A robust normative evaluation of India's performance in allocating risks of death

Nicolas Gravel[*], Abhiroop Mukhopadhyay[†] and Benoît Tarroux[‡]

This version January 20th 2008

JEL Classification numbers: C12, C25, D3, D63, D81, I1, I3, N35

Keywords: Poverty, welfare, risk, dominance, consumption health.

Abstract

This paper provides a robust normative evaluation of the recent evolution of Indians' exposures to health-related risks. Specifically, the paper compares the distribution of individuals' risks of death in India in 2002 with that of 1995 using two new ethically robust criteria. A probability of death is assigned to each individual as an estimated probit function of several explanatory variables, including the individual's district of residence. The criteria uses ranks distributions of individuals risks in the same way than would all Von Neumann-Morgenstern (VNM) social planners who respect, in the usual Pareto sense - individuals VNM preferences over state-dependent lotteries. The first criterion assumes that individuals' VNM utilities are increasing in money and value marginally money more when it is received in the bad state than in the good one. The second criterion makes the extra assumption that individuals are risk averse and have VNM utility functions that are more concave in the bad state than in the good one. It happens that the distribution of risk of death in India in 2002 dominates that of 1995 according to the second criterion but not according to the first.

[*]Université de la Méditerranée & IDEP-GREQAM, Centre de la Vieille Charité, 2, rue de la Charité, 13002 Marseille, France, nicolas.gravel@univmed.fr.

[†]Indian Statistical Institute, Delhi Center, 7, S. J. S. Sansanwal Marg, 11 0016 New Delhi - India, abhiroop@isid.ac.in.

[‡]CREM, Université de Caen, Campus 4, 19 rue Claude Bloch, BP 5186, 14032 Caen, Cedex 05, France, benoit.tarroux@unicaen.fr.

1. Introduction

The spectacular episode of economic growth experienced by India in the last twenty years has been the object of a significant effort of normative evaluation (e.g. see the contributions in the collective volume of Deaton & Kozel (2005) and Datt & Ravallion (2002)). Much of this effort has focused on pecuniary poverty and/or inequality and has done so by using rather specific indices (Gini and Coefficient of variation for inequality, headcount and poverty measures based on a specific poverty line for poverty). The use of specific inequality indices or poverty measures based on a given poverty line suffers from a lack of robustness that has been illustrated in many controversies surrounding the impact of Indian growth on poverty and inequality. For instance Deaton & Drèze (2005), using both the mean relative standard deviation and the log of the variance as inequality indices, conclude that income inequality in many Indian states as well as in the whole of India increased in the nineties. In contrast, other researchers, such as Bhalla (2005), conclude that income inequality has reduced when the measure of inequality used is the Gini coefficient. A similar lack of robustness concerns the choice of the poverty line in the differing conclusions obtained by Himanshu & Sen (2005) and Deaton (2005) on the extent to which headcount poverty in India has gone down during the same period. Whatever the appraisal one wants to perform, it is certainly preferable for the appraisal to be robust with respect to the choice of the poverty line or the poverty index.

Moreover, pecuniary poverty and inequality, important as they can be, are certainly not the only feature of the Indian economy that is relevant for normative evaluation. The ultimate goal of any economic system is to enable indeed individuals to fulfill better their goals, desires and ambitions (irrespective of what these may be). The pecuniary variables considered in conventional poverty or inequality evaluation are only contributing elements to these achievements. Other attributes such as health, education, public safety, political freedom or access to a cleaner environment, to mention just a few, are also important contributors to individuals success in life for which pecuniary variables like expenditures, wealth, or income are imperfect indicators. While this "multidimensionality" of the individuals' attributes that matter for normative evaluation has been long recognized from a theoretical standpoint (see e.g. Kolm (1977), Atkinson & Bourguignon (1982), Atkinson & Bourguignon (1987), Rawls (1971), Sen (1987) and Sen (1992)) it is only recently that these concerns have been incorporated in empirical studies (such as Crawford (2005), Duclos *et al.* (2006) and Gravel *et al.* (2007)). Can we say that the recent Indian growth experience has improved the distribution of relevant attributes among Indian citizens when one adopts such a *multidimensional* perspective?

