CIRPÉE
Centre interuniversitaire sur le risque, les politiques économiques et l'emploi

Cahier de recherche/Working Paper **04-25**

What is "Pro-Poor"?

Jean-Yves Duclos
Quentin Wodon

Octobre/October 2004

Duclos: CIRPÉE, Département d'économique, Université Laval, Québec. Corresponding address: CIRPÉE, Département d'économique, Pavillon De Sève, Université Laval, Québec, Canada, G1K 7P4. Tel. (418)656-7096; Fax (418)656-7798
jyves@ecn.ulaval.ca
Wodon: LCSPR, World Bank, 1818 H Street, NW, Washington, DC 20433, USA
qwodon@worldbank.org

This paper was funded through a World Bank Research Support Budget and has also benefitted from the support of SSHRC, FQRSC, and the Poverty and Economic Policy research network of IDRC. We are grateful to Boniface Essama-Nssah and Christian Calmes for helpful comments.

Abstract:
Assessing whether distributional changes are « pro-poor » has become increasingly widespread in academic and policy circles. Starting from relatively general ethical axioms, this paper proposes simple graphical methods to test whether distributional changes are indeed pro-poor. Pro-poor standards are first defined. An important issue is whether these standards should be absolute or relative. Another issue is whether pro-poor judgements should put relatively more emphasis on the impact of growth upon the poorer of the poor. Having formalized the treatment of these issues, the paper describes various ways for checking whether broad *classes* of ethical judgements will declare a distributional change to be pro-poor.

Keywords: Poverty, Inequality, Pro-Poor growth

JEL Classification: D31, D63, H53, I32, I38

1 Introduction

Is growth good for the poor? This is certainly an issue on which much policy and academic debate has taken place recently[1]. But, in what sense can growth be declared "pro-poor"? Answering this question properly would seem to require going beyond the frequent use of simple average relationships between growth and some summary poverty statistics – as is being increasingly recognized in the literature.

There are two reasons for this. The first one is that the usefulness of summary poverty statistics depends on whether there is variability in the impact of growth upon the poor. The second reason is that summary poverty statistics invariably incorporate arbitrary and disputable normative judgements. This is true for all poverty statistics, but it is particularly valid for the most commonly used pro-poor measure: the change in a poverty headcount following growth. This change hides the variability of the impact of growth among the poor; it also largely depends on the impact of growth on those closest to the poverty line. For instance, growth may very well reduce on average the proportion of the poor in a population, but in some cases this may be at the cost of adverse and severe impacts on the very poor. Taking this cost into account would certainly seem important.

A related reason for caution is that the link between growth and changes in poverty indices can be highly sensitive to the choice of poverty lines. For instance, even if the poor's incomes always increased in line with average growth in the economy, the impact of growth on the headcount would vary erratically across countries according to their respective densities of income around the poverty line, and thus according to the choice of that poverty line. It can be shown for instance that, for a constant Lorenz curve and thus for constant relative inequality, the elasticity of the poverty headcount to growth will tend mechanically to increase with average income in the economy and to decrease with the poverty line.

An additional central issue in the discussion of growth is whether we should be interested in its impact on absolute poverty or on relative inequality[2]. As we discuss below, the same critical issue arises in the discussion of whether these changes are pro-poor. To assess whether growth is pro-poor, it is thus first important to distinguish between growth that is expected to change the incomes of the

[1] See, among many others, Bourguignon (2003), Bruno *et al.* (1998), Dollar and Kraay (2002), Eastwood and Lipton (2001), Ravallion (2002), United Nations (2000), and World Bank (2002).

[2] Note that concerns for relative inequality are closely linked to concerns for relative poverty, as has recently been discussed by Foster and Sen (1997), Zheng, Formby, Smith and Chow (2000), and Duclos and Makdissi (2004).

poor either by the same absolute or by the same proportional amount.

Absolute poverty is usually of greater concern in developing countries. Interest in relative poverty has nevertheless gained significant ground in developed economies[3]. It is also emerging as an important issue in developing countries too. One reason for this is that inequality may be potentially bad for growth[4]. Higher inequality may also be bad because, *ceteris paribus*, it usually makes poverty fall slower for a given level of economic growth. Inequality further breeds relative deprivation, economic isolation and social exclusion, which may be of concern for social cohesion and political stability. Finally, relative inequality can be deemed bad on its own for well-known ethical reasons – such as those developed in Rawls (1971).

The nature of the impact of growth on inequality and poverty will thus depend on numerous factors, such as the initial distributional conditions (namely, inequality and average income levels), the type of growth experienced, the functioning of markets, and the ability of the poor to partake in the growth process. Because of this, we can expect a high degree of heterogeneity of the effects of growth on absolute and relative poverty across and within countries.

It is not an objective of this paper to explore the empirical context-specific evidence for that heterogeneity. Instead, we propose methods that can help shed light on its magnitude. This paper does this by investigating how pro-poor judgements can be made robust to the choice of pro-poor evaluation functions and to the choice of poverty lines. This is done by considering classes of pro-poor evaluation functions which show varying distribution-sensitivity to the assessment of the impact of growth, and by considering ranges of possible poverty lines over which to define the sets of the poor. This is in contrast to much of the earlier literature which focussed on summary pro-poor measures with fixed poverty lines[5]. In con-

---

[3]See among many others Atkinson *et. al.* (2002).

[4]See for instance Alesina and Rodrik (1994) and Deiniger and Squire (1998).

[5]See, for instance, McCulloch and Baulch (1999) for the difference between a post-change poverty headcount with that headcount which would have occurred if all had gained equally, Kakwani *et al.* (2003) for a "poverty equivalent growth rate", Kakwani and Pernia (2000) for a ratio of the actual change in poverty over the change that would have been observed under distributional neutrality, Dollar and Kraay (2002) for a comparison of the growth rate in average income to the growth rate of incomes in the lowest quintile, Ravallion and Chen (2003) for a comparison of the growth rate in average income to a "population weighted" average growth rate of the initially poor percentiles of the population (for more, see page 8), Klasen (2003) for a comparison of the growth rate in average income to "population" and "poverty weighted" average growth rates, Essama-Nssah (2004) for the use of an ethically-flexible weighted average of individual growth rates that does not make use of poverty lines, and Ravallion and Datt (2002) for an example of the

trast to the earlier literature, we also distinguish formally between absolute and relative pro-poor judgements. Note finally that the derived tools can be applied equally well to understanding the impact of any distributional change, including that of negative growth – for instance, ”are recessions pro-poor?” – and of public expenditures.

The rest of the paper runs as follows. Section 2 formalizes our measurement of pro-poor changes, using either absolute or relative standards. Section 3 concludes. The Appendix regroups the proofs of the most important methodological results.

2 Measuring pro-poor changes

2.1 The general setting

Let $\mathbf{y}^1 = \left(y_1^1, y_2^1, \cdots, y_{n_1}^1\right) \in \Re_+^{n_1}$ be a vector of non-negative initial incomes[6] (at time 1) of size n_1, and let $\mathbf{y}^2 = \left(y_1^2, y_2^2, \cdots, y_{n_2}^2\right)$ be an analogous vector of posterior incomes (at time 2) of size n_2.

To determine whether the movement from $\mathbf{y}^1$ to $\mathbf{y}^2$ is pro-poor, we first need to define a standard with which this assessment can be made. A formal treatment of these standards is provided in Axioms 7 (relative) and 13 (absolute). Take the case of a relative standard, which we will denote as $1+g$. Roughly speaking, $1+g$ is some change in living standards that we wish the poor to undergo to ”catch up” with the change in the overall distribution of income. It will often be a function of the evolution of the entire income distributions, but it does not need to be. This standard will play a crucial role in the analysis below. Various interpretations can be given to it:

- It can be set on an ”ethical” basis. An example of this would be $1 + g$ set as the ratio of the mean of $\mathbf{y}^2$ over that of $\mathbf{y}^1$. It would then be felt desirable that the incomes of the poor increase in proportion to average growth in the population (see again Axiom 7). As we discuss below in Section 2.2, this has links with concerns for relative poverty. Another example of relativity would be the ratio of changes in ”equally distributed equivalent incomes” of the well-known Kolm-Atkinson-Sen type. The pro-poor standard can also be absolute: it would then be felt desirable that the incomes of the poor

very common use of growth elasticities of poverty measures. A recent study which goes beyond focussing on summary pro-poor measures with a fixed poverty line is Son (2004) – see also the discussion on page 13.

[6]Or consumption, wealth, or any other welfare indicator of interest.

increase by at least the same amount as some indicators of welfare in the population — see Axiom 13.

- $1+g$ can also be set on a more "statistical" basis. An example of this would be the use of the ratio of the median of $\mathbf{y}^2$ over that of $\mathbf{y}^1$ — it is often argued that, for relative poverty comparisons, estimators of the median are statistically more robust and less subject to sampling variability than estimators of the mean.

- g can also be set on the basis of political or administrative criteria. Government agencies can, for instance, fix pro-poor objectives for changes in the living standards of the poor, and then wish to assess whether these objectives have been met.

Note that the above framework is general enough to accommodate negative as well as positive growth.

Denote by $z>0$ a poverty line defined in real terms. Let $W\left(\mathbf{y}^1,\mathbf{y}^2,g,z\right)$ be a pro-poor evaluation function. This is defined as the difference between two evaluation functions $P\left(\mathbf{y}^1,z\right)$ and $P^*\left(\mathbf{y}^2,1+g,z\right)$, each for time 1 and time 2, respectively:

$$W\left(\mathbf{y}^1,\mathbf{y}^2,g,z\right)\equiv P^*\left(\mathbf{y}^2,1+g,z\right)-P\left(\mathbf{y}^1,z\right). \tag{1}$$

We can interpret P and P^* as assessing the "ill-fare" of the poor in the initial and posterior distributions, respectively. Note that the evaluation function $P^*\left(\cdot\right)$ differs from $P\left(\cdot\right)$ in part because of the use of the pro-poor standard g in assessing the second distribution. We then have:

Definition 1 *The change from* $\mathbf{y}^1$ *to* $\mathbf{y}^{\mathbf{2}}$ *is pro-poor if* $W\left(\mathbf{y}^1,\mathbf{y}^2,g,z\right)\leq 0$.

