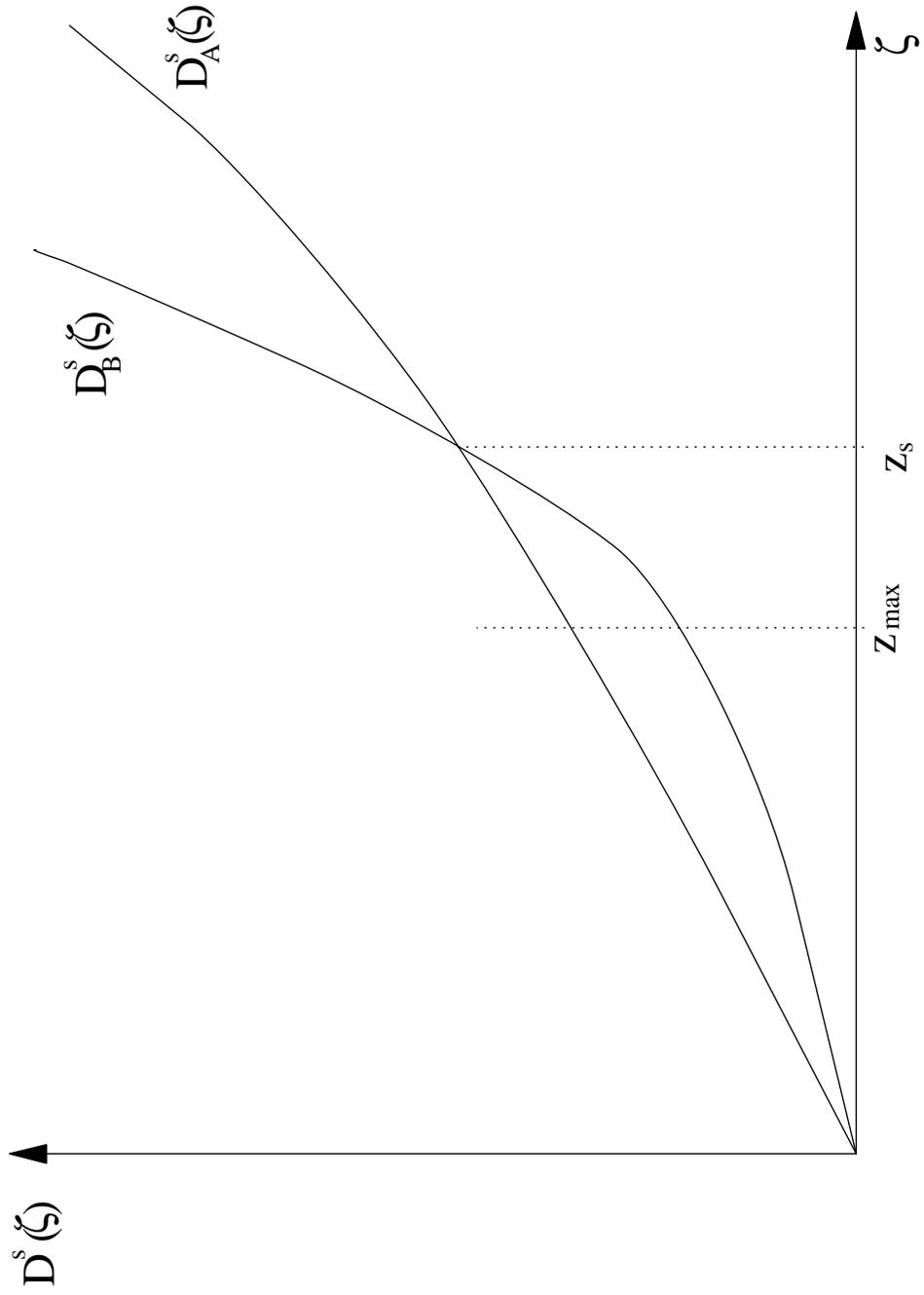


Figure 13: Quantile curve for discrete distribution

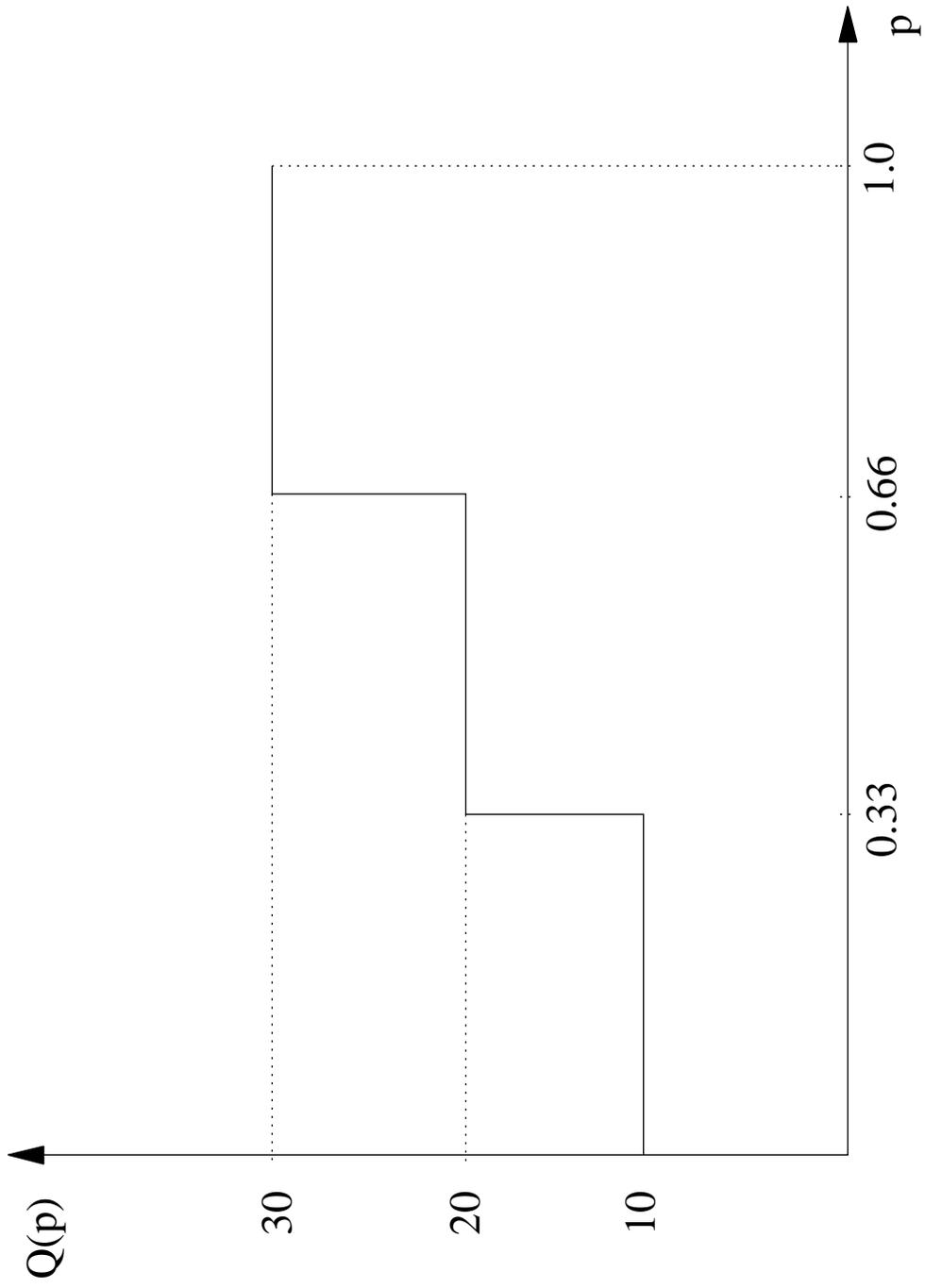


Figure 14: Quantile curve for continuous distribution

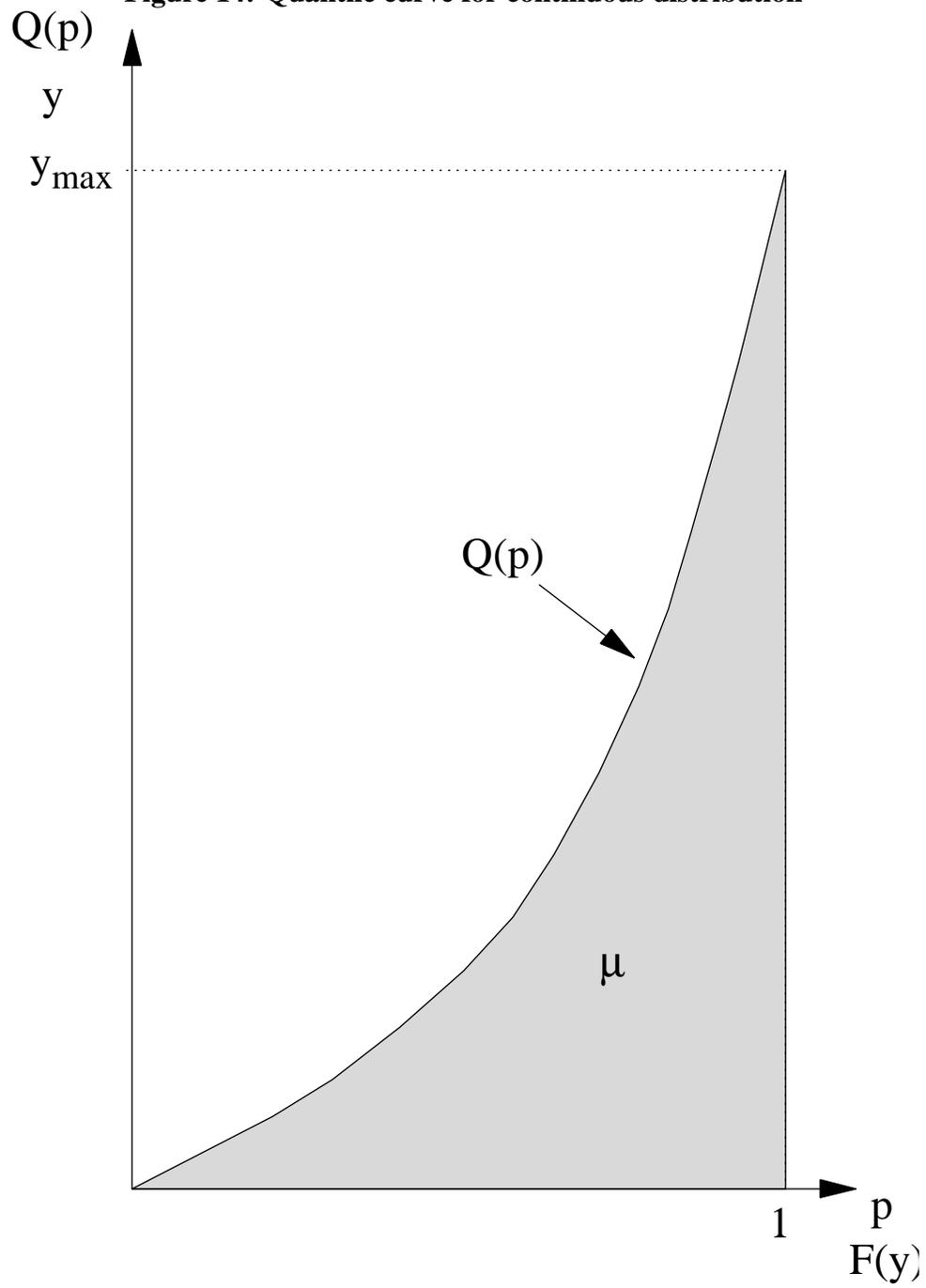