In a recent paper (see Gravel & Mukhopadhyay (2007)), we propose a partial, but somewhat robust, *positive* answer to this question. Specifically, we compare, over the period 1987-2002, three samples of the joint distribution of *three* non-pecuniary individual attributes along with the conventional pecuniary individual consumption's expenditures. The non-pecuniary attributes considered are infant mortality, literacy and crime rates measured at the level of the individual's district of residence. These non-pecuniary attributes are interpreted as indicators of district level public goods. Comparing the three distributions using ethically robust dominance criteria developed along the lines of Atkinson & Bourguignon (1982), we conclude in Gravel & Mukhopadhyay (2007)) that *all ethical judgements* that are increasing and inequality averse with respect to an individual's achievement would agree that the distribution of individuals' achievements has improved in India over the period. This conclusion at least holds *if* one assumes that individuals' achievements depend *only* upon the above four attributes *and* that the dependence obeys relatively mild conditions discussed in Gravel & Mukhopadhyay (2007).

The non-pecuniary attributes considered in Gravel & Mukhopadhyay (2007) are all measured at the district level and are therefore common to all individuals living in the same district. Hence, these attributes are connected to individuals' achievements only through the assumption that all individuals of a district face the same risk of crime or of loosing their children and that they are all equally affected by the average district literacy rate. In the current paper, we provide another normative evaluation of the recent evolution of the distribution of individuals' achievements in India by focusing on purely individual attributes.

Specifically, we concentrate on an aspect of the individual's situation that has not been studied much in normative economics: exposure to risk. The protection against various kinds of risks that societies provide to their members is arguably an important ingredient for normative evaluation. An important risk faced by any individual is the potential loss of life. This risk is clearly affected by various things: access to competent health care, better information, quality of the environment (less dangerous roads, less dangerous workplace, etc.). Individuals are obviously very unequal in terms of their exposure to the risks of dying. How has this inequality evolved in India in the last fifteen year ? Is the distributions of exposure to risks of death among Indians better now than it was 10 years ago ? These are the questions that are answered in this paper.

We answer these questions by using robust normative dominance criteria developed in Gravel & Tarroux (2007) that apply precisely to distributions of risks among individuals. The risks handled by the approach of Gravel & Tarroux (2007) are those that can be described by a finite number of states, a probability distribution over these states, and a pecuniary outcome contingent upon the state. In the case of risks of death however, there are only two states - being alive or dead. Further, it can be assumed that the pecuniary outcome - say consumption - is independent of the state.

Assuming that individuals have Von Neuman-Mongenstern (VNM) preferences over these risks, and acknowledging that a *distribution of* these *individual risks* can be seen as a *socially risky situation*, Gravel & Tarroux (2007) derive *robust* empirically implementable criteria for comparing socially risky situations that coincide with the unanimity, taken over (suitably large) classes of individual VNM preferences, of all Pareto consistent social rankings that satisfy VNM properties. Because of Harsanyi's aggregation theorem (Harsanyi (1955)), we know that a VNM ranking of socially risky situations that respects, in the usual Pareto-sense, individual VNM preferences over risks *can* - but need not - be thought of as resulting from the comparisons of the *sum* of the individuals' VNM expected utility functions. Hence, a ranking of socially risky situations agreed upon by all social VNM rankings which respects an individuals' VNM rankings is nothing else than a ranking that commands unanimity over all sums of individual VNM expected utility functions taken in some suitable class.

Two empirically implementable criteria are characterized in Gravel & Tarroux (2007), each of which corresponds to a specific class of the VNM individual utility functions over which the unanimity of Paretian social rankings is sought. The first criterion coincides with the ranking that commands unanimity over all VNM and Pareto-consistent social planners who assume that individuals VNM utility functions are increasing with respect to consumption and are such that marginal utility of consumption is, weakly, lower if alive than if dead. The empirically implementable criterion that corresponds to the Paretian unanimity over this class of VNM utilities is what we call "Sequential Expected Headcount Poverty" (SEHP) dominance. According to this criterion, a distribution of risks A is better than distribution B if, for any poverty line, the *expected number of individuals* who will both be dead and below the line is no greater in A than in B and if the expected number of individuals who are below the line irrespective of whether they are alive or dead is also weakly smaller in A than in B. The sequential aspect of this criterion arises from the fact that in order for A to dominate B, one looks first at the expected number of poor people who will die and, in a second step, at the total expected number of poor that are either alive or dead. This criterion therefore gives priority to poverty among the dead as compared to poverty among those alive. This reflects of course the assumption made on the VNM utility function that the marginal utility of consumption is greater in the bad state than in the good one.