Clearly, whether the distributional change will be deemed pro-poor will depend on the way in which z, P, and P^* will be chosen. The central goal of this paper is to explore how we may circumvent this dependence by imposing suitable general conditions on these objects, for a given g. Note that the weak inequality in Definition 1 could alternatively be replaced by a strict inequality, to make it a strict definition of pro-poor changes. A consequence of this would be to change the axioms to be defined below to strict axioms, and the inequalities in the conditions of the Theorems to strict conditions.

We start with a focus axiom:

Axiom 2 (Focus on the poor) *Let* $\mathbf{y} = (y_1, \ldots, y_n)$ *and* $\dot{\mathbf{y}} = (\min(y_1, z), \ldots, \min(y_n, z))$. *Then* $W(\mathbf{y}, \mathbf{y}^2, g, z) = W(\dot{\mathbf{y}}, \mathbf{y}^2, g, z)$.

This is a rather uncontroversial axiom when the objective is to assess the well-being of the poor and of its evolution. Note, however, that this focus axiom does not imply that pro-poor judgements are necessarily made irrespective of the evolution of the well-being of the rich. As discussed above, taking into account the evolution of the overall distribution would and can appear through the standard g.

Axiom 3 (Population invariance) *Adding a replication of a population* $\mathbf{y}^j, j = 1, 2,$ *(initial or posterior) to that same population has no impact on* W.

This is a common axiom in welfare economics. For our purposes, it makes it possible to make pro-poor judgements even when the absolute population size varies across the distributions. Note that this axiom may not always be appropriate. One example is the case of the assessment of the pro-poor effect of the AIDS epidemics. This epidemics might in some circumstances have the impact of increasing the average well-being of the surviving poor, but at the presumably substantial social costs of decreasing total population size.

Axiom 4 (Population symmetry or anonymity) *Let a (initial or posterior) distribution of size* n *be given by* $\mathbf{y}$. *Let* M *be an* $n \,\mathrm{x}\, n$ *permutation matrix*[7] *and let* $\dot{\mathbf{y}} = M\mathbf{y}'$. *Then we should have that* $W(\mathbf{y}, \mathbf{y}^2, g, z) = W(\dot{\mathbf{y}}, \mathbf{y}^2, g, z)$.

This axiom is also standard in welfare economics. Permuting the incomes of any two persons in any given distribution does not affect pro-poor judgements. Note that this axiom leads to violations of the well-known Pareto efficiency-criterion: for growth to be declared pro-poor, for instance, it is not needed that none of the poor be penalized by the change. Hence, pro-poor growth (as defined here) is compatible in principle with a fair amount of "horizontal inequality" (for a discussion of this, see for instance Ravallion (2003) and Bibi and Duclos (2003)).

Axiom 5 (Monotonicity) *Let* $\mathbf{y}$ *be an income vector, let* $\epsilon > 0$ *be any positive constant, and let* $\dot{\mathbf{y}} = (y_1, \ldots, y_j + \epsilon, \ldots, y_n)$. *Then* $W(\mathbf{y}^1, \mathbf{y}, g, z) \geq W(\mathbf{y}^1, \dot{\mathbf{y}}, g, z)$.

[7] A permutation matrix is composed of 0's and 1's, with each row and each column summing to 1.

Axiom 5 is reminiscent of the Pareto principle: for a given g, if anyone's posterior income increases, W should not increase, and may sometimes fall. Because of the anonymity axiom 4, it does not follow, however, that only those changes that are Pareto efficient will be judged pro-poor. As hinted to already, many of the pro-poor changes that we will be able to identify in the aggregate will in fact involve adverse changes for many of the poor.

The following is a normalization axiom: if there has been no distributional change, and if the pro-poor standard g equals 0, then the pro-poor judgement is necessarily neutral.

Axiom 6 (No distributional change combined with $g = 0$ implies pro-poor neutrality) *For any* $\mathbf{y} \in \Re^{\mathbf{n}}_{+}$*, we must have* $W(\mathbf{y}, \mathbf{y}, 0, z) = 0$.

2.2 Pro-poor judgements using relative pro-poor standards

We start with the first of two main approaches to assessing pro-poor changes: the relative (or proportional) one. (The second approach will be discussed in section 2.3.) This first approach is also the most widespread and probably the least controversial of the two. It essentially says that pro-poor judgements should be made by comparing the growth of the poor's living standards to a standard g. It is consistent, for instance, with the view of Kakwani and Pernia (2000) that "promoting pro-poor growth requires a strategy that is deliberately biased in favor of the poor so that the poor benefit proportionately more than the rich. (p.3)". The relevant axiom is as follows:

Axiom 7 (Relative pro-poor standards) *Consider two posterior distributions,* $\mathbf{y}$ *and* $\dot{\mathbf{y}}$*, both of sizes* n*, with respective pro-poor standards* g *and* $\dot{g}$*. Suppose that* $\mathbf{y}/(1+g) = \dot{\mathbf{y}}/(1+\dot{g})$*. Then,* $\mathbf{y}$ *and* $\dot{\mathbf{y}}$ *should be judged equally pro-poor by* W *regardless of the initial distribution* $\mathbf{y}^1$*, that is, we should have that*

$$W\left(\mathbf{y}^1, \mathbf{y}, g, z\right) = W\left(\mathbf{y}^1, \dot{\mathbf{y}}, \dot{g}, z\right) \tag{2}$$

for any choice of $\mathbf{y^1}$.

2.2.1 First-order pro-poor judgements

The axioms above define a first class of pro-poor evaluation functions.

Definition 8 *The class of pro-poor evaluation functions* $\Omega^1(g, z^+)$ *is made of all of the functions* $W(\cdot, \cdot, g, z)$

- *which satisfy the focus-on-the-poor axiom 2, the population axiom 3, the anonymity axiom 4, the monotonicity axiom 5, the normalization axiom 6, and the proportionality axiom 7,*
- *and for which* $z \leq z^{+}$.

The problem is then to check whether all such pro-poor evaluation functions will unanimously declare some distributional change to be pro-poor. To check this, let $I(\cdot)$ be an indicator function that takes the value 1 if its argument is true and 0 otherwise. The distribution function $F^{j}(y)$ is then defined as

$$F^{j}(y) = n_{j}^{-1} \sum_{i=1}^{n_{j}} I(y_{i}^{j} \leq y). \tag{3}$$

The distribution function for $\mathbf{y}^{j} / (1+g)$ is given by $\overline{F}^{j}(y) = F^{j}((1+g)\,y)$. Using $\overline{F}^{j}(y)$ is equivalent to using the distribution of incomes $\mathbf{y}^{j}$ divided by $(1+g)$. First-order pro-poor distributional changes are then identified as follows[8]:

Theorem 9 (First-order relative pro-poor judgements) *A movement from* $\mathbf{y^1}$ *to* $\mathbf{y^2}$ *is judged pro-poor by all pro-poor evaluation functions* $W(\cdot,\cdot,g,z)$ *that are members of* $\Omega^{1}(g,z^{+})$ *if and only if*

$$F^{2}\left((1+g)z\right) \leq F^{1}\left(z\right) \text{ for all } z \in [0, z^{+}]. \tag{4}$$

A distributional change that satisfies (4) is called first-order pro-poor since all pro-poor evaluation functions within $\Omega^{1}(g,z^{+})$ will find that it is pro-poor, and this, for any choice of poverty line within $[0,z^{+}]$. (The term ”first-order” is used in reference to the stochastic dominance literature, where utility or social welfare functions are deemed to be of the first-order type if they are monotonically increasing in returns or in incomes.)

Verifying (4) simply involves checking whether – over the range of poverty lines $[0,z^{+}]$ – the headcount index in the initial distribution is larger than the headcount index in the posterior distribution when that distribution is normalized by $1+g$. An example of this is shown on Figure 1. The movement from distribution 1 to distribution 2 is first-order pro-poor for all choices of poverty lines up to z^{++}, which in this case includes z^{+}.

[8]The proofs of Theorems 9 and 12 appear in the appendix.

2.2.2 Discussion

There are several alternative (though equivalent) ways of checking whether a distributional change can be declared first-order pro-poor. These alternative procedures may be deemed attractive on intuitive, expositional, computational or statistical grounds. To describe them, define as $Q^j(p)$ the quantile function for distribution F^j.[9] In a continuous setting and with a strictly positive income density, $Q(p)$ is simply the inverse of the distribution function, that is, it equals $F^{(-1)}(p)$. Roughly speaking, $Q(p)$ is the income of that individual who is at rank p in the distribution. The normalized quantile for F^j is $\overline{Q}^j(p) = Q^j(p)/(1+g)$. The normalized poverty gap (or deprivation) at rank p is then given by $d^j(p,z) = z^{-1}\max(0,(z-Q^j(p)))$.

Checking condition (4) can then be shown to be equivalent to checking, for all $p \in [0, F^1(z^+)]$,

1. that
$$\overline{Q}^2(p) \geq Q^1(p), \tag{5}$$
viz, the normalized posterior incomes at rank p are larger than the incomes at the same rank p before the change;

2. that
$$\frac{Q^2(p) - Q^1(p)}{Q^1(p)} \geq g \tag{6}$$
– the income growth rate at rank p is larger than g;

3. or that
$$d^2(p,(1+g)z^+) \leq d^1(p,z^+) \tag{7}$$
i.e., the posterior poverty gap with $(1+g)z^+$ is lower than the poverty gap before with z^+.

Ravallion and Chen (2003) suggests the use of ”growth incidence curves” to check whether growth is pro-poor. These curves show the growth rates of living standards at different ranks in the population. In our notation, they are defined as

$$\Gamma(p) = \frac{Q^2(p) - Q^1(p)}{Q^1(p)}. \tag{8}$$

[9]It is formally defined as $Q^j(p) = \inf\{s \geq 0 | F^j(s) \geq p\}$ for $p \in [0,1]$.

An alternative name for these curves would be ”income growth curves”. Such a name could avoid confusion with the well-known ”poverty incidence curves”, since the income growth curves do not in themselves say much about the incidence of poverty, or about its change.

Ravallion and Chen show that the average height of these curves is linked to changes in the Watts index of poverty following distributional changes. Although this interpretation is certainly useful, the main disadvantage of this average height is indeed that it is strictly valid only for the Watts index. The Watts index has, indeed, properties with which not all pro-poor analysts will necessarily agree. To see this, assume a poverty line equal to 100,000 (in whatever units). Also assume two individuals with income 1 and 20 respectively. For most analysts, these two individuals will not seem very different in terms of deprivation since their income's distance from the poverty line is roughly the same. Yet, the Watts index *falls* following a distributional change that gives 1 unit of income to the first individual and withdraws 9 units from the second individual. This also shows why (assuming $g = 0$) the Watts index would say that such a change is pro-poor, even though it decreases significantly the average incomes of the very poor. Other pro-poor judgements may clearly not agree with this.