Figure 15: Living standards and poverty at different percentiles

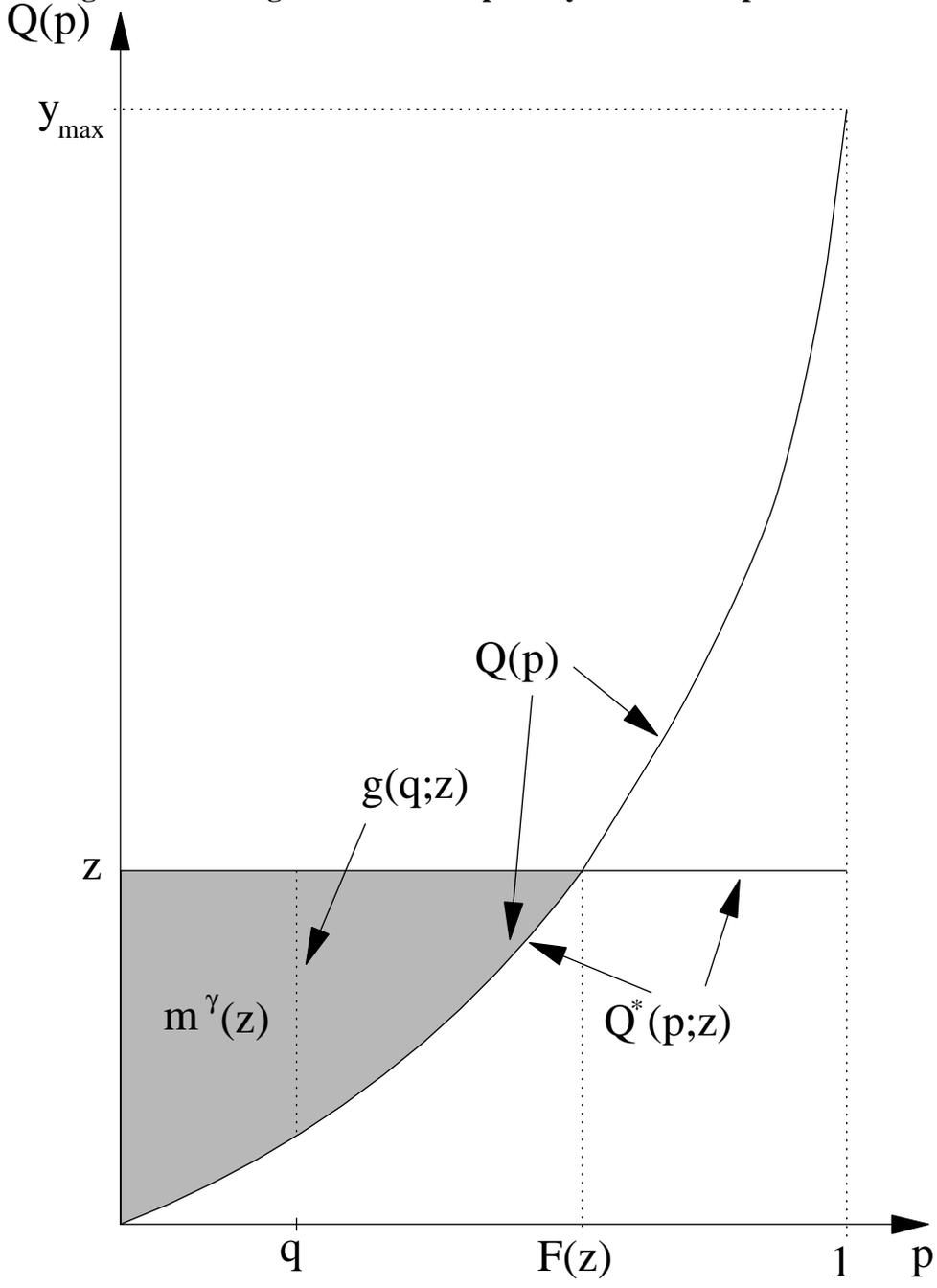


Figure 16: Price adjustments and well-being

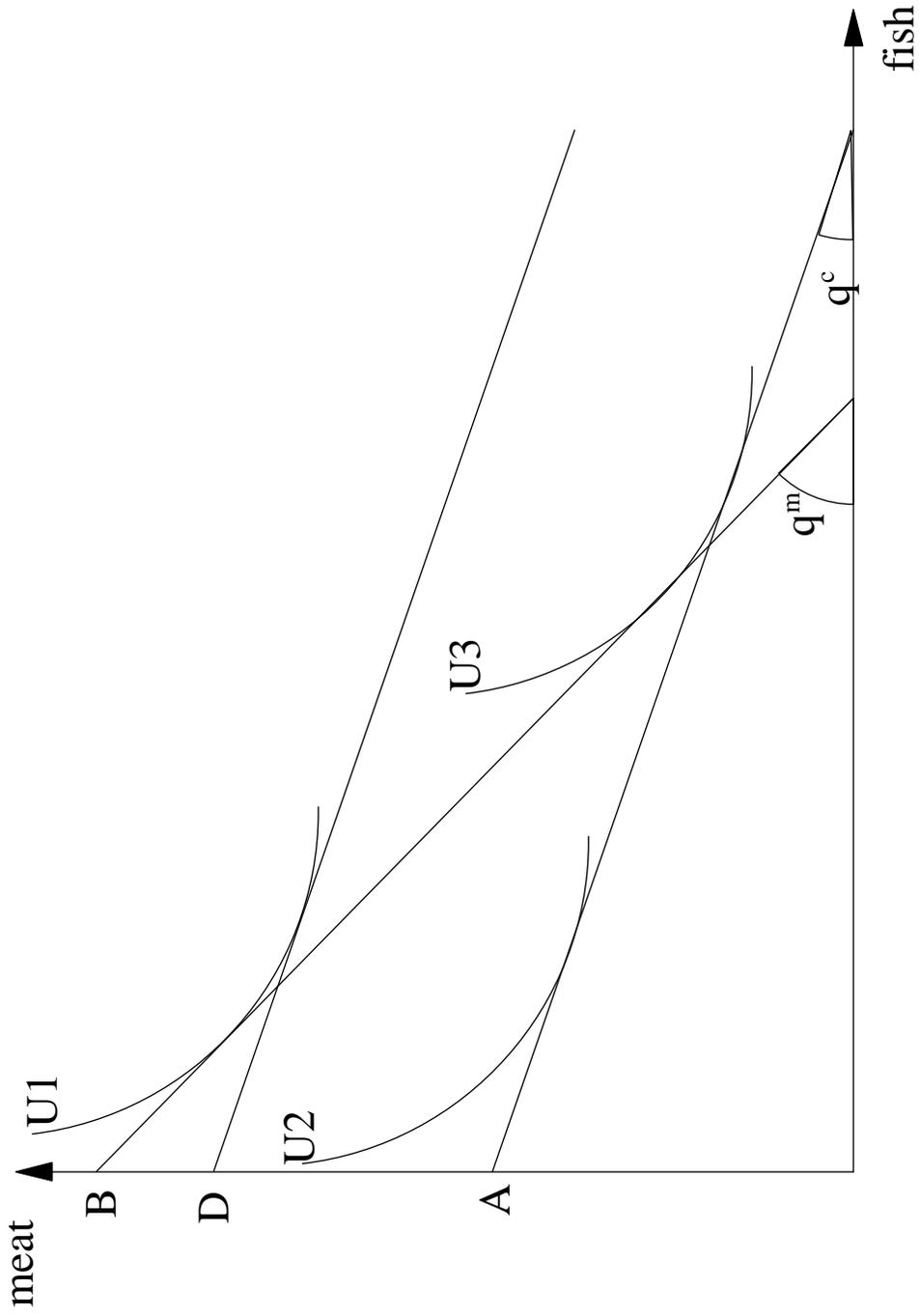


Figure 17: Equivalence scales and reference well-being