The second more restricted family of VNM utility functions considered satisfy, in addition to the above properties, the requirement that the marginal utility of income is, in every state, decreasing with consumption - reflecting risk aversion - and that this decrease in the marginal utility of consumption is more important in the dead state than in the alive one. Put differently, it is assumed that the state dependent utility function is concave in every state and is "more concave" in the dead state than in the alive one. It is then shown in Gravel & Tarroux (2007) that the implementable criterion that corresponds to Paretian unanimity over this class of VNM utility function, assuming again VNM preference from the part of the social evaluator, is the so-called "Sequential Expected Poverty Gap" (SEPG) dominance. This criterion works just like the SEHP one, but with the poverty gap rather than headcount poverty used as the relevant poverty measure.

In the present paper, we use these two criteria to compare, between the years 1995 and 2002, the joint distributions of individual consumptions (as obtained from the National Sample Survey Organization (NSSO) of India in the 52nd and 58th round respectively) and individual's probability of dying. In 1995, a probability of dying is assigned to each individual of the 52nd round on the basis of his or her observable characteristics (gender, age, district of residence, social group, schooling, etc.) as predicted from the estimation, on the 52nd round of the Morbidity and Health Care data of the NSSO, of a probit model explaining the living status of the individual as a function of some observable. While a similar procedure is adopted for the year 2002, we use the 60th round of Morbidity and Health Care Data (2004) rather than the 58th round to estimate the probit model explaining death by observable individual characteristics. The reason for not choosing the

58th round is that the latter does not contain information on the number of dead in each household.

We then show that while there is no dominance of 2002 over 1995 using the more robust SEHP criterion, there is dominance if one uses the more discriminatory SEPG dominance criterion. Of course, the gain of discriminatory power of this criterion is achieved at the cost of making extra assumptions on the individual VNM utility functions. Yet, the range of individual VNM preferences that are covered by these assumptions is quite large. For this reason, it can be said that, in a somewhat strong sense, the distribution of exposures of Indian citizens to risks of dying has improved over the period 1995-2002. At least any ethics that ranks social lotteries by means of VNM preferences, and that respects individuals VNM preferences for these risks, would agree with this statement if it can be assumed that individuals are risk averse and are more risk averse when dead than when alive.

In our view, these results illustrate the usefulness of the dominance methodology for addressing policy issues with a redistributive impact. Dominance analysis is sometimes dismissed as being a nice theoretical construction with little practical interest. These arguments are held on the basis of both the incompleteness of the dominance criteria - e.g. they sometimes fail to rank distributions - and their alleged complexity of implementation. This skepticism is held even more strongly when dominance criteria are used to compare distributions of several attributes, be it simply because the incompleteness of the criteria is even more severe in that case than in the standard one-dimensional one. What the results of this paper and others suggest is that this skepticism is misplaced. In many instances - and the recent growth experience in India is one of them - the multidimensional generalizations of dominance criteria of the kind used in this paper are capable of providing firm conclusions on the impact of specific policies on the distribution of well-being across individuals. The ethical robustness of the conclusions thus achieved is, of course, not to be disdained.

The plan of the rest of the paper is as follows. In the next section, we present the theoretical criteria used to perform the comparisons. Section 3 discusses the data, the statistical methodology and the results of the comparisons. Section 4 concludes.

2. Presentation of the criteria

2.1. Theoretical criteria

Assume there are n individuals indexed by i. Each individual is exposed to a risk of death and can therefore fall into *two* mutually exclusive states: "dead" (d) and "alive" (a). An individual receives, as a function of the state in which he or she falls, a non-negative pecuniary consequence which we shall refer to as "income". Of course we do not preclude *a priori* the possibility that individuals also value the states in a non-pecuniary fashion. Having 10 000 rupees is obviously not the same thing when alive and when dead. We call *socially risky situation* a specific pattern of exposure of individuals to risks of death.

Formally, such a socially risky situation is depicted as a probability distribution - or a *lottery* - p on the set $\mathbb{X} = (\{a, d\} \times \mathbb{R}_+)^n$ of all vectors of state-income pairs, one such pair for each individual. A typical element x of $\mathbb{X}$ is given by:

$$x = (s_1, y_1, ...s_n, y_n),$$

where, for $i = 1, ..., n$, $s_i \in \{a, d\}$ denotes a state in which i falls and y_i denotes i's income in that state. Hence $p(s_1, y_1, ...s_n, y_n)$ is interpreted as the joint probability that individuals $i = 1, ..., n$ be in state s_i and receives in that state an income y_i. Denoting by $(s_i, y_i; s_{-i}, y_{-i})$ the vector of state-income pairs where individual i gets the pair $(s_i, y_i) \in \{a, d\} \times \mathbb{R}_+$, and all individuals other than i get the vector of state-income pairs $(s_{-i}, y_{-i}) \in (\{a, d\} \times \mathbb{R}_+)^{n-1}$, we further require lotteries to satisfy the following condition:

Condition 1. *For every individual i, if $p(s_i, y_i; s_{-i}, y_{-i}) > 0$ for some $(s_i, y_i) \in \{a, d\} \times \mathbb{R}_+$ and $(s_{-i}, y_{-i}) \in (\{a, d\} \times \mathbb{R})^{n-1}$, then $p(s_i, y'_i; s'_{-i}, y'_{-i}) = 0$ for all $y'_i \in \mathbb{R}_+$ such that $y'_i \neq y_i$ and $(s'_{-i}, y'_{-i}) \in (\{a, d\} \times \mathbb{R})^{n-1}$.*

This condition says that lotteries have finite support, and are such that they never assign a positive probability to two different incomes received by an individual in a given state, no matter what the income and state of other individuals are. Let $\mathbb{L}$ be the set of all probability distributions on $\mathbb{X}$ that satisfy condition 1. Since a lottery p in $\mathbb{L}$ assigns a positive probability to a unique income for individual i in every state in which i can fall with positive probability, we denote by $y^p_{\sigma i}$ the value of this unique income that is given positive probabilities by p in the state $\sigma \in \{a, d\}$. Moreover, since for every individual and state achieved with positive probability there is a unique income received by the individual in that state, we can think of a lottery p in $\mathbb{L}$ as defined only on the set $\{a, d\}^n$ of combinations of individual states. Accordingly, we denote by $p(s)$ the (joint) probability of the individuals being in the configuration of states $s = (s_1, ..., s_n)$.

Every individual i has a VNM preference ordering $\succsim_i$ on $\mathbb{L}$, with asymmetric and symmetric factors $\succ_i$ and $\sim_i$ respectively. This means that there exists a utility function $\Phi_i : (\{a, d\} \times \mathbb{R}_+)^n \to \mathbb{R}$ such that, for every lotteries p and q in $\mathbb{L}$:

$$p \succsim_i q \Leftrightarrow \sum_{s \in \{a,d\}^n} p(s)\Phi_i(s_1, y^p_{s_1 1}, ..., s_n, y^p_{s_n n}) \geq \sum_{s \in \{a,d\}^n} q(s)\Phi_i(s_1, y^q_{s_1 1}, ..., s_n, y^q_{s_n n}). \tag{2.1}$$

We further assume that individuals are selfish and, therefore, only care about the state in which they fall and the income they get in that state, and that they have the *same* selfish preference. In order to write this condition formally, we notice that any lottery p in $\mathbb{L}$ can be viewed as a lottery on $\{a, d\}^n$. This induces an individual i's (marginal) binary lottery p_i on $\{a, d\}$,which is defined, for $\sigma \in \{a, d\}$, by:

$$p_i(\sigma) = \sum_{\{s \in \{a,d\}^n : s_i = \sigma\}} p(s).$$

For notational convenience, and using the fact that probabilities sum to 1, we write p_i instead of $p_i(a)$ and $1 - p_i$ instead of $p_i(d)$. With this piece of notation, the assumption that individuals have the same selfish VNM preference means that there exists a function, $U : \{a, d\} \times \mathbb{R}_+ \to \mathbb{R}$, such that, for every i and every lotteries p and q in $\mathbb{L}$:

$$p \succsim_i q \Leftrightarrow p_i U(a, y^p_{ai}) + (1-p_i)U(d, y^p_{di}) \geq q_i U(a, y^q_{ai}) + (1-q_i)U(d, y^q_{di}). \tag{2.2}$$

Also for notational convenience, we write, for every $y \in \mathbb{R}$, and $\sigma \in \{a, d\}$, $U_\sigma(y) = U(\sigma, y)$.

Lotteries in $\mathbb{L}$ are compared by a social ranking $\succsim$ (with asymmetric and symmetric factors $\succ$ and $\sim$ respectively) that satisfies the VNM properties and respects the weak Pareto principle with respect to individual preferences. In the same fashion as in (2.1), the first property means that there exists a function $\Phi : (\{a, d\} \times \mathbb{R}_+)^n \to \mathbb{R}$ such that, for every lottery p and

q in $\mathbb{L}$,

$$p \succsim q \Leftrightarrow \sum_{s\in\{a,d\}^n} p(s)\Phi(s_1, y^p_{s_11}, ..., s_n, y^p_{s_nn}) \geq \sum_{s\in\{a,d\}^n} q(s)\Phi(s_1, y^q_{s_11}, ..., s_n, y^q_{s_nn}). \tag{2.3}$$

The second property requires $\succsim$ to be such that, for two lotteries p and q in $\mathbb{L}$, if $p \succ_i q$ for all i, then $p \succ q$.