Note also that the link between the area under the income growth curves and the change in the Watts poverty index is only valid for *marginal* distributional changes. Instead of taking the average of income growth rates, it would seem safer to consider the entire income growth curve $\Gamma(p)$ of equation (8). This is done by condition (6) and is again equivalent to checking whether a distributional change is unambiguously first-order pro-poor.

Care must also be taken in the interpretation of relative pro-poor comparisons when these are made across countries with varying headcounts. Assume that the pro-poor standard $1 + g$ is set as the ratio of mean incomes. By definition, it will then be much more difficult to have a ”pro-poor growth rate” (and thus a ”rich-averse growth rate”) in societies in which there are *few* rich. At the limit, if everyone is initially poor, it will be impossible to verify condition (4)– this is because it is impossible for everyone's income to grow faster than average income. Relative pro-poor judgements would then seem to make sense only in distributions in which there is a significant number of non-poor individuals to whom the poor can be compared.

The use of the above conditions is illustrated on Figures 2, 3 and 4. The filled line on Figure 2 shows the values of the p-quantiles in the posterior distribution (on the vertical axis) against the values of the p-quantiles (of the same percentile p)

of the initial distribution (on the horizontal axis). The condition $\overline{Q}^2(p) \geq Q^1(p)$ requires that this line be above a line that starts from the origin with a slope of $1+g$. Two such pro-poor standards $1+g$ are shown on Figure 2: the first one, m_2/m_1, is the ratio of the medians, and the second one, μ_2/μ_1, is the ratio of the means. The distributional change is deemed first-order pro-poor for all choices of poverty lines within a range $[0, z^{++}]$ when the ratio of the medians is considered to be the relevant pro-poor standard. That range extends beyond z^+ when the ratio of the mean is used instead.

An equivalent statement is obtained by looking instead at the income growth curve $\Gamma(p)$ of Figure 3. Recall that we need to check whether that curve is above g. On Figure 3, g is taken to be either the growth in median or in average income. When growth in median income is considered ($g = m_2/m_1 - 1$), the distributional change is considered first-order pro-poor over all poverty lines within $[0, z^{++}]$; that range extends again further (and in fact beyond z^+) when growth in mean income ($g = \mu_2/\mu_1 - 1$) is taken as the pro-poor standard. An alternative way to affect the range of poverty lines over which the distributional change is first-order pro-poor is to ask that the incomes of the poor grow by more than a proportion γ^+ of the growth of median income ($g = \gamma^+ (\mu_2/\mu_1 - 1)$). With some $\gamma^+ < 1$, the growth shown on Figure 3 is judged pro-poor until z^+.

The link between income growth and changes in poverty gaps is illustrated on Figure 4. The values of the p-quantiles are shown on the left vertical axis and those of the poverty gaps appear on the right vertical axis. For all $p \in [0, F^1(z^+)]$, we have that $\overline{Q}^2(p) \geq Q^1(p)$. Thus, the use of the quantiles $Q^1(p)$ and $\overline{Q}^2(p)$ in Figure 4 shows first-order pro-poorness until (at least) z^+. This is also verified by "inverting" the axes and noting on the horizontal axis that $\overline{F}^2(z) \leq F^1(z)$ for all $z \in [0, z^+]$ (condition (4)). Condition (7) is verified on Figure 4 by noting that the posterior gaps with $(1+g)z^+$ are always larger than the initial gaps with z^+ whatever the percentiles p considered.

2.2.3 Second-order pro-poor judgements

First-order pro-poor judgements are demanding. They require *all* quantiles of the poor to undergo a rate of growth at least as large as g. Some pro-poor analysts may be willing to relax this condition on the basis that a large rate of growth for the poorer among the poor may sometimes be ethically sufficient to offset a rate of growth for the not-so-poor that may be below g. This is captured by the following axiom.

Axiom 10 (Distribution sensitivity) *Let* $\mathbf{y}$ *be an ordered income vector,* $\epsilon > 0$ *be any positive value, and let* $\dot{\mathbf{y}} = (y_1, \ldots, y_j + \epsilon, \ldots, y_k - \epsilon, \ldots, y_n)$*, with* $y_j + \epsilon \leq y_k$*. Then* $W(\mathbf{y}^1, \mathbf{y}, g, z) \geq W(\mathbf{y}^1, \dot{\mathbf{y}}, g, z)$.

This axiom is analogous to the well-known Pigou-Dalton principle of transfers in welfare economics. It says that the evaluation functions P should give more weight to the poorer than to the not-so-poor among the poor. "By how much more?" does not need to be specified in our general context (since we are interested in classes of pro-poor judgements). Axiom 10 thus leads to "distribution-sensitive" pro-poor judgements: shifting incomes from the richer to the poorer is a pro-poor distributional change.

The monotonicity and the distribution-sensitive axioms lead to two different orders of pro-poor judgements. The first order (monotonicity) says that distributional impacts on the poor are independently important at all poor individuals' initial income levels. The second order (distribution sensitivity) imposes that the distributional impacts on the poorer individuals cannot be ethically less important than similar distributional impacts on the richer individuals.

Definition 11 *The class of pro-poor evaluation functions* $\Omega^2(g, z^+)$ *is made of all functions* $W(\cdot, \cdot, g, z)$

- *which satisfy the focus-on-the-poor axiom 2, the population axiom 3, the anonymity axiom 4, the monotonicity axiom 5, the normalization axiom 6, the proportionality axiom 7, and the distribution-sensitivity axiom 10,*
- *and for which* $z \leq z^+$.

Now define the poverty deficit $D^j(z)$ as:

$$D^j(z) = n_j^{-1} \sum_{i=1}^{n_j} z^{-1} \left(z - y_i^j\right) I\left(y_i^j \leq z\right). \tag{9}$$

This leads to:

Theorem 12 (Second-order relative pro-poor judgements) *A movement from* $\mathbf{y^1}$ *to* $\mathbf{y^2}$ *is judged pro-poor by all pro-poor evaluation functions* $W(\cdot, \cdot, g, z)$ *that are members of* $\Omega^2(g, z^+)$ *if and only if*

$$D^2\left((1+g)\, z\right) \leq D^1(z) \text{ for all } z \in [0, z^+]. \tag{10}$$

A distributional change that satisfies condition (10) is called second-order pro-poor since all relative pro-poor evaluation functions that are distribution-sensitive will find that it is pro-poor, and this, for any choice of poverty line within $[0, z^+]$. To reach this conclusion, it must simply be checked that the initial poverty deficit using z is larger than the posterior poverty deficit with $(1+g)z$, over a range of poverty lines $z \in [0, z^+]$.

2.2.4 Discussion

As for first-order pro-poor judgements, there are alternative equivalent ways of checking condition (10). Define the cumulative poverty gap[10] up to rank p as

$$G^j(p,z) = \int_0^p d^j(q,z)dq. \tag{11}$$

Note that $G^j(p,z)$ attains its maximum value of $D^j(z)$ at $p = F^j(z)$. Checking condition (10) is then equivalent to checking that $G^2(p,(1+g)\,z^+) \leq G^1(p,z^+)$ for all $p \in [0,1]$. This is illustrated on Figure 5. Because $G^2(p,(1+g)\,z^+) \leq G^1(p,z^+)$ for all $p \in [0,1]$, the distributional change is deemed second-order pro-poor for any choice of poverty lines between 0 and z^+. Note that this implies graphically that $D^2\left((1+g)\,z^+\right) \leq D^1\left(z^+\right)$ since $G^j(1,z) \equiv D^j(z)$. But it does not follow that $\overline{F}^2\left(z^+\right) \leq F^1\left(z^+\right)$ (recall condition (4)). In fact, the opposite is shown on Figure 5: the headcount (with $(1+g)z^+$) after the change is larger than the headcount (with z^+) before the change. First-order pro-poorness implies second-order pro-poorness, but not the reverse.

Similarly to (11), the cumulative income up to rank p (the Generalized Lorenz curve at p) is given by

$$C^j(p) = \int_0^p Q^j(q)dq. \tag{12}$$

The use of the Generalized Lorenz curve provides an intuitive *sufficient* condition for checking second-order pro-poor change. A distributional change is indeed second-order pro-poor if for all $p \in [0, \overline{F}^2(z^+)]$,

$$\lambda(p) \equiv \frac{C^2(p) - C^1(p)}{C^1(p)} \geq g. \tag{13}$$

[10] This is also called a TIP curve by Jenkins and Lambert (1997), and a poverty gap profile by Shorrocks (1998); see also Spencer and Fisher (1992).

Expression (13) involves computing the growth rates in the cumulative incomes of proportions p of the poorest, and to compare those growth rates to g. If the cumulative incomes of the poor increase faster than the pro-poor standard, then growth is pro-poor for all relative distribution-sensitive pro-poor assessments. Note also that when $1+g$ equals the ratio of mean income, condition (13) is equivalent to checking whether the Lorenz curve for $\mathbf{y}^2$ is above that of $\mathbf{y}^1$ for the range of $p \in [0, \overline{F}^2(z^+)]$.

The use above of the Lorenz and generalized Lorenz curves to check for second-order pro-poor changes is reminiscent of Son (2004). There are three main differences between this section's contribution and that of Son. First, Son does not condition her comparisons of the Lorenz and generalized Lorenz curves on those incomes falling below an upper poverty line z^+ — she implicitly sets $z^+ = \infty$. Given that we can generally agree on some finite upper bounds for ranges of possible poverty lines, setting z^+ to infinity would seem to be unnecessarily strong and to limit too much one's ability to identify second-order pro-poor changes. Second, she considers only additive poverty evaluation functions. Finally, she assumes that relative pro-poor standards $1+g$ equal the ratio of mean incomes. The analysis here is thus more general on these three aspects.