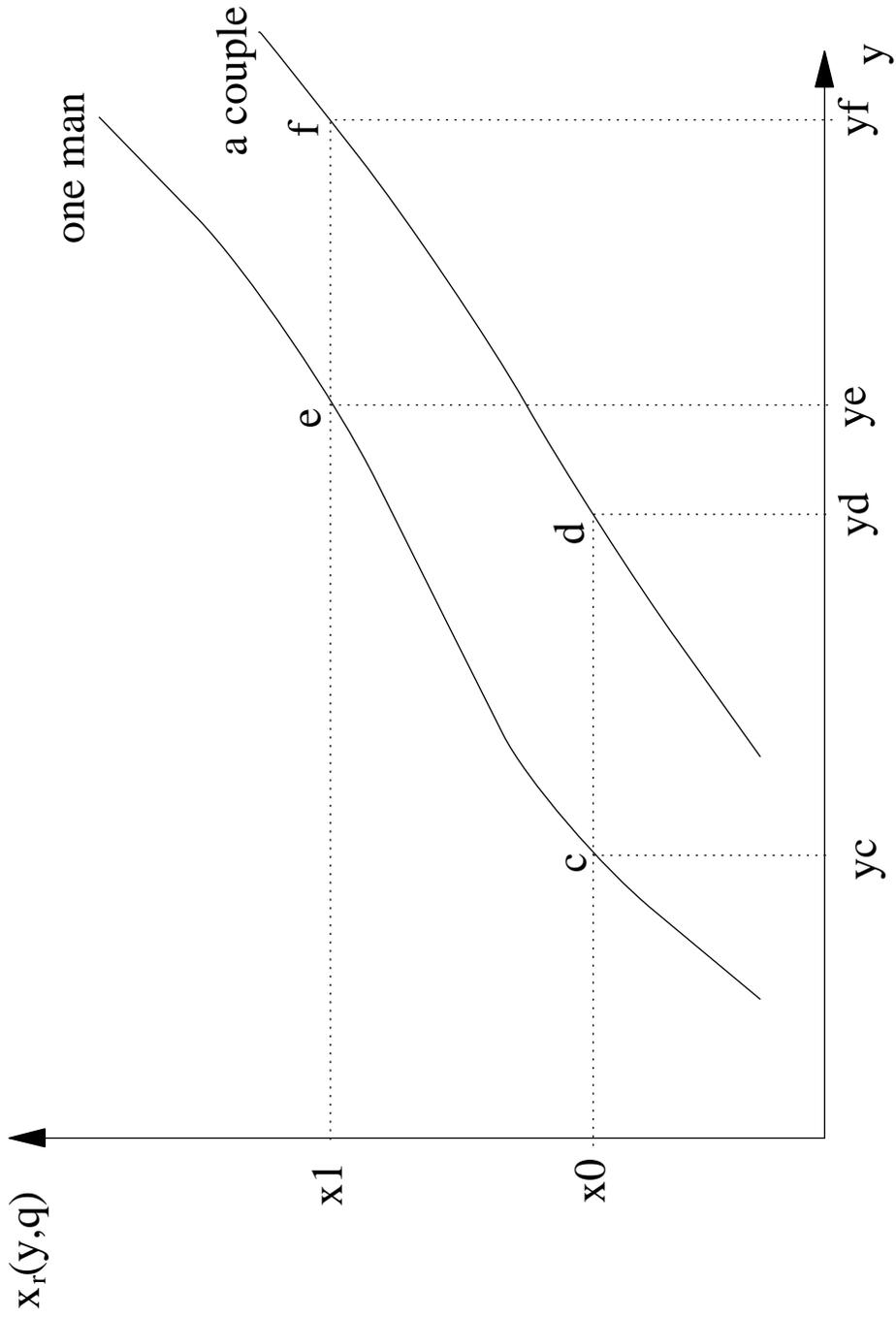


Figure 18: Atkinson social evaluation functions and the cost of inequality

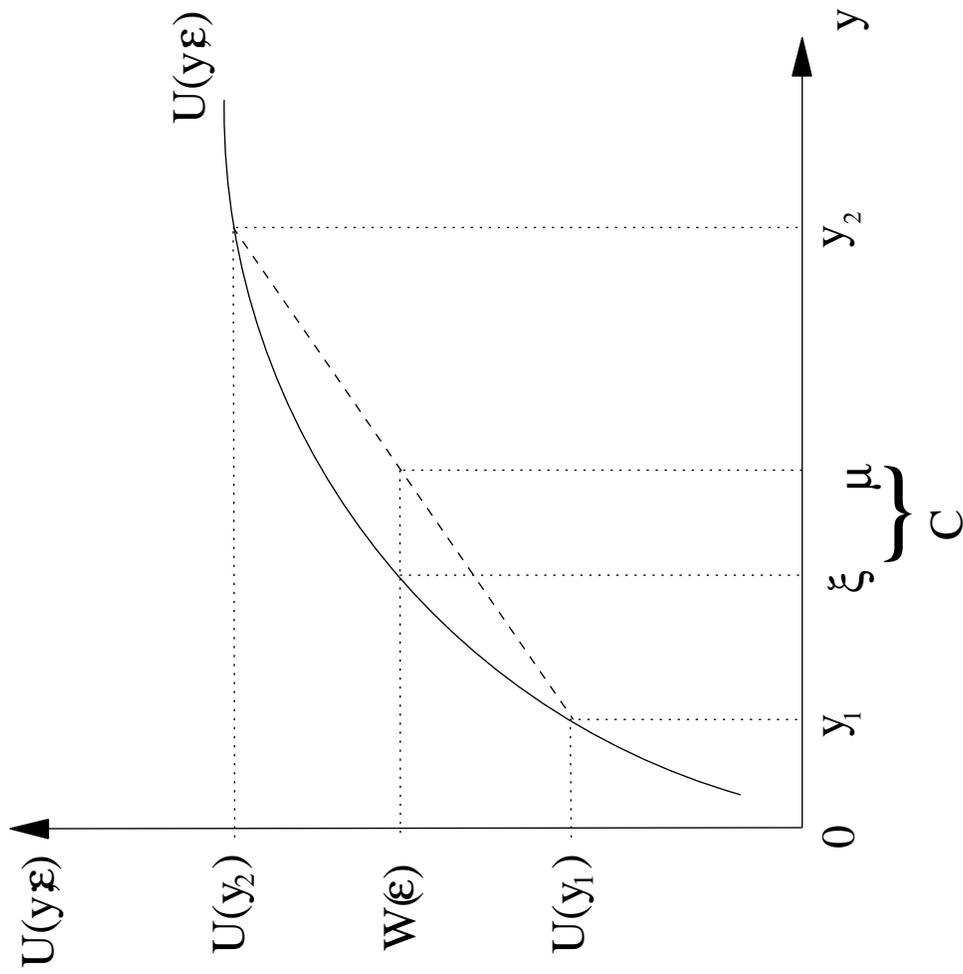


Figure 19: Inequality aversion and the cost of inequality

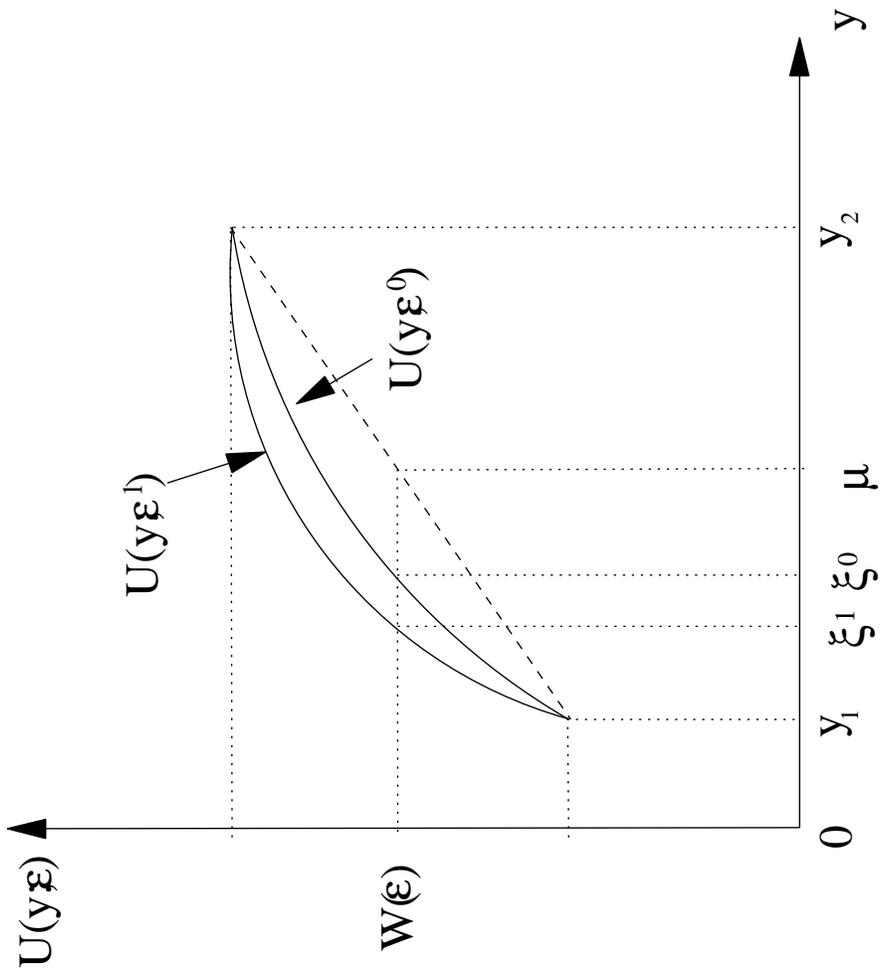