By virtue of a version of Harsanyi's aggregation theorem (Harsanyi (1955)) due to Weymark (1993), any VNM social ordering $\succsim$ of $\mathbb{L}$ that respects the weak Pareto principle can be written as the sum of the individual's expected utility representations of their VNM preferences. That is to say, if $\succsim$ is VNM and satisfies the Pareto principle, it can be written as:

$$p \succsim q \Leftrightarrow \sum_{i=1}^{n}[p_iU_a(y^p_{ai}) + (1-p_i)U_d(y^p_{di})] \geq \sum_{i=1}^{n}[q_iU_a(y^q_{ai}) + (1-q_i)U_d(y^q_{di})]. \tag{2.4}$$

The criteria used in this paper coincide with the ranking of lotteries in $\mathbb{L}$ that would be agreed upon by *all* social orderings that can be written as per (2.4) for a large class of individual expected utility functions. Specifically, we define as follows the notion of normative dominance with respect to a class of expected utility functions.

Definition 1 (Normative dominance). *Social risk p normatively dominate social risk q for a class $\mathbb{U}$ of utility functions defined on $\{a, l\} \times \mathbb{R}_+$, denoted $p \succsim_{\mathbb{U}} q$, if for all functions U in the class we have,*

$$\sum_{i=1}^{n}[p_iU_a(y^p_{ai}) + (1-p_i)U_d(y^p_{di})] \geq \sum_{i=1}^{n}[q_iU_a(y^q_{ai}) + (1-q_i)U_d(y^q_{di})].$$

Assuming, for simplification, the differentiability of U_σ and denoting by U'_σ and U''_σ its first and second order derivative respectively (for $\sigma = a, d$), the classes of expected utility functions are the following:

$\mathbb{U}_1 = \{U : \{a, d\} \times \mathbb{R}_+ \to \mathbb{R} : U'_d(y) \geq U'_a(y) \geq 0$ and $U_a(y) \geq U_d(y)\ \forall\ y \in \mathbb{R}_+ \}$,

$\mathbb{U}_2 = \mathbb{U}_1 \cap \{U : \{a, d\} \times \mathbb{R}_+ \to \mathbb{R} : U''_d(y) \leq U''_a(y) \leq 0$ for all $y \in \mathbb{R}_+\}$.

The class $\mathbb{U}_1$ contains therefore all utility functions that are, in every state, *increasing with income* and satisfy the additional requirements that, *i*) for any given income level, the utility enjoyed alive is no smaller than that enjoyed when dead and

ii) the marginal utility of income is weakly larger when dead than when alive.

The class $\mathbb{U}_2$ contains all functions that belong to $\mathbb{U}_1$ and satisfy the additional requirements that:

iii) the marginal utility of income is decreasing, in every state, with income and

iv) the decrease in the marginal utility of income is weakly more important when dead than when alive.

It requires obviously some imagination to think about the relationship between utility when dead and money. Yet one must recall that the utility functions U_a and U_d defined by the classes $\mathbb{U}_1$ and $\mathbb{U}_2$ describe the *ex ante* preference of an individual who is facing a risk of death when alive. Specifically, the function U_d describes the evaluation made by the individual alive of the money received if he or she were to die. It is in this sense not implausible that the individual would value more at the margin a rupee received in the bad state than in the good one (class $\mathbb{U}_1$), or would have a decrease in marginal utility that is more important in the bad state than in the good one (class $\mathbb{U}_2$).

2.2. Implementable criteria

The first implementable criterion that is considered is the Sequential Expected Headcount Poverty (SEHP) dominance criterion. It is formally defined as follows.