The combined use of conditions (6) and (13) is illustrated on Figure 6. Note first (as for Figure 5) that $\overline{F}^2(z^+) > F^1(z^+)$. Moreover, the income growth curve $\Gamma(p)$ clearly shows that income growth is sometimes lower than g for some of the percentiles below $F^1(z^+)$. Hence, the distributional change of Figure 6 is not first-order pro-poor for all poverty lines up to z^+. That change could be deemed first-order pro-poor only if we relaxed our pro-poor standard (by decreasing G to $\Gamma(F^1(z^+))$, or if we chose z^{++} instead of z^+ as the upper bound of the range of possible poverty lines. An alternative route to generating pro-poorness would be to add the distribution-sensitivity Axiom 10. Doing this indeed makes the distributional change of Figure 6 second-order pro-poor for all poverty lines up to z^+ since $\lambda(p) \geq g$ for all p between 0 and $\overline{F}^2(z^+)$.

2.3 Pro-poor judgements using absolute pro-poor standards

The second of the two main approaches alluded to before is an *absolute* one. It says that pro-poor judgements should be made by comparing the absolute change in the poor's living standards to some absolute pro-poor standard a. Although a is an absolute standard in the sense of Axiom 13, it need not be independent of the distribution of living standards. It could represent, for instance, the absolute change in average living standards or in some equally-distributed living standards.

Axiom 13 (Absolute pro-poor standards) *Consider two posterior distributions,* $\mathbf{y}$ *and* $\dot{\mathbf{y}}$*, both of sizes* n*, with respective pro-poor standards* a *and* $\dot{a}$*. Suppose that* $\mathbf{y} + a = \dot{\mathbf{y}} + \dot{a}$*. Then,* $\mathbf{y}$ *and* $\dot{\mathbf{y}}$ *should be judged equally pro-poor by* W*, that is, we should have that*

$$W\left(\mathbf{y}^1, \mathbf{y}, a, z\right) = W\left(\mathbf{y}^1, \dot{\mathbf{y}}, \dot{a}, z\right) \tag{14}$$

for any choice of $\mathbf{y^1} \in \Re^{\mathbf{n}}_{+}$*.*

This alternative axiomatization says essentially that P^* should be "translation invariant" in **y** and a. The pro-poor judgement should be neutral whenever the poor gain in absolute terms the same as the standard a. This axiom allows us to define the following class of absolute pro-poor evaluation functions.

Definition 14 *The class of pro-poor evaluation functions* $\tilde{\Omega}^1(a, z^+)$ *is made of all of the functions* $W(\cdot, \cdot, a, z)$

- *which satisfy the focus-on-the-poor axiom 2, the population axiom 3, the anonymity axiom 4, the monotonicity axiom 5, the normalization axiom 6, and the absoluteness axiom 13,*
- *and for which* $z \leq z^+$*.*

Then:

Theorem 15 (First-order absolute pro-poor judgements) *A movement from* $\mathbf{y^1}$ *to* $\mathbf{y^2}$ *is judged pro-poor by all pro-poor evaluation functions* $W(\cdot, \cdot, a, z)$ *that are members of* $\tilde{\Omega}^1(a, z^+)$ *if and only if*

$$F^2(z + a) \leq F^1(z) \text{ for all } z \in [0, z^+]. \tag{15}$$

An analogous result holds for absolute pro-poor judgements that are distribution sensitive.

Definition 16 *The class of pro-poor evaluation functions* $\tilde{\Omega}^2(a, z^+)$ *is made of all functions* $W(\cdot, \cdot, a, z)$

- *which satisfy the focus-on-the-poor axiom 2, the population axiom 3, the anonymity axiom 4, the monotonicity axiom 5, the normalization axiom 6, the distribution-sensitivity axiom 10, and the absoluteness axiom 13,*

- *and for which* $z \leq z^{+}$.

Let $\tilde{D}^{j}(z)$ be defined as $zD^{j}(z)$. This leads to:

Theorem 17 (Second-order absolute pro-poor judgements) *A movement from* $\mathbf{y^1}$ *to* $\mathbf{y^2}$ *is judged pro-poor by all pro-poor evaluation functions* $W(\cdot, \cdot, a, z)$ *that are members of* $\tilde{\Omega}^2(a, z^{+})$ *if and only if*

$$\tilde{D}^{2}(z+a) \leq \tilde{D}^{1}(z) \text{ for all } z \in [0, z^{+}]. \tag{16}$$

Note that the change in the average income of the bottom p proportion of the population is given by $\left(C^2(p) - C^1(p)\right)/p$. A sufficient condition for Condition (16) is then to verify whether that change exceeds a whatever the value of $p \in [0, \tilde{F}^2(z^{+})]$.

3 Concluding remarks

The paper has proposed simple graphical tests to test whether distributional changes are "robustly" pro-poor, in the sense of whether broad classes of ethical judgements would declare a distributional change to be pro-poor. An important issue is whether pro-poor judgements should put relatively more emphasis on the impact of growth upon the poorer of the poor. Another issue is whether these standards should be absolute or relative. A number of tests logically equivalent to those of Theorems 9 and 12 have also been outlined for relative pro-poor judgements. Analogous equivalent conditions can further be derived for absolute pro-poor judgements of the first and second order. It is also possible to derive tests for pro-poor judgements of any higher order desired using curves of normalized FGT (Foster, Greer and Thorbecke (1984)) indices, as is done in the stochastic dominance literature (see for instance Duclos and Makdissi (2004)).

As was mentioned in the introduction, the formulation of the paper is general enough to accommodate negative and positive growth as well as whether public policy and public expenditures are pro-poor. Note finally that the property of the pro-poor standard does not need to be independent of whether growth is negative or not. For instance, one may choose for positive growth a relative pro-poor standard, and for negative growth an absolute one. Then, positive growth will be deemed pro-poor only if it increases the incomes of the poor by proportionately more than the relative standard, but a recession will be deemed pro-poor only if does not lead to an absolute decrease of those same incomes — independently of whether the relative distribution of incomes has moved in favor of the poor.

4 Appendix: proofs of the theorems

Preliminary remarks:

1. Axiom 7 implies that $W(\mathbf{y}^1, \mathbf{y}^2, g, z)$ should be homogeneous of degree 0 in $\mathbf{y}^2$ and in $1+g$. Hence, we have that

$$W\left(\mathbf{y}^1, \mathbf{y}^2, 1, z\right) = P^*\left(\mathbf{y}^2/\left(1+g\right), 1, z\right) - P\left(\mathbf{y}^1, z\right). \tag{17}$$

2. Axiom 6 implies that $P^*(\mathbf{y}, 1, z) = P(\mathbf{y}, z)$, and we can therefore substitute $P^*(\mathbf{y}^2/(1+g), 1, z)$ in (17) by $P(\mathbf{y}^2/(1+g), z)$.

3. We will assume in the proofs below that we are comparing two distributions of the same size n. Achieving a common size for any two distributions can always be achieved by suitable replications of these two distributions. Such replications have no impact on W (by axiom 3), nor on the distribution functions (this can be readily verified by considering equation (3)). (Using distribution functions effectively normalizes population size to 1, which also has expositional advantages.)

4. We will generally assume that $\mathbf{y}^1$ and $\mathbf{y}^2$ have been anonymously ordered in non-decreasing values, which will guarantee that axiom 4 will be obeyed. Interchanging the values of any two y_i and y_j will also leave unchanged their distribution function $F(y)$.

5. Recall that by axiom 6, we have that $W(\mathbf{y}^1, \mathbf{y}^2, g, z) = P(\mathbf{y}^2/(1+g), z) - P(\mathbf{y}^1, z)$. For expositional simplicity, we may therefore work in the proofs with $F^2(y)$ and $\mathbf{y}^2$ instead of $\overline{F}^2(y)$ and $\mathbf{y}^2/g$, and we use $\Omega^1(1, z^+)$ instead of $\Omega^1(g, z^+)$. We can then reinterpret pro-poor judgements as the more general problem of comparing a poverty index across two distributions. The previous literature has considered this problem in an additive context (see Foster and Shorrocks (1988a,b) and Duclos and Makdissi (2004) for instance), but not to our knowledge in the context of the more general non-additive formulation considered in this paper.

4.1 Proof of Theorem 9

4.1.1 Sufficiency of condition (4):

If $F^2(z) \leq F^1(z)$, $\forall z \in [0, z^+]$, then $W(\mathbf{y}^1, \mathbf{y}^2, 1, z) \leq 0$, $\forall W \in \Omega^1(1, z^+)$.

Proof:

Denote first by $\omega^1(z^*)$ the subset of all $W(\cdot,\cdot,1,z^*)$ which belong to $\Omega^1(1,z^+)$ for some $z^* \leq z^+$. Since $F^2(y) \leq F^1(y),\, \forall y \in [0,z^*]$, it must be that $y_i^1 \leq y_i^2$ whenever $y_i^1 \leq z^*$.
(Suppose otherwise: if $y_i^1 > y_i^2$, then by (3) we will have that $F^2(y) > F^1(y)$ for all $y \in [y_i^2, y_i^1[$.)

Let $h = nF^2(z^*)$. By the focus axiom 2, we have that $P(y_1^2,\ldots,y_n^2) = P(y_1^2,\ldots,y_h^2,z,\ldots,z)$. Combining this to the monotonicity principle of Axiom 5, we necessarily have $\forall W \in \omega^1(z^*)$ that

$$\begin{aligned} P(y_1^1,\ldots,y_n^1) &\geq P(y_1^2,\ldots,y_n^1) \geq P(y_1^2,y_2^2,\ldots,y_n^1) \\ &\geq P(y_1^2,\ldots,y_h^2,\ldots,y_n^1) \geq P(y_1^2,\ldots,y_h^2,z,\ldots,z) \\ &= P(y_1^2,\ldots,y_n^2). \end{aligned} \tag{18}$$

Hence, $W(\mathbf{y}^1,\mathbf{y}^2,1,z^*) \leq 0\, \forall W \in \omega^1(z^*)$. This argument can be repeated for any other choice of $z^* \in [0,z^+]$. Therefore, the result must hold for all $W \in \Omega^1(1,z^+)$. ∎

4.1.2 Necessity of condition (4):

Only if $F^2(y) \leq F^1(y),\, \forall y \in [0,z^+]$, will $W(\mathbf{y}^1,\mathbf{y}^2,1,z) \leq 0\, \forall W \in \Omega^1(1,z^+)$.