Figure 20: Homothetic social evaluation functions

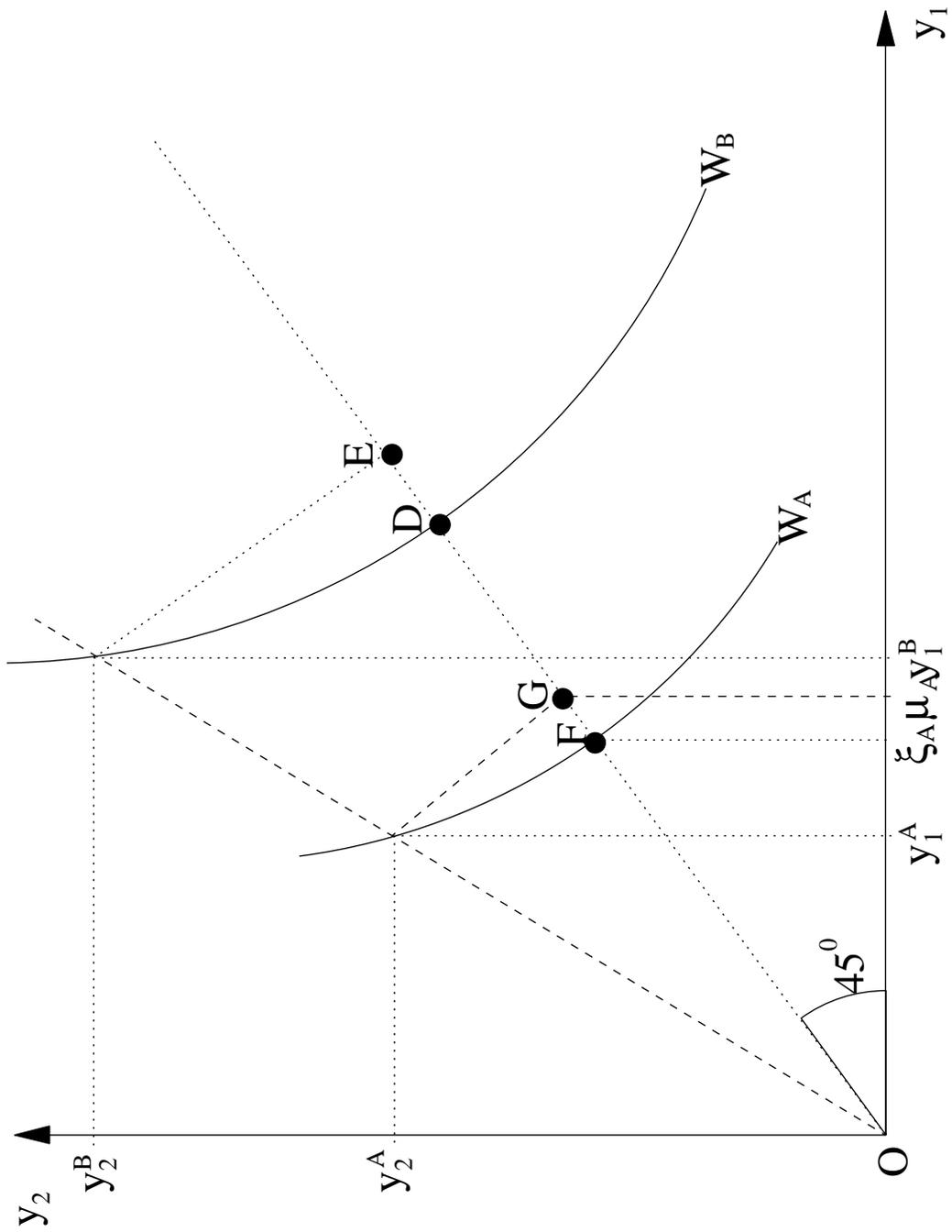


Figure 21: Social utility and living standards

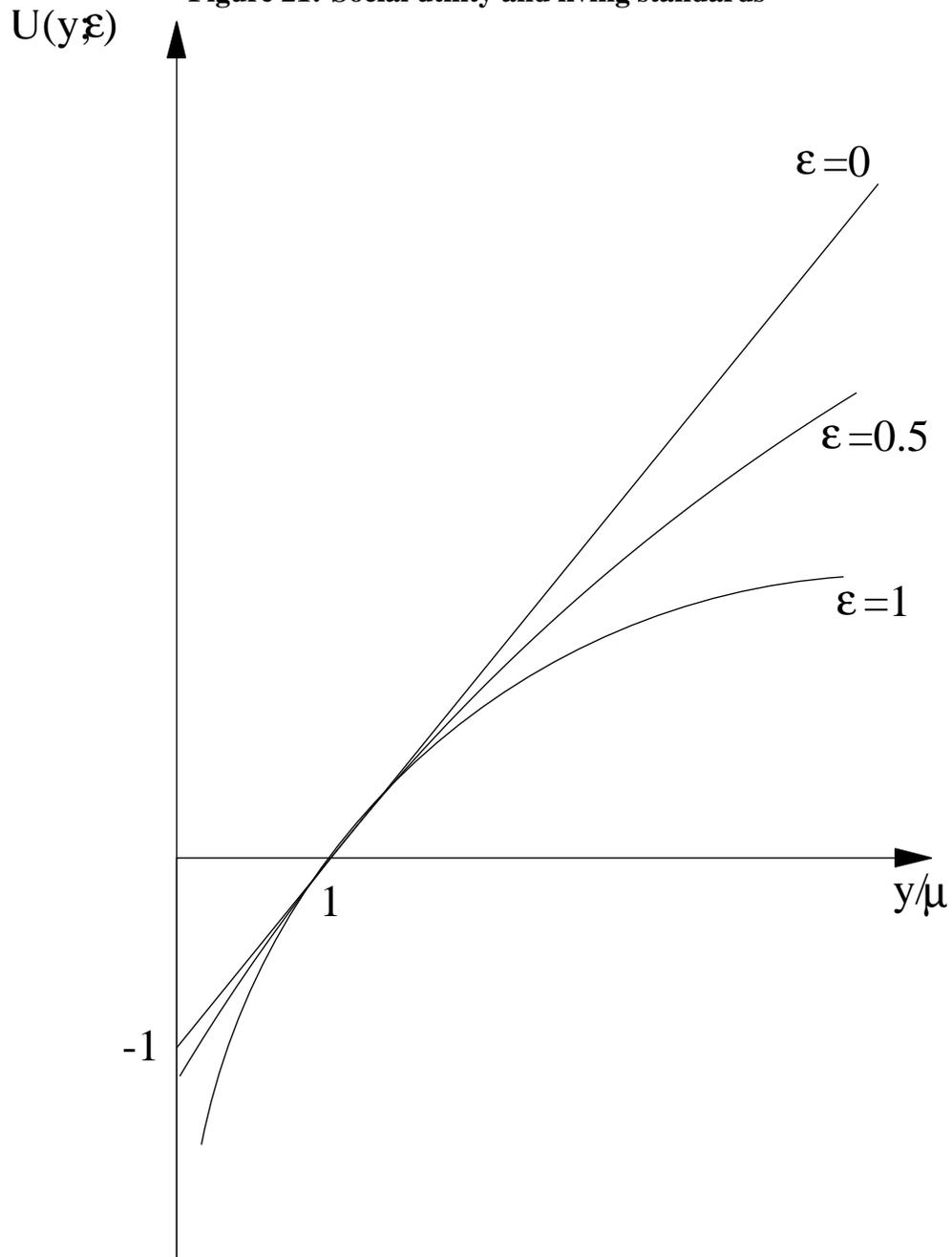


Figure 22: Marginal social utility and living standards