Definition 2 (Sequential Expected Headcount Poverty dominance). *For every p and $q \in \mathbb{L}$, we say that p SEHP dominates q, denoted $p \succsim_{SEHP} q$, if for every poverty line t,*

$$\sum_{\{i:y^p_{di}\leq t\}} (1-p_i) \leq \sum_{\{i:y^q_{di}\leq t\}} (1-q_i), \tag{2.5}$$

and,

$$\sum_{\{i:y^p_{di}\leq t\}} (1-p_i) + \sum_{\{i:y^p_{ai}\leq t\}} p_i \leq \sum_{\{i:y^q_{di}\leq t\}} (1-q_i) + \sum_{\{i:y^q_{ai}\leq t\}} q_i. \tag{2.6}$$

In words, socially risky situation p dominates socially risky situation q for the SEHP criterion if, no matter how one defines the poverty line, the *expected numbers of individuals who will be dead and poor for this poverty line* is no greater in p than in q (condition (2.5)) and (condition (2.6)), if the expected number of individuals who are poor in either state is also smaller, weakly, in p than in q. The importance of the requirement, expressed in condition (2.5), that headcount poverty be calculated, for those who are dead, *in expectation only* is worth stressing. If one uses a poverty line sufficiently large, it is easy to see that condition (2.5) requires the expected number of dead individuals to be weakly smaller in the dominating distribution than in the dominated one. If, as will be the case here, individual are assumed to receive the same income in the two states, then condition (2.6) amounts to requiring headcount poverty irrespective of the state to be lower in the dominating distribution than in the dominated one for all poverty lines. Further discussion of this criterion is provided in Gravel & Tarroux (2007).

The second empirically implementable dominance criterion used in this paper is the analogue of SEHP dominance, but using the poverty gap instead of headcount poverty as a measure of poverty. For this reason it is called the Sequential Expected Poverty Gap (SEPG) criterion. In order to define this criterion, we denote for any income y and poverty line t, the *poverty gap* $P(t, y)$ by,

$$P(t, y) = \max(t - y, 0]. \tag{2.7}$$

The poverty gap is, as usual, interpreted as the minimal amount of income that is required to get a person with an income of y out of poverty when poverty is defined as falling short of having an income of t. With this notation, we define sequential expected poverty gap as follows.

Definition 3 (Sequential Expected Poverty Gap dominance). *For every p and $q \in \mathbb{L}$, we say that p SEPG dominates q, denoted $p \succsim_{SEPG} q$ if, for every poverty line t,*

$$\sum_{i=1}^{n}(1 - p_i)P(t, y^p_{di}) \leq \sum_{i=1}^{n}(1 - q_i)P(t, y^q_{di}), \tag{2.8}$$

$$\sum_{i=1}^{n}[(1 - p_i)P(t, y^p_{di}) + p_i P(t, y^p_{ai}) \leq \sum_{i=1}^{n}[(1 - q_i)P(t, y^q_{di}) + q_i P(t, y^q_{ai}), \tag{2.9}$$

and

$$\sum_{i=1}^{n} p_i \geq \sum_{i=1}^{n} q_i. \tag{2.10}$$

In words, a socially risky situation p dominates situation q for the sequential poverty gap dominance criterion if the *expected amount of money* that is required to eliminate poverty *among all dead people* is lower in p than in q for every conceivable poverty line (condition (2.8)) and if the expected amount of money that is required to eliminate poverty *in either state* is also weakly smaller in p than in q for every poverty line. Notice that the criterion of SEPG dominance also requires – as expressed in condition (2.10) – that the expected number of people alive be no smaller in the dominating distribution than in the dominated one. It should also be stressed that, just like with the SEHP criterion, if an individual receives the same income in the two states, then condition (2.9) amounts to requiring poverty gap irrespective of the state to be smaller in the dominating distribution than in the dominated one for all poverty lines.

We now recall the two equivalences, established in Gravel & Tarroux (2007), between each of these two implementable criteria on the one hand and the unanimity of all anonymous Paretian and VNM social rankings who assume that individuals' VNM utility functions are in each of the two classes $\mathbb{U}_1$ and $\mathbb{U}_2$ on the other. These two equivalence provide indeed the normative justification for using the SEHP and the SEPG criterion to compare alternative distributions of risks.

Theorem 1. *Let p and q be two socially risky situations in $\widehat{\mathbb{L}}$. Then:*
(1) $p \succsim_{\mathbb{U}_1} q$ if and only if $p \succsim_{SEHP} q$ and
(2) $p \succsim_{\mathbb{U}_2} q$ if and only if $p \succsim_{SEPG} q$.

3. Empirical implementation

3.1. Data

In order to use the criteria described in the previous section for appraising the temporal evolution of the exposure to risks of Indian citizens, data on an individual's probability of death as well as on income in either states are needed. It is difficult to appraise the impact of death on an individual's income. For this reason, we consider that individual income is independent from the state - life or death - in which the individual is. Our data sources for individuals pecuniary consequences are households' consumption figures as taken from the 52nd (1995-96) and 58th rounds (2002) of consumption expenditure surveys of the National Sample Survey Organization (NSSO).