Proof: Assume that $F^2(y) > F^1(y)$ for some range $[\underline{y},\overline{y}]$ with $\overline{y} \geq \underline{y} < z^+$. Then it must be that, for some i, $z^+ \geq y_i^2 < y_i^1$. (Assume that this is not the case: then, by (3), it must be $F^2(y) \leq F^1(y)$ for all $y \leq z^+$, which violates the above assumption.) Choose for $P(\cdot,z)$ the distribution function $F(z)$, with $z = y_i^2$. $F(z)$ belongs to the class $\Omega^1(1,z^+)$ since it obeys all of the relevant axioms and since $z \leq z^+$ by assumption. This particular function is, however, such that $F^2(y_i^2) - F^1(y_i^2) = i/n - F^1(y_i^2) > i/n - F^1(y_i^1) = 0$ since it was chosen such that $y_i^1 > y_i^2$. This exercise can be done for any other range $[\underline{y},\overline{y}]$ such that $\underline{y} < z^+$. Hence condition $F^2(y) \leq F^1(y)\, \forall y \in [0,z^+]$ is necessary for the result. ∎

4.2 Proof of Theorem 12

4.2.1 Sufficiency of condition (10):

If $D^2(y) \leq D^1(y),\, \forall y \in [0,z^+]$, then $W(\mathbf{y}^1,\mathbf{y}^2,1,z) \leq 0,\, \forall W \in \Omega^2(1,z^+)$.

Proof:

To prove sufficiency, we will work in steps. Each step creates a new distribution from the initial distribution F^1 through a series of beneficial Pigou-Dalton transfers.

Step 1

Let $\mathbf{y}^{1,0} = \mathbf{y}^1$. Let $\mathbf{z^0} = (z_1^0, z_2^0, \cdots, z_{2n}^0)$ be the $2n$-vector $\left(y_1^{1,0}, y_2^{1,0}, \cdots, y_n^{1,0}, y_1^2, y_2^2, \cdots, y_n^2\right)$ ordered in increasing value. Let $\Delta^j F(z) = F^{1,j}(z) - F^2(z)$ and $\Delta^j D(z) = D^{1,j}(z) - D^2(z)$, where $F^{1,j}(z)$ and $D^{1,j}(z)$ are the distribution and deficit functions for a vector $\mathbf{y}^{1,j}$. Since $D^2(y) \leq D^1(y)$ for all $y \in [0, z^+]$, it must then be that $\Delta D^0(y) \geq 0$ for all $y \in [0, z^+]$

First, note that we must have that $y_1^{1,0} \leq y_1^2$ since otherwise we would have that $\Delta^0 D(z) < 0$ for all $z \in [y_1^2, y_1^{1,0}]$. Hence, it must also be that $\Delta^0 F(z) \geq 0 \, \forall z \in [0, z_2^0]$.

Now, either we have that

- $z_2^0 = y_2^{1,0}$, in which case $\Delta^0 F(z_2^0) = 2/n > 0$,
- or that $z_2^0 = y_1^2$, in which case $\Delta^0 F(z_2^0) = 0$.

In either case, we have that $\Delta^0 F(z_2^0) \geq 0 \, \forall z \in [0, z_3^0[$.

Now define $z^{*,0}$ as the smallest z_i^0 such that $\Delta^0 F(z_i^0) < 0$. If $z^{*,0} > z^+$, then we can move directly to "***Final Step***" on page 20. Let $s = nF^2(z^{*,0})$. Hence, $z^{*,0} = y_s^2$. We will consider a series of a maximum of $s-1$ equalizing Pigou-Dalton transfers that will move the initial distribution from $F^{1,0}$ to $F^{1,1}$ while ensuring that $\Delta^1 F(z) \geq 0 \, \forall z \in [0, z^{*,0}]$.

Since $\Delta^0 F(z) \geq 0$ for all $z < y_s^2$, it must be that $y_i^2 \geq y_i^{1,0}, i = 1, \cdots, s-1$. Let $\tau = \min(z^+, y_s^{1,0} - y_s^2)$. τ must be strictly positive since by definition of s we have that $F^{1,0}\left(y_s^{1,0}\right) = F^2\left(y_s^2\right) = s/n$ and $s/n > F^{1,0}\left(y_s^2\right)$. Then, define a series of $s-1$ non-negative transfers to those in distribution $F^{1,0}$ with incomes lower than $y_s^{1,0}$:

$$
\begin{array}{l}
\tau_1 = \min\left(\tau, y_1^2 - y_1^{1,0}\right) \\
\tau_2 = \min\left(\tau - \tau_1, y_2^2 - y_2^{1,0}\right) \\
\tau_3 = \min\left(\tau - \tau_1 - \tau_2, y_3^2 - y_3^{1,0}\right) \\
\vdots \\
\tau_{s-2} = \min\left(\tau - \sum_{j=1}^{s-3}\tau_j, y_{s-2}^2 - y_{s-2}^{1,0}\right) \\
\tau_{s-1} = \tau - \sum_{j=1}^{s-2}\tau_j.
\end{array}
\tag{19}
$$

Note from the last line of (19) that $\tau = \sum_{j=1}^{s-1}\tau_j$. This is illustrated in Figure 7 for the case of $s = 3$.

Define a new $\mathbf{y}^{1,1}$ distribution as

$$\mathbf{y}^{1,1} = \left(y_1^{1,0} + \tau_1, \cdots, y_{s-1}^{1,0} + \tau_{s-1}, y_s^{1,0} - \tau, y_{s+1}^{1,0}, \cdots, y_n^{1,0}\right). \tag{20}$$

$\mathbf{y}^{1,1}$ is thus obtained from $\mathbf{y}^{1,0}$ by a series of Pigou-Dalton transfers. By Axiom 10 and by the focus and monotonicity axioms, we therefore have that

$$P(\mathbf{y}^{1,1}) \leq P(\mathbf{y}^{1,0}). \tag{21}$$

Note from (19) and (20) that $y_i^{1,1} \leq y_i^2$, $\forall i = 1, \cdots, s-2$ and that $y_s^{1,1} \leq y_s^2$ by construction.

Suppose that $y_s^{1,0} \leq z^+$. To verify that $y_{s-1}^{1,1} \leq y_{s-1}^2$, note from (19) that

$$\sum_{j=1}^{s-2} \tau_j = \min\left(\tau, \sum_{j=1}^{s-2} \left(y_j^2 - y_j^{1,0}\right)\right). \tag{22}$$

Hence, since $\tau_{s-1} = \tau - \sum_{j=1}^{s-2} \tau_j$, we have that

$$\tau_{s-1} = \begin{cases} 0 & \text{if } \tau < \sum_{j=1}^{s-2} \left(y_j^2 - y_j^{1,0}\right), \\ y_s^{1,0} - y_s^2 - \sum_{j=1}^{s-2} y_j^2 - y_j^{1,0} & \text{otherwise.} \end{cases} \tag{23}$$

In the first case, we have that $y_{s-1}^{1,1} = y_{s-1}^{1,0} < y_{s-1}^2$. For the second case, recall that $\Delta D^0(y) \geq 0$, $\forall y \leq z^+$, which implies in particular that $\Delta D^0(y_s^{1,0}) \geq 0$. From (9), this latter inequality implies that

$$\begin{aligned} \sum_{j=1}^{s} \left(y_s^{1,0} - y_j^{1,0}\right) & \geq \sum_{j=1}^{s} \left(y_s^{1,0} - y_j^2\right) + \sum_{j=s}^{n} \left(y_s^{1,0} - y_j^2\right) I\left[y_s^{1,0} > y_j^2\right] \\ & \geq \sum_{j=1}^{s} \left(y_s^{1,0} - y_j^2\right). \end{aligned} \tag{24}$$

Transforming (24)leads to:

$$\sum_{j=1}^{s} \left(y_j^{1,0} - y_j^2\right) \leq 0. \tag{25}$$

Combining (23) and (25), we obtain the desired result that $y_{s-1}^{1,1} \leq y_{s-1}^2$. Hence, $\Delta^1 F(z) \geq 0 \, \forall z \leq y_s^2$. When $y_s^{1,0} > z^+$, an exactly analogous demonstration shows that $\Delta^1 F(z) \geq 0 \, \forall z \leq z^+$.

We must now show that $\Delta D^1(z) \geq 0 \, \forall z \in [0, z^+]$. We start by showing that $\Delta D^0(z) = \Delta D^1(z) \, \forall z \geq y_s^{1,0}$. $D^2(z)$ is clearly unchanged by the above movement from $\mathbf{y}^{1,0}$ to $\mathbf{y}^{1,1}$. For all $z \geq y_s^{1,0}$, we obtain that

$$\begin{aligned}
D^{1,1}(z) &= n^{-1}\left(\sum_{j=1}^{s}\left(z - y_j^{1,1}\right) + \sum_{j=s+1}^{n}\left(z - y_j^{1,1}\right) I\left[z \geq y_j^{1,1}\right]\right) \\
&= n^{-1}\left(\sum_{j=1}^{s-1}\left(z - y_j^{1,0} - \tau_j\right) + \left(z - y_s^{1,0} + \tau\right) + \sum_{j=s+1}^{n}\left(z - y_j^{1,0}\right) I\left[z \geq y_j^{1,1}\right]\right) \\
&= n^{-1}\left(\sum_{j=1}^{s}\left(z - y_j^{1,0}\right) - \sum_{j=1}^{s}\tau_j + \tau + \sum_{j=s+1}^{n}\left(z - y_j^{1,0}\right) I\left[z \geq y_j^{1,1}\right]\right) \\
&= D^{1,0}(z).
\end{aligned} \tag{26}$$

The second line follows because $y_j^{1,1} = y_j^{1,0}$ for $j > s$ and the last line from $\sum_{j=1}^{s}\tau_j = \tau$.

It also follows that $\Delta^1 D(z) \geq 0 \, \forall z \leq y_s^{1,0}$. To see this, note that, for $z < y_s^{1,1}$, we have

$$\begin{aligned}
D^{1,1}(z) &= n^{-1}\sum_{j=1}^{s-1}\left(z - y_j^{1,1}\right) I(y_j^{1,1} < z) \\
&= n^{-1}\sum_{j=1}^{s-1}\left(z - y_j^{1,0} - \tau_j\right) I(y_j^{1,1} < z) \\
&\leq D^{1,0}(z) - n^{-1}\sum_{j=1}^{s-1}\tau_j I(y_j^{1,1} < z) \\
&\leq D^{1,0}(z),
\end{aligned} \tag{27}$$

and for $z \in [y_s^{1,1}, y_s^{1,0}]$, we have

$$\begin{aligned}
D^{1,1}(z) &= n^{-1}\left(\sum_{j=1}^{s-1}\left(z - y_j^{1,0} - \tau_j\right) + z - y_s^{1,1}\right) \\
&= n^{-1}\left(-\sum_{j=1}^{s-1}\tau_j + \sum_{j=1}^{s-1}\left(z - y_j^{1,0}\right) + z - y_s^{1,1}\right) \\
&\leq n^{-1}\left(-\sum_{j=1}^{s-1}\tau_j + \sum_{j=1}^{s-1}\left(z - y_j^{1,0}\right) + z - y_s^{1,0} + \tau\right) \\
&\leq D^{1,0}(z)
\end{aligned} \tag{28}$$

since $\tau = \sum_{j=1}^{s-1}\tau_j$.