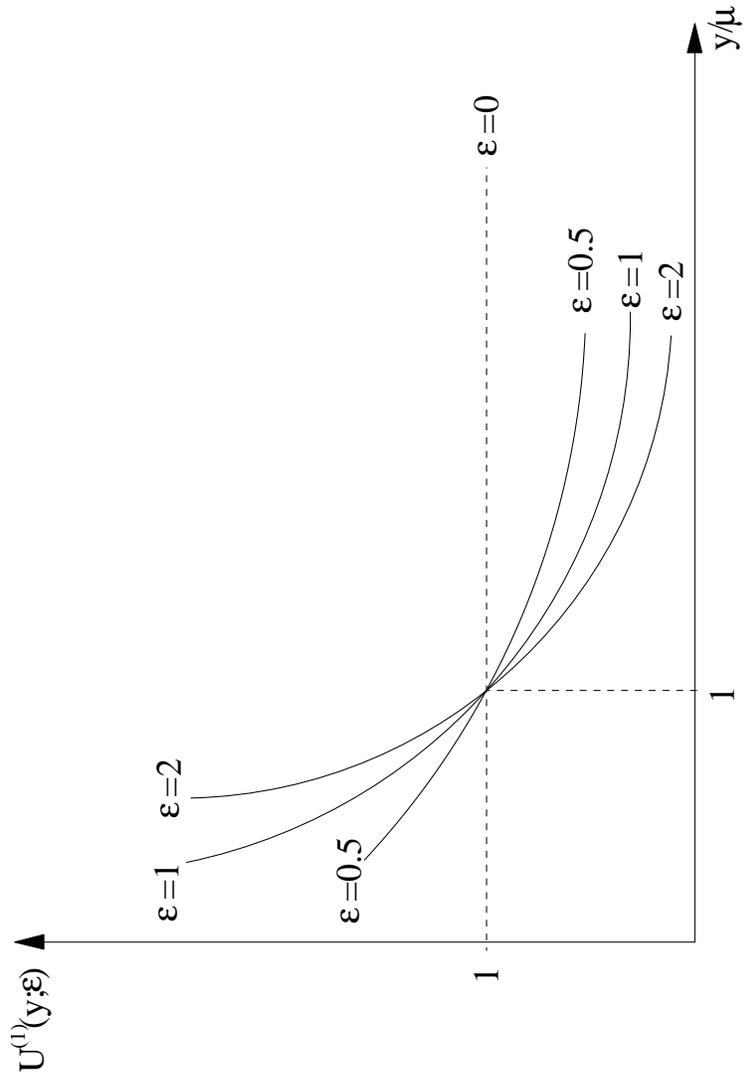


Figure 23: Mean living standard and inequality for constant social welfare

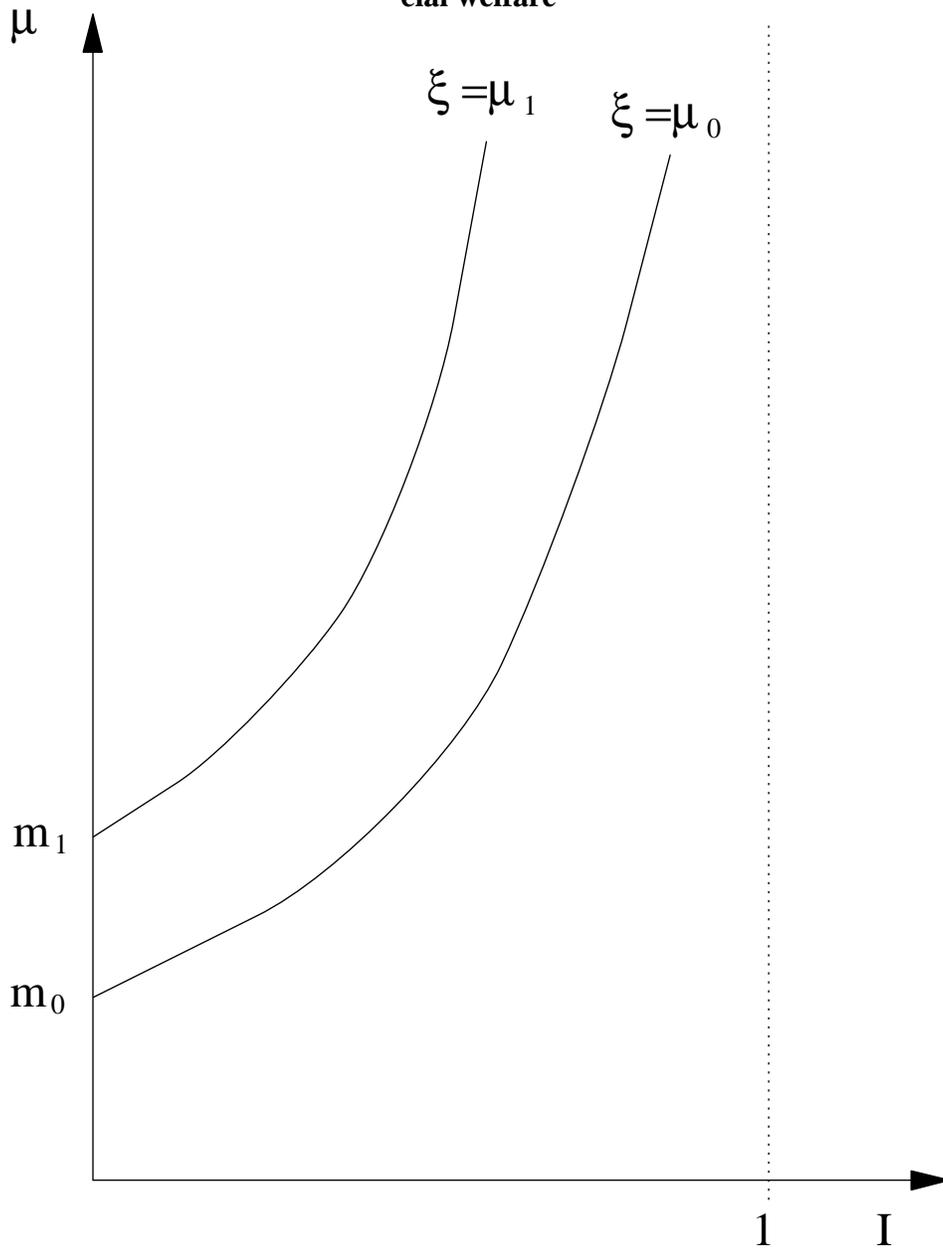
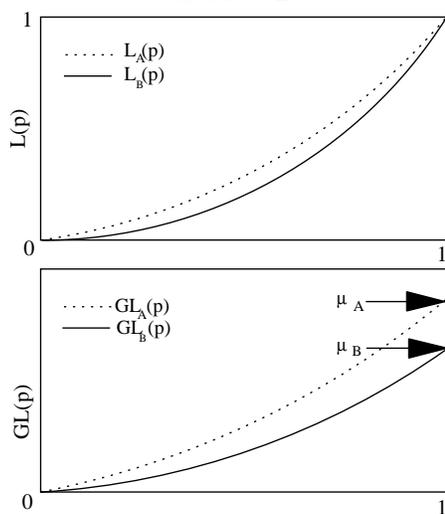
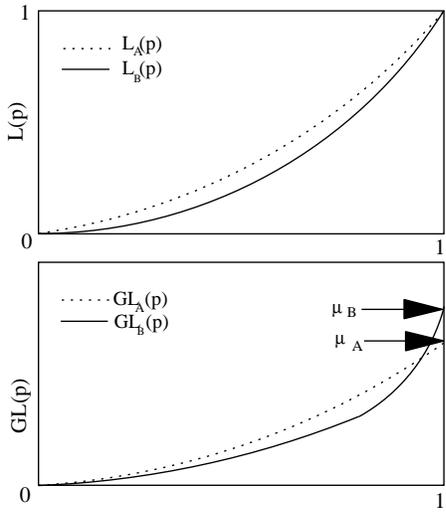