Individual consumption expenditures are derived from household surveys using the Oxford equivalent scales and are in 2002 urban rupees (see Gravel & Mukhopadhyay (2007) for a more detailed discussion of NSS data and the choice of these two years). While these surveys are conducted for the whole country, we exclude Jammu and Kashmir as well as all the North Eastern States of India because of well-known problems with data collection in these area. Thus our analysis is conducted over 397 districts in 1995-96 and the 434 districts in 2002 that correspond to the same geographical area (the larger number of districts in 2002 are due to formation of new districts by subdivision of some of the 1995-96 districts).

To each individual below the age of 65 in the consumption expenditure data set, we assign for the two periods a probability of death[1]. To do so, we first estimate a probability model of death. We use for this sake the data collected from the Morbidity and Health Care surveys conducted by the NSSO for the years 1995-96 (469,451 individuals) and 2004 (239,423 individuals)[2]. These surveys include information on the household history of deaths in the previous 365 days (the gender and age of the deceased member are reported). We drop individuals from this data set that do not live in the 397 districts in 1995-96 and the corresponding 434 districts in 2004.

Using a probit specification, we model the probability of death as a function of the individual's age, the square of this age, his or her gender, his or her social group (scheduled caste and tribe) and the maximum level of education achieved in the household. We also consider all district dummies in order to account for the fact that the district of residence affects quite significantly the individual probability of death. We use the information on the individuals who died over the last 365 days as well as those currently alive in a household to estimate the probit model. The dependent variable for our analysis takes the value 1 for the individual who died in the last 365 days and 0 for the individual who was alive at the end of this period. As is standard in probit estimation, we assume the unobservable part of the latent health of an individual follows a normal distribution. We estimate the probit model separately for the individuals sampled in rural and urban areas and separately for each year

[1] We drop individuals above 65 as we do not want to include the age group where there is a greater probability of dying due to natural causes.

[2] The health data collected in 2004 (60th round) are the closest available data for imputation of the probablity of death for individuals sampled in 2002.

Except for the district dummies, the results of the probit are reported in Table 1. These results should be seen more as descriptive than causative. Also, it is worth mentioning that the results of the different columns are not comparable since the base category changes in each probit estimation. Yet, the results are generally in line with what could be expected. Age initially lowers the probability of death. However the positive second derivative implies that this trend slows down after a certain age. Being a female reduces the probability of death. Being from a disadvantaged social group increases the probability of mortality (viz. the omitted category in each year and sector) except in the case of urban sector in 2004. Education happens significantly to reduce the probability of death only if the level achieved is sufficiently high (as compared to a household whose most educated member is illiterate). Yet the effect is much stronger in 1995-1996 than in 2004 where only one category of education level happens to significantly reduce the probability of death.

Table 1: Probit: Probability of Mortality

Dependent Variable: ***Dead=1 if individual is dead***	**1995-96**		**2004**	
	Rural	**Urban**	**Rural**	**Urban**
Age	-0.00037* (0.00)	-0.0002* (0.00)	-0.0002* (0.00)	-0.0001* (0.00)
Square of Age	0.000007* (0.00)	0.000005* (0.00)	0.000004* (0.00)	0.0000036* (0.00)
$D_{Female}=1$ (Omitted Category: Male)	-0.0004* (0.036)	-0.001* (0.00)	-0.0006* (0.005)	-0.0008* (0.003)
Social Group **(Omitted Category: Non SC and ST households)***				
$D_{ST}=1$	0.0015* (0.00)	0.002* (0.00)	0.001* (0.038)	-0.0006 (0.54)
$D_{SC}=1$	0.0005* (0.036)	0.0003 (0.22)	0.0005 (0.12)	0.0002 (0.578)
$D_{OBC}=1$			0.00007 (0.79)	0.0002 (0.563)
Maximum Education Level in the Household **(Omitted Category: Illiterates)**				
$D_{\text{Literate without Formal Schooling}}$	0.0006 (0.52)	-0.0002 (0.87)	0.0012 (0.474)	0.00011 (0.944)
$D_{\text{Below Primary}}$	-0.0002 (0.40)	-0.000004 (0.99)	0.0003 (0.539)	-0.0004 (0.55)
D_{Primary}	-0.0004 (0.16)	-0.0003 (0.37)	0.0003 (0.395)	-0.0002 (0.711)
D_{Middle}	-0.0006* (0.02)	-0.0006* (0.09)	-0.00004 (0.917)	-0.0005 (0.343)
D_{Sec}	-0.0009* (0.004)	-0.0012* (0.00)	-0.0007* (0.10)	-0.0006 (0.265)
$D_{\text{Higher Secondary}}$	-0.0006 (0.14)	-0.0010* (0.004)	-0.0003 (0.492)	-0.0006 (0.303)
$D_{\text{Graduate and Above}}$	-0.0008* (0.06)	-0.0017* (0.00)	-0.001 (0.13)	-0.0016* (0.001)
Not shown on this table: District Dummies				
Number of Observations	281786	187665	162800	76623
LR χ^2	1369.16 (0.00)	1527.25 (0.00)	679.56 (0.00)	311.57 (0.00)

For 1995-96: the Social Group OBC was not reported separately from the General Category
(robust p values reported in parentheses below the coefficients)
**: Significant at 10% level of significance*

We perform our distributional analysis on samples taken from the consumer's expenditure survey. These samples contain 126,903 individuals in 2002 and 199,622 in 1995-96. While the number of individuals in the consumer expenditure surveys is less than in the health surveys, we thought it was better to use the consumer expenditure survey data to construct our sample because the sampling design ensures that it is representative of the population in so far as consumption is concerned.

3.2. Statistical inference

In order to account for the fact that the compared distributions of individual consumptions and probability of death are samples drawn from a larger population, we perform statistical inference based on the Union-Intersection (UI) method as advocated and initiated by Bishop *et al.* (1989) or Bishop & Formby (1999). Suppose that we want to compare two socially risky situations P and Q according to, say the SEHP criterion (the method works just as well for the SEPG one), and that we have samples p and q of these two situations. Denote by $\widehat{ED}_p^d(t)$ and $\widehat{ED}_p(t)$ the *estimated* expected poverty in the state of death and the expected poverty in either state in p respectively defined by:

$$\begin{aligned}\widehat{ED}_p^d(t) &= \frac{1}{n_p}\sum_{i=1}^{n_p}(1-p_i)\cdot I(x_i^p \leq t),\\ \widehat{ED}_p(t) &= \frac{1}{n_p}\sum_{i=1}^{n_p}p_i\cdot I(y_i^p \leq t)+\frac{1}{n_p}\sum_{i=1}^{n_p}(1-p_i)\cdot I(y_i^p \leq t),\end{aligned}$$

where n_p is the number of individuals in the sample corresponding to p. One can define $\widehat{ED}_q^d(t)$ and $\widehat{ED}_p(t)$ for q in an analogous fashion. We conclude that distribution p dominates q if $\widehat{ED}_p^d(t) \leq \widehat{ED}_q^d(t)$ and $\widehat{ED}_p(t) \leq \widehat{ED}_q(t)$ for all possible existing levels of income t. In order to be sure that this conclusion is valid for the true situations P and Q in a statistically significant way, we need to perform statistical inference on it. For this sake, consider a grid of K poverty thresholds, $(t_1, ..., t_K)$, and define the statistics:

$$
\begin{aligned}
T_k^d &= \frac{\widehat{ED}_p^d(t_k) - \widehat{ED}_q^d(t_k)}{\left(\widehat{\omega}_p^d(t_k)/n_p + \widehat{\omega}_q^d(t_k)/n_q\right)} \\
T_k &= \frac{\widehat{ED}_p(t_k) - \widehat{ED}_q(t_k)}{\left(\widehat{\omega}_p(t_k)/n_p + \widehat{\omega}_q(t_k)/n_q\right)},
\end{aligned}
$$

where $\widehat{\omega}_p^d(t_k)$ and $\widehat{\omega}_p(t_k)$ are the estimations of the asymptotic variance of $\widehat{ED}_p^d(t_k)$ and $\widehat{ED}_p(t_k)$, respectively. Following the reasoning of Davidson & Duclos (2000) and Duclos *et al.* (2006) , based on the law of large numbers and the central limit theorem we can derive estimations of the asymptotic variance of the estimated expected poverty as:

$$
\begin{aligned}
\widehat{\omega}_p^d(t_k) &= \frac{1}{n_p}\sum_{i=1}^{n_p}\left[(1-p_i)\cdot I(x_i^p \leq t)\right]^2 - \left[\widehat{ED}_p^d(t_k)\right]^2 \\
\widehat{\omega}_p(t_k) &= \frac{1}{n_p}\sum_{i=1}^{n_p}\left[(1-p_i)\cdot