We have therefore obtained a vector $\mathbf{y}^{1,1}$ such that $\Delta^1 D(z) \geq 0$ for all $z < z^+$ and such that $P(\mathbf{y}^{1,1}) \leq P(\mathbf{y}^{1,0})$.

End of Step 1 If $\Delta^1 F(z) \geq 0$ for all $z \leq z^+$, we move to ***Final Step***. If not, we move to ***Next Step*** by reordering $\mathbf{y}^{\mathbf{1,1}}$ if needed.

Next Step We proceed as for ***Step 1***, iteratively until the point is reached at which $\Delta^J F(z) \geq 0 \, \forall z \leq z^+$.

Final Step Say that the above procedure has taken J steps. By the focus axiom, we obtain that $P(\mathbf{y}^{1,J}) = P(\mathbf{y}^{1,J+1})$ where $\mathbf{y}^{J+1}$ is defined as

$$\mathbf{y}^{1,J+1} = \left(y_1^{1,J}, \cdots, y_{nF^1(z^+)}^{1,J}, z, \cdots, z\right). \tag{29}$$

Let

$$\mathbf{y}^{*2} = \left(y_1^2, y_2^2, \cdots, y_{nF^2(z^+)}^2, z, \cdots, z \right). \tag{30}$$

By the focus axiom, we have that $P(\mathbf{y}^{*2}) = P(\mathbf{y}^2)$. Since $\Delta^J F(z) \leq 0 \, \forall z \leq z^+$, by the monotonicity axiom and Theorem 9, and considering the recursive process started by (21), we have $P(\mathbf{y}^{*2}) \leq P(\mathbf{y}^{1,J+1})$. Hence, considering (21), we obtain

$$W = P(\mathbf{y}^2) - P(\mathbf{y}^1) \leq 0. \tag{31}$$

Since this is true for any of the P that obeys the conditions of Theorem 12, (31) must be true for all $W \in \Omega^2(z^+)$.

4.2.2 Necessity of condition (10):

Only if $D^2(y) \leq D^1(y), \ \forall y \in [0, z^+]$ will $W(\mathbf{y}^1, \mathbf{y}^2, 1, z) \leq 0, \ \forall W \in \Omega^2(1, z^+)$.

Proof:

As for the proof in 4.1.2, but choosing instead the deficit function as the P function. ∎

References

[1] Atkinson, A.B., B. Cantillon, E. Marlier and B. Nolan (2002), *Social Indicators - The EU and Social Inclusion*, Oxford University Press.

[2] Alesina, A. and D. Rodrik (1994), ”Distributive Policies and Economic Growth”, *Quarterly Journal of Economics* , **109**, 465–490.

[3] Bibi, Sami, and Jean-Yves Duclos (2003), ”Equity and Policy Effectiveness with Imperfect Targeting”, mimeo, CIRPÉE and Université Laval.

[4] Bourguignon, François (2003), ”The Poverty-Growth-Inequality Triangle”, paper prepared for a Conference on Poverty, Inequality and Growth, Agence française de développement, Paris, November.

[5] Bruno, M., L. Squire, and M. Ravallion (1998), ”Equity and Growth in Developing Countries: Old and New Perspectives on the Policy Issues”, in Tanzi, V. and K. Chu (eds.) *Income Distribution and High-Quality Growth*, Cambridge, MIT Press.

[6] Deininger, K. and L. Squire (1998), ”New ways of Looking at Old Issues: Inequality and Growth”, *Journal of Development Economics*, **57**, 259–287.

[7] Dollar, D. and A. Kraay (2002), ”Growth is Good for the Poor”, *Journal of Economic Growth*, **7**, 195–225.

[8] Duclos, J.-Y., and P. Makdissi (2004), ”Restricted and Unrestricted Dominance for Welfare, Inequality and Poverty Orderings”, *Journal of Public Economic Theory* **6**, 145–164.

[9] Eastwood R. and M. Lipton (2001), ”Pro-poor growth and pro-growth poverty reduction: meaning, evidence, and policy implications”, *Asian Development Review*, **19**, 1–37.

[10] Essama-Nssah, B. (2004), ”Framing Pro-Poor Growth within the Logic of Social Evaluation”, mimeo, Poverty Reduction Group, The World Bank.

[11] Foster, J.E., J. Greer and E. Thorbecke (1984), ”A Class of Decomposable Poverty Measures”, *Econometrica*, **52 (3)**, 761–776.

[12] Foster, J.E. and A. Sen (1997), ”On Economic Inequality after a Quarter Century”, in *On Economic Inequality*, (expanded edition), Oxford: Clarendon Press.

[13] Foster, J.E. and A.F. Shorrocks (1988a), ”Poverty Orderings”, *Econometrica*, **56**, 173–177.

[14] Foster, J.E. and A.F. Shorrocks (1988b), ”Poverty Orderings and Welfare Dominance”, *Social Choice Welfare*, **5**, 179–198.

[15] 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.

[16] Kakwani, N. and E. Pernia (2000), ”What is Pro Poor Growth?”, *Asian Development Review*, **18**, 1–16.

[17] Kakwani, Nanak, Shahidur Khandker, and Hyun H. Son (2003), ”Poverty Equivalent Growth Rate: With Applications to Korea and Thailand”, Economic Commission for Africa, mimeo.

[18] Klasen, Stephan (2003), ”In Search of the Holy Grail: How to Achieve Pro-Poor Growth?”, Discussion Paper #96, Ibero-America Institute for Economic Research, Georg-August-Universitt, Gttingen.

[19] McCulloch, N. and B. Baulch (1999), ”Tracking pro-poor growth”, ID21 insights #31, Sussex, Institute of Development Studies.

[20] Ravallion, M. (2002), ”Growth, Inequality, and Poverty: Looking beyond Averages”, mimeo, the World Bank.

[21] Ravallion, M. (2003), ”Competing concepts of inequality in the globalization debate”, mimeo, the World Bank.

[22] Ravallion, M. and G. Datt (2002), ”Why has economic growth been more pro-poor in some states of India than others?”, *Journal of Development Economics*, **68**, 381–400.

[23] Ravallion, M. and S. Chen (2003), ”Measuring Pro-Poor Growth”, *Economics Letters*, **78**, 93–99.

[24] Rawls, J. (1971), *A Theory of Justice*, Cambridge: MA: Harvard University Press.

[25] 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.

[26] Son, Hyun Hwa (2004), ”A note on pro-poor growth”, *Economics Letters*, **82**, 307–314.

[27] Spencer, B. and S. Fisher (1992), ”On Comparing Distributions of Poverty Gaps” *Sankhya: The Indian Journal of Statistics*, Series B, **54**, 114–126.

[28] United Nations (2000), *A Better World for All*, New York, United Nations.

[29] World Bank (2000), *The Quality of Growth*, New York, Oxford University Press.

[30] Zheng, B., J.P. Formby, W.J. Smith and V.K. Chow (2000), ”Inequality Orderings, Stochastic Dominance and Statistical Inference”, *Journal of Business and Economic Statistics*, **18**, 479-488.