Figure 24: Inequality and social welfare dominance

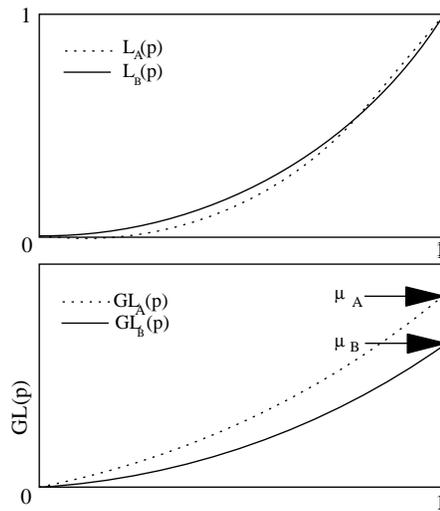
Case 1



Case 2



Case 3



Case 4

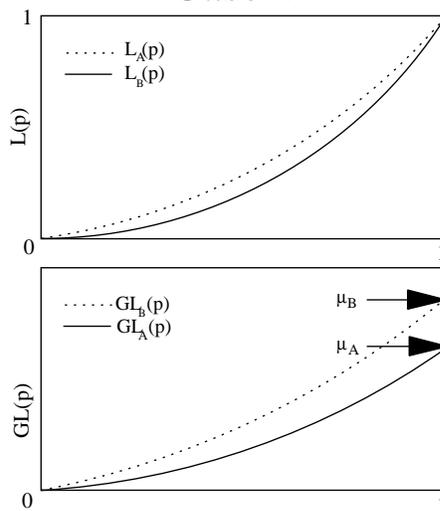


Figure 25: Primal stochastic dominance curves

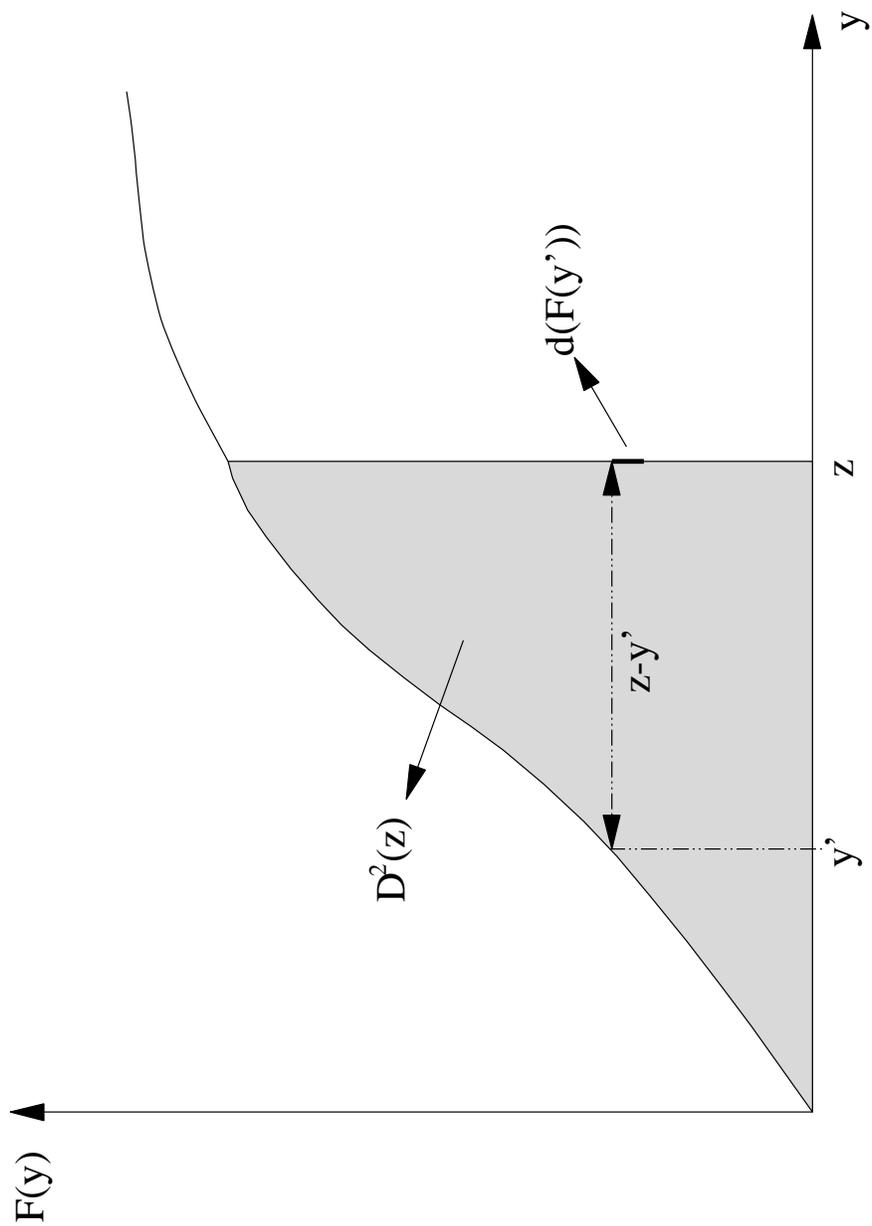


Figure 26: Histograms and density functions

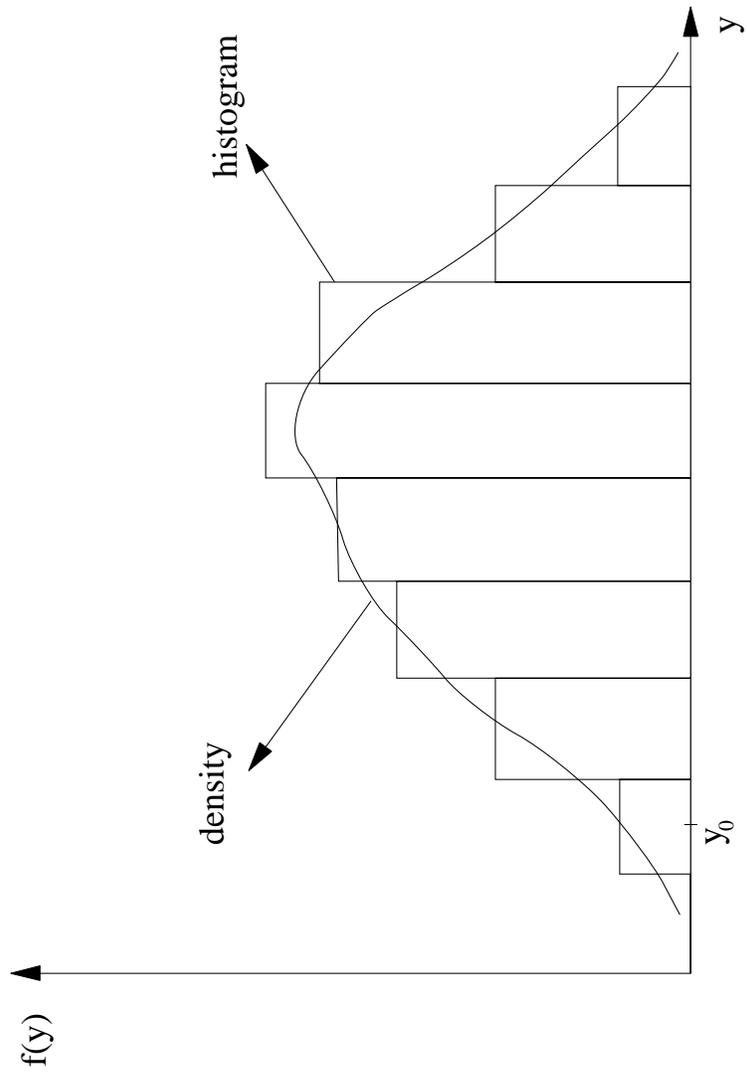


Figure 27: Poverty gaps and FGT indices

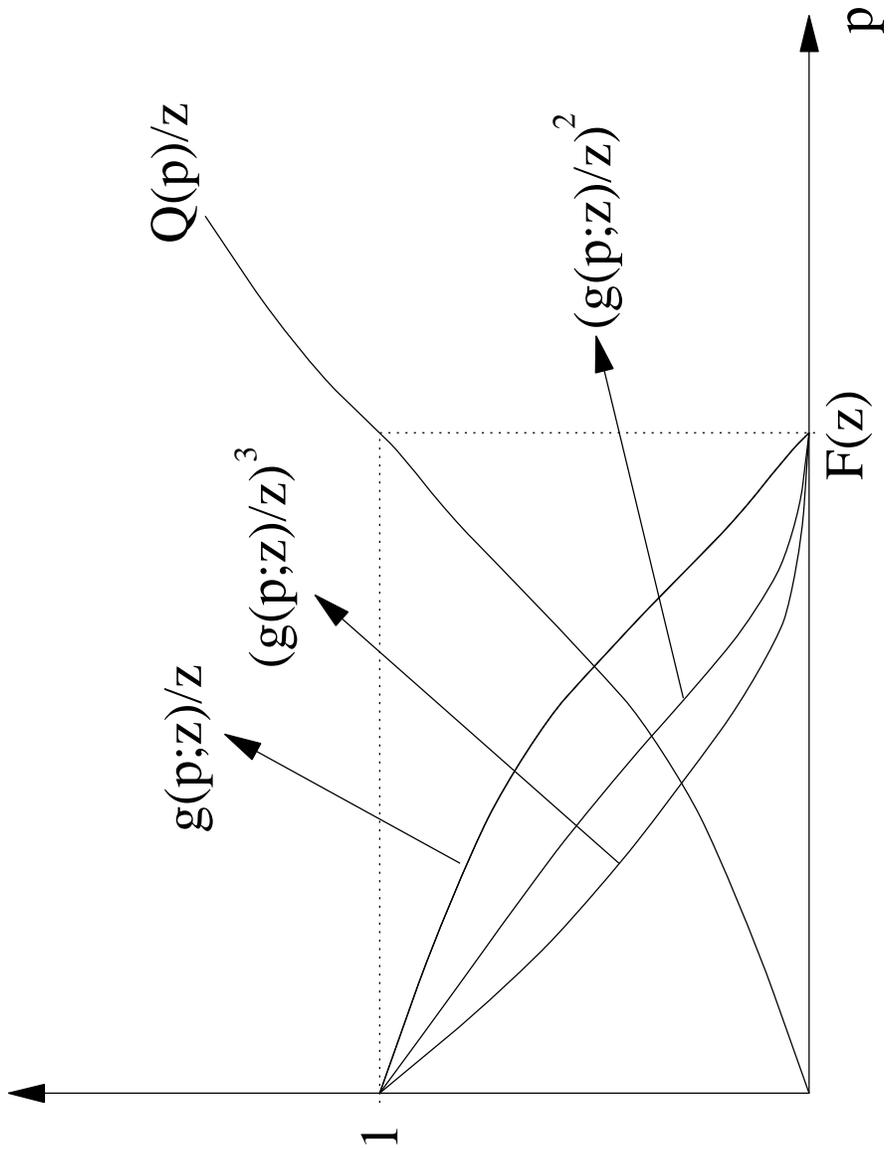


Figure 28: The relative contribution of the poor to FGT indices

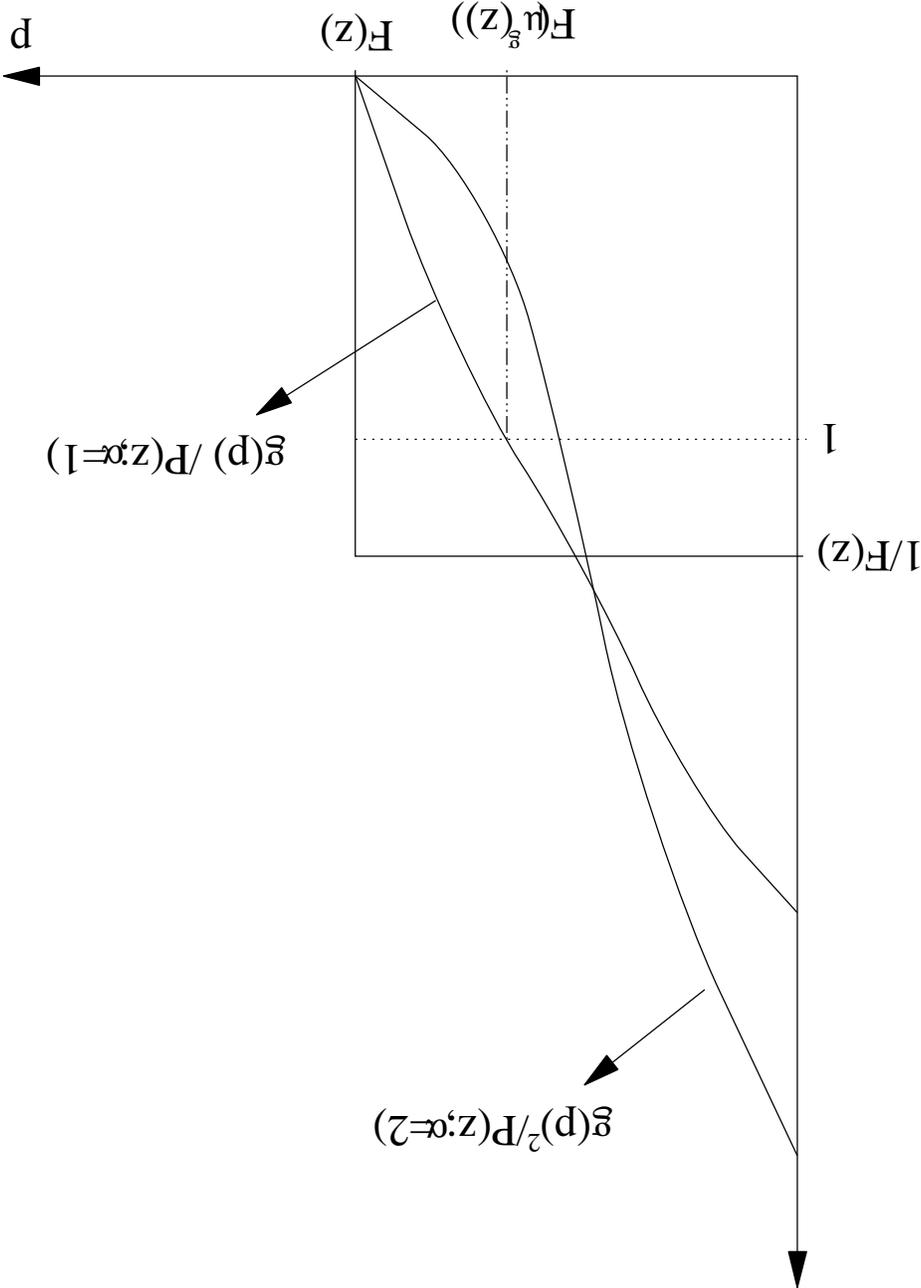


Figure 29: Socially-representative poverty gaps

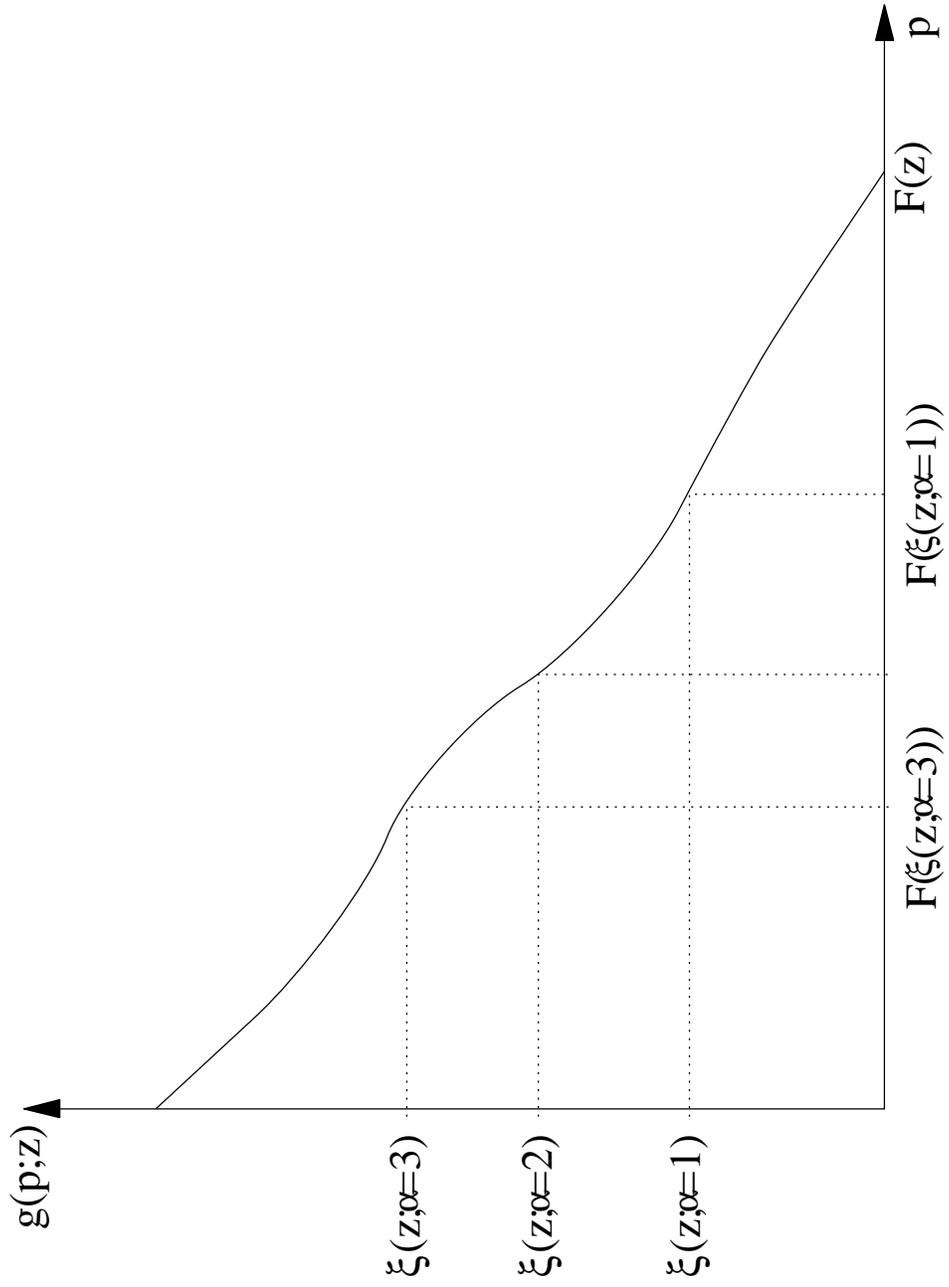


Figure 30: Growth elasticity of the poverty headcount

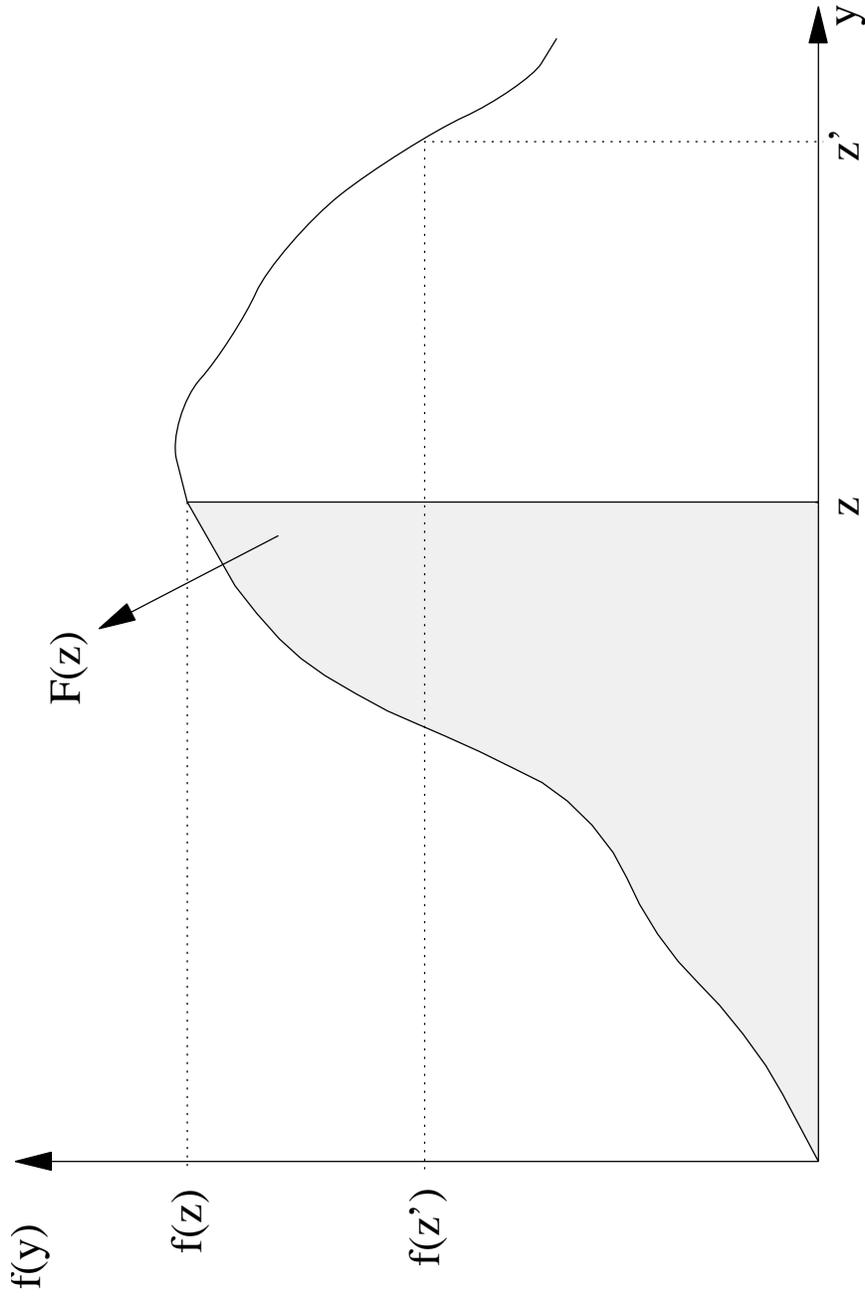


Table 1: Equivalent income and price changes

$q_1 = q_2 = 1$

i	y_i	x_1	x_2	U	y/z
1	150.00	50.00	100.00	79.37	0.50
2	210.00	70.00	140.00	111.12	0.70
3	300.00	100.00	200.00	158.74	1.00
4	380.00	126.67	253.33	201.07	1.27
5	500.00	166.67	333.33	264.57	1.67
6	510.00	170.00	340.00	269.86	1.70
7	550.00	183.33	366.67	291.02	1.83
8	600.00	200.00	400.00	317.48	2.00
9	800.00	266.67	533.33	423.31	2.67
10	1000.00	333.33	666.67	529.13	3.33