Figure 1: Comparing distribution functions for first-order pro-poor growth

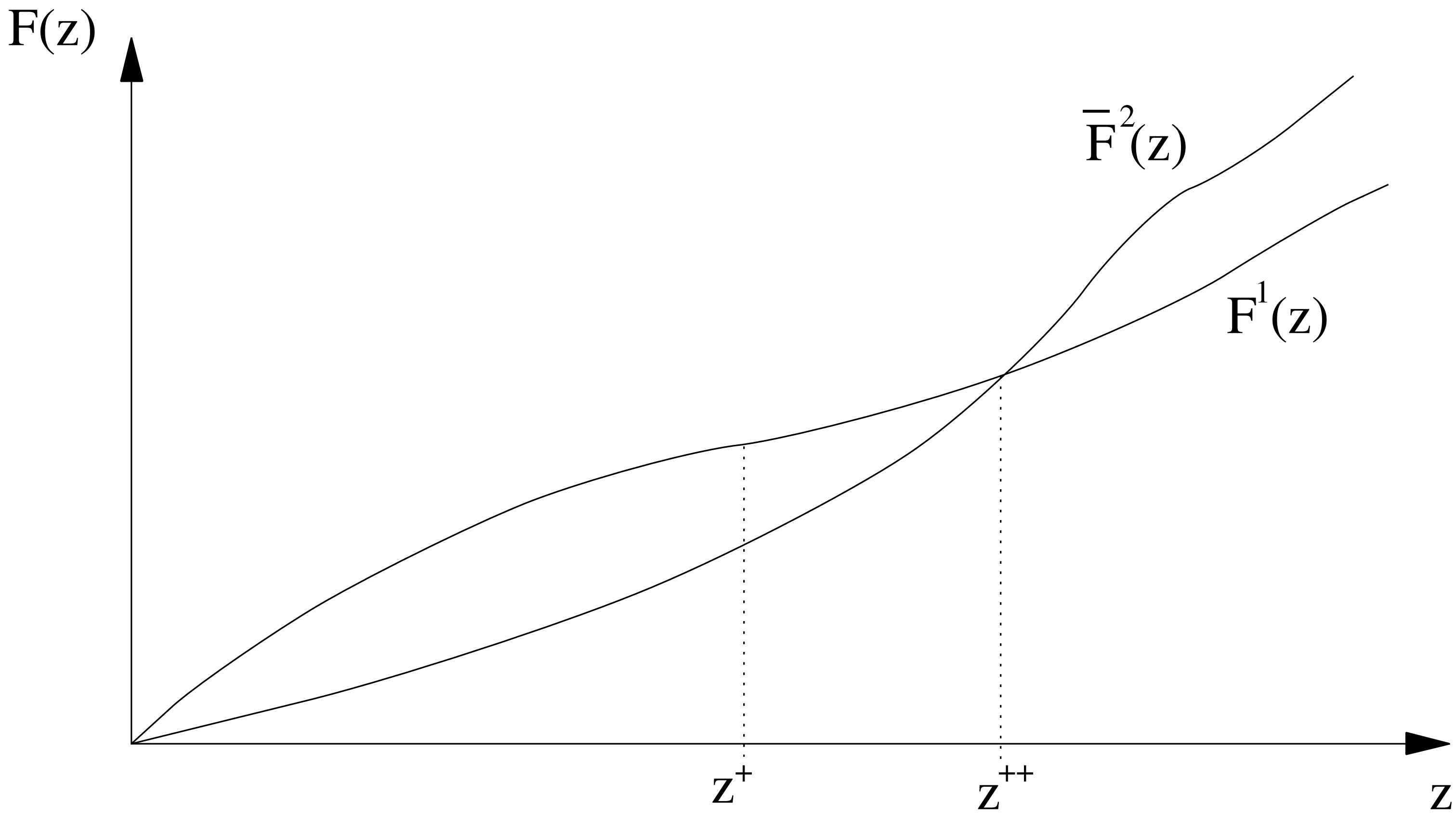

Figure 2: Ratios of quantiles and pro-poor standards

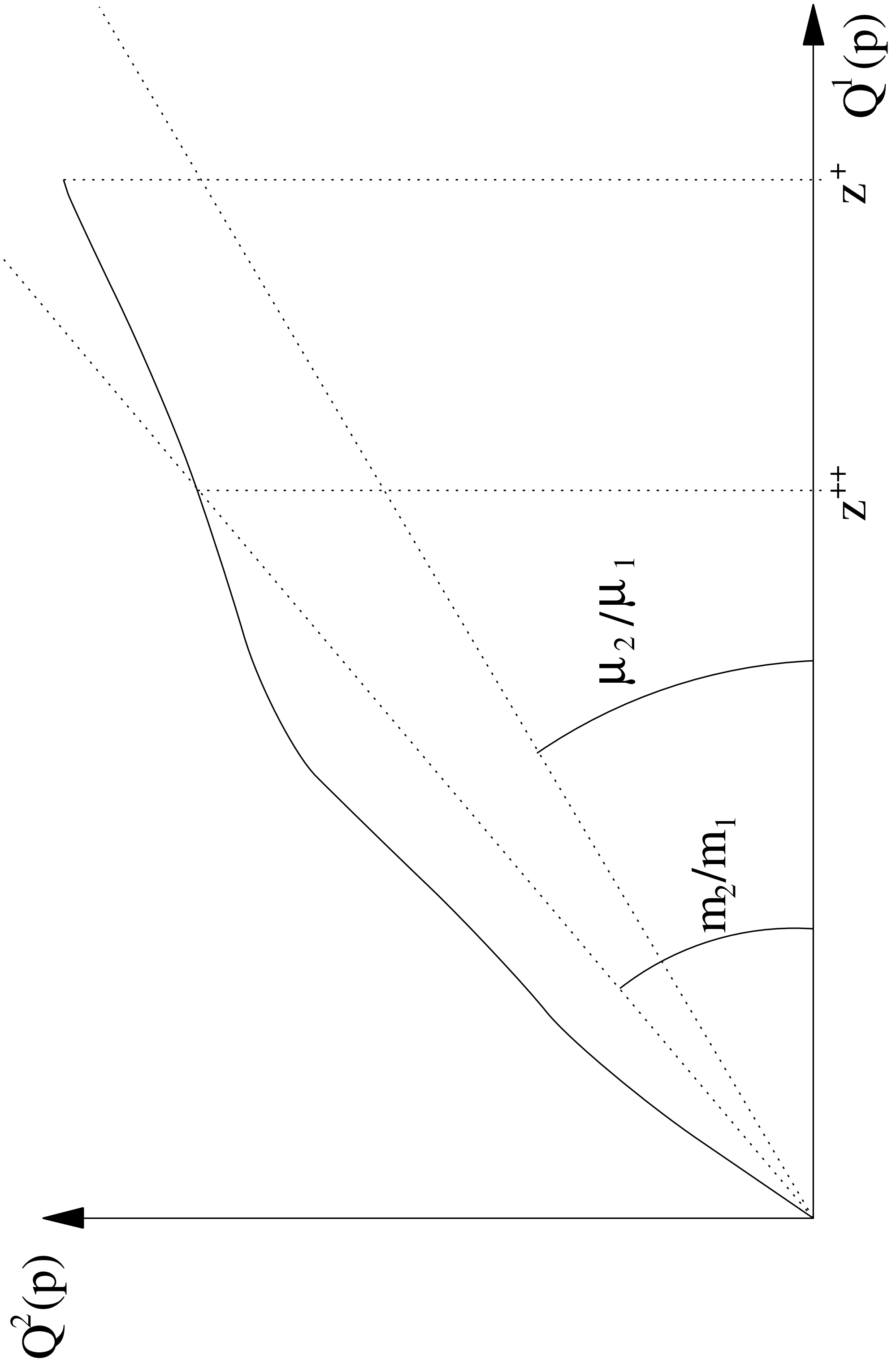

Figure 3: Income growth curves and pro-poor standards

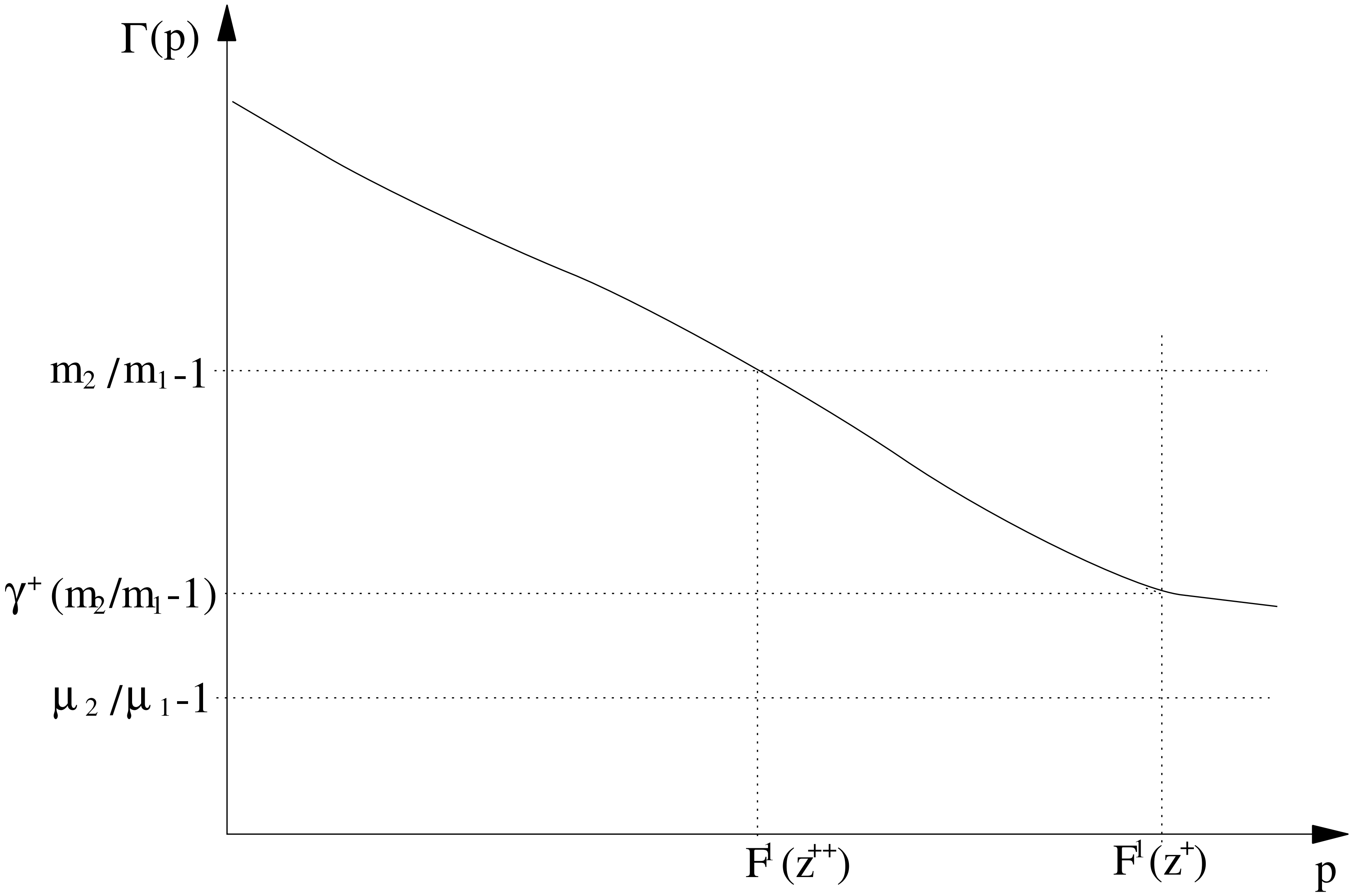

Figure 4: Income quantiles and poverty gaps

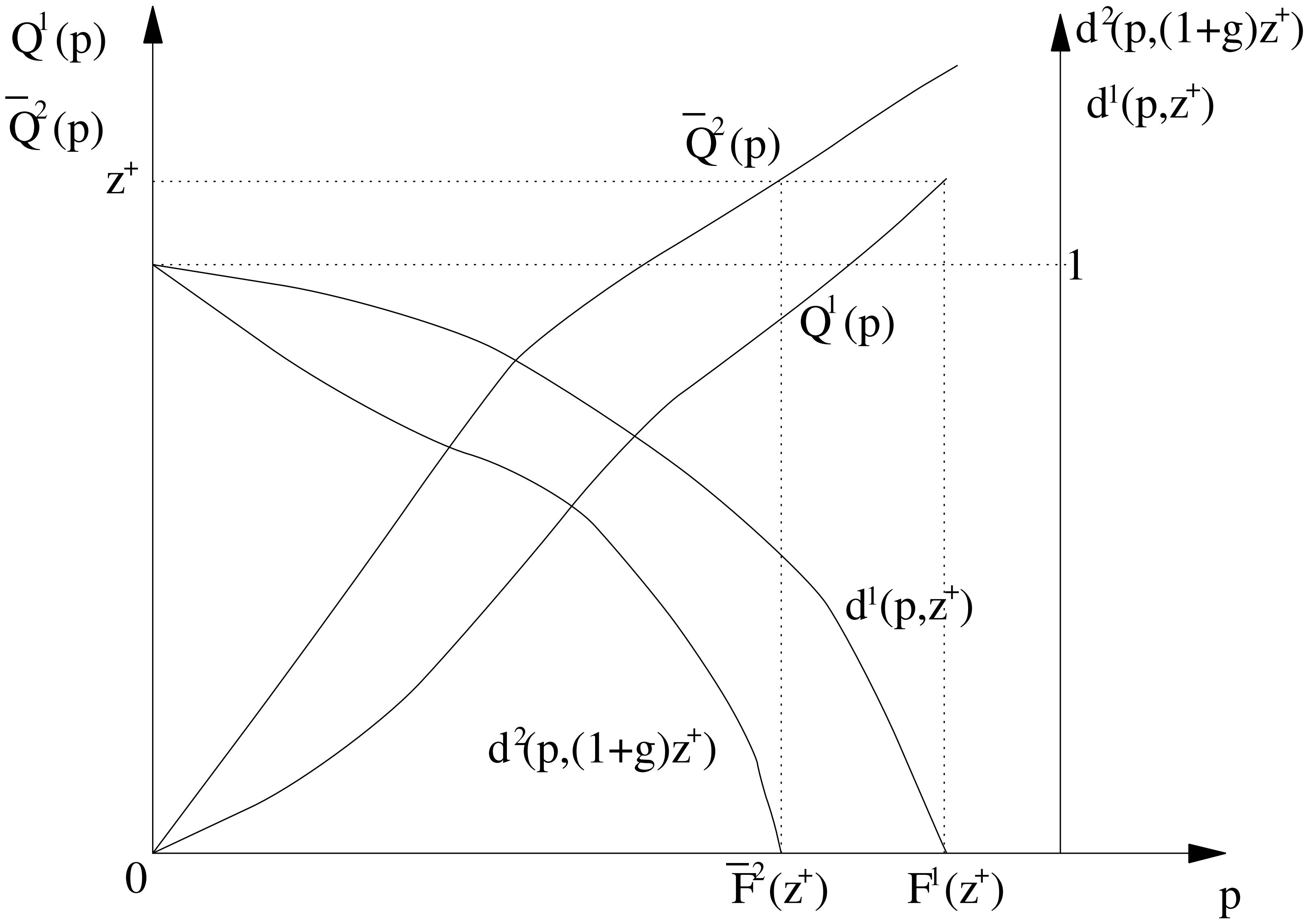

Figure 5: Cumulative poverty gap curves and second-order pro-poor growth

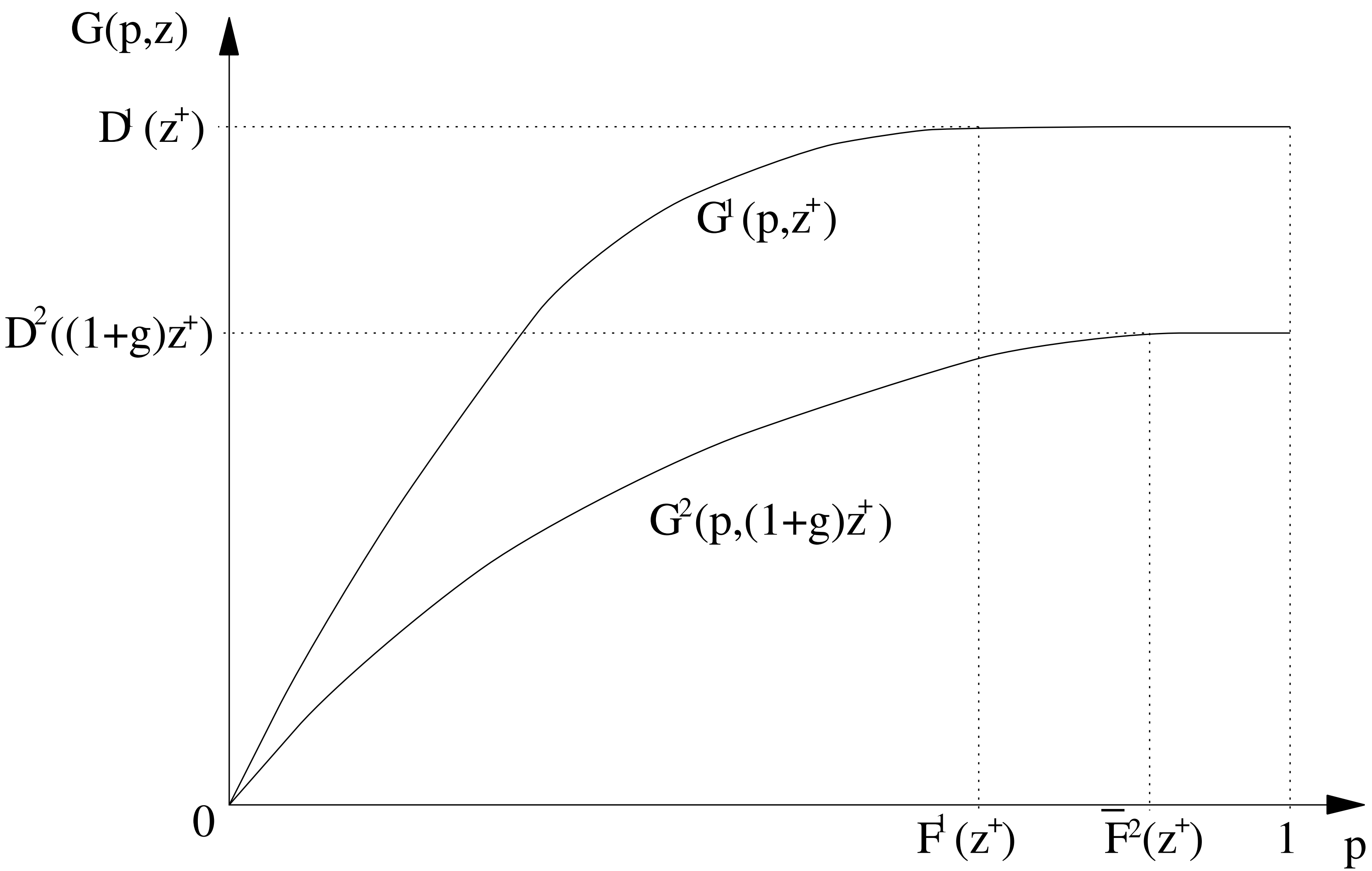

Figure 6: Cumulative income growth curves and second-order pro-poor growth

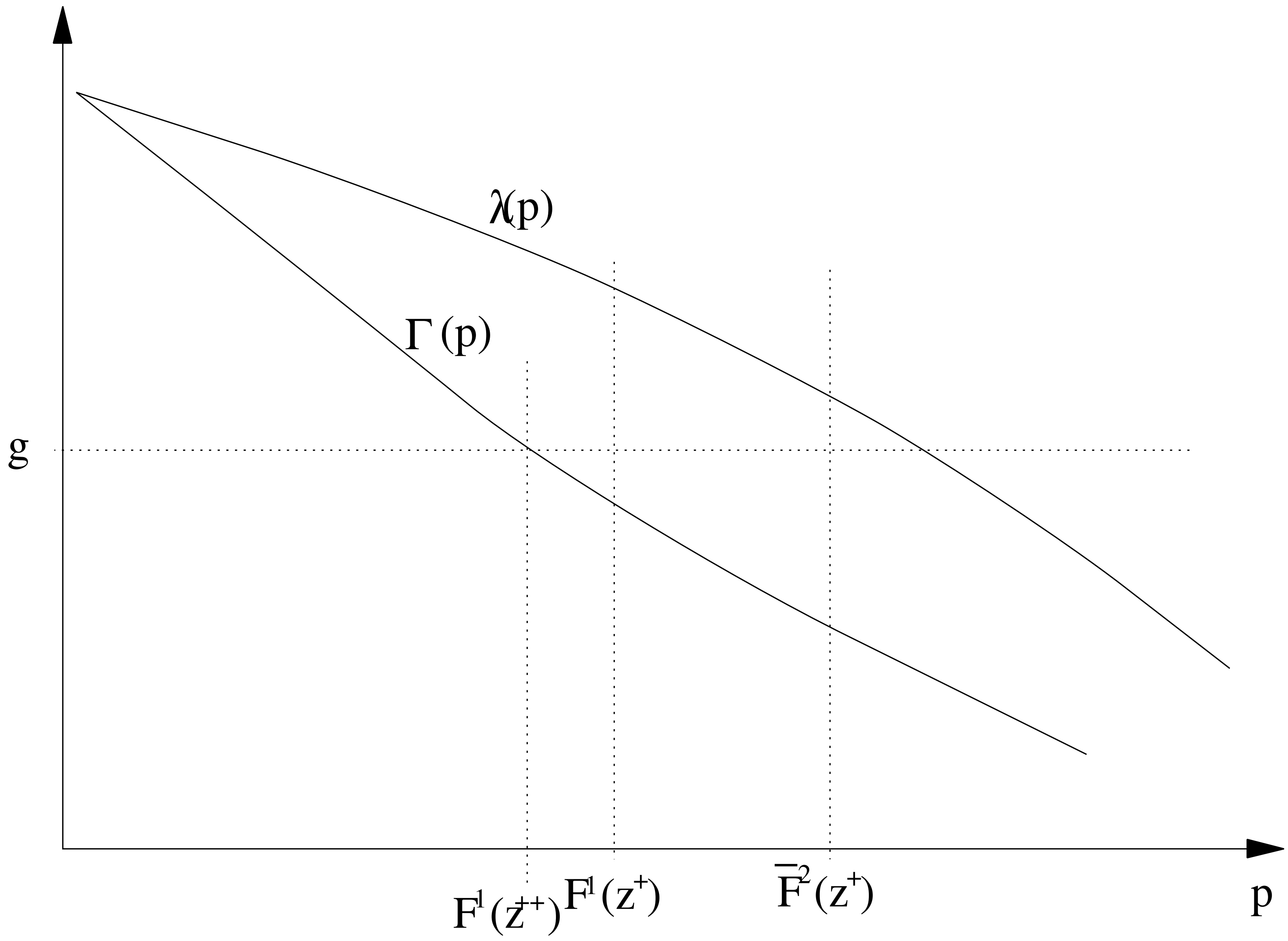

Figure 7: Two equalizing Pigou-Dalton transfers