$q_1 = 1$ and $q_2 = 3$

i	y_i	x_1	x_2	U	y_i^R	y/z	y_i^R / z_r	$y/700$
1	160.00	53.33	35.56	40.70	76.92	0.26	0.26	0.23
2	200.00	66.67	44.44	50.88	96.15	0.32	0.32	0.29
3	500.00	166.67	111.11	127.19	240.37	0.80	0.80	0.71
4	630.00	210.00	140.00	160.26	302.87	1.01	1.01	0.90
5	1100.00	366.67	244.44	279.82	528.82	1.76	1.76	1.57
6	1240.00	413.33	275.56	315.43	596.13	1.99	1.99	1.77
7	1300.00	433.33	288.89	330.70	624.97	2.08	2.08	1.86
8	1500.00	500.00	333.33	381.57	721.12	2.40	2.40	2.14
9	1600.00	533.33	355.56	407.01	769.20	2.56	2.56	2.29
10	2770.00	923.33	615.56	704.64	1331.68	4.44	4.44	3.96

Table 2: The weighting function $\omega(p; \rho)$

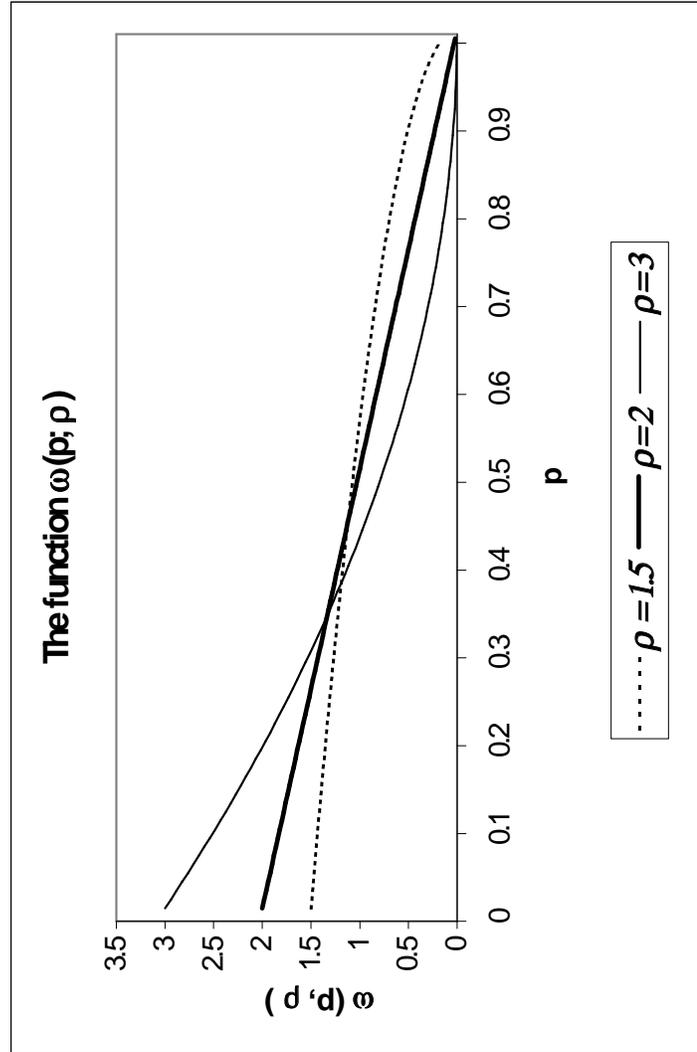


Table 3: The weighting function $\kappa(p; \rho)$

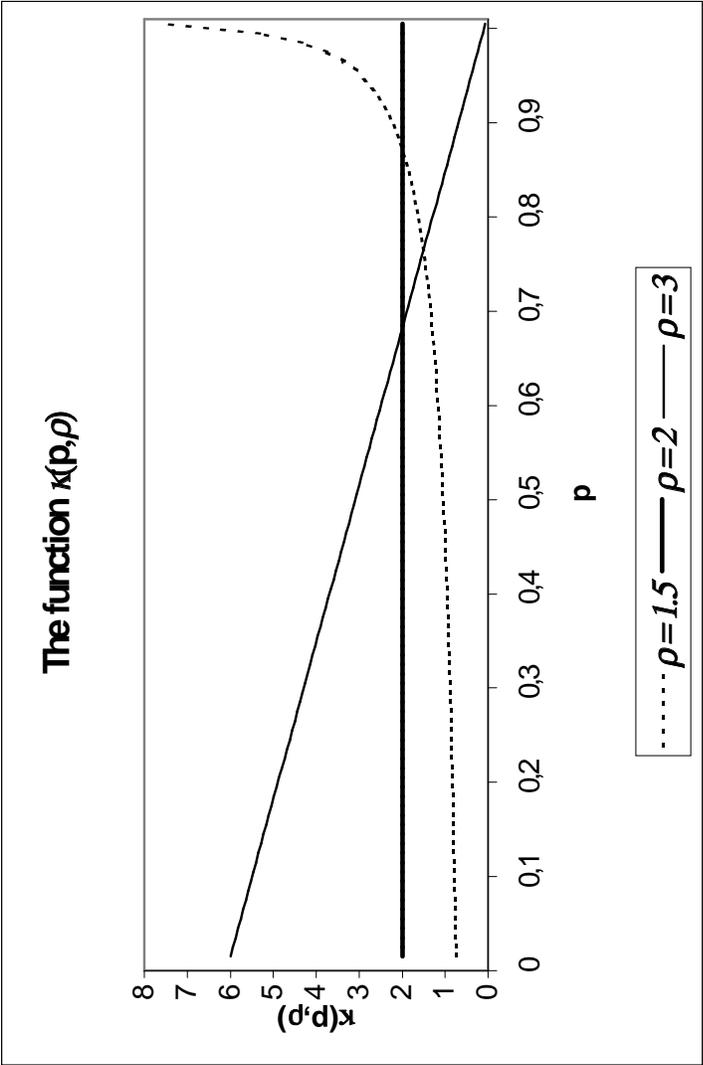


Table 4: Estimated poverty lines in Cameroon according to different methods (Francs CFA/day/adult equivalent)

	FEl food poverty line	Lower non-food poverty line	Lower total poverty line	Upper non-food poverty line	Upper total poverty line
Cameroon	256	117	373	278	534
Yaoundé	337	143	480	412	749
Douala	408	181	589	588	995
Other cities	347	152	499	385	732
Forests	259	134	393	214	473
Highlands	170	65	235	186	357
Savana	204	78	282	190	394

Table 5: Headcount according to alternative measurement methods and for different regions in Cameroon (%)

	Calorie poverty using common calorie poverty line	Food poverty using common food poverty line	Food poverty using regional food poverty lines	Total expenditure poverty using common lower CBN poverty line	Total expenditure poverty using regional lower CBN poverty line	Total expenditure poverty using common upper CBN poverty line	Total expenditure poverty using regional upper CBN poverty line	Proportion of region in total population
Cameroon	68.1	69.5	66.4	43.9	33.9	68.0	60.1	100 %
Yaoundé	73.4	53.1	67.9	19.2	34.7	41.6	59.6	7.1 %
Douala	67.3	42.0	67.5	16.5	38.1	33.4	59.0	9.6 %
Autres villes	59.9	44.5	63.2	16.0	31.8	36.5	58.8	12.7 %
Forêt	86.5	82.5	82.5	57.7	62.6	83.8	78.1	18.5 %
Hauts plateaux	64.6	82.5	61.1	58.7	19.0	81.1	53.1	27.8 %
Savane	61.1	74.0	61.2	49.0	29.7	78.7	55.8	24.2 %

Table 6: Distribution of the poor according to calorie, food and total expenditures poverty (% of the population)

	Calorie poor	Calorie non-poor
Poor in food expenditure	58.5 %	9.6 %
Non poor in food expenditure	11.2 %	20.7 %

	Poor in total expenditure	Non poor in total expenditure
Poor in food expenditure	56.6 %	9.8 %
Non poor in food expenditure	11.3 %	22.2 %

	Poor in total expenditure	Non poor in total expenditure
Calorie poor	55.8 %	12.3 %
Calorie non poor	12.2 %	19.7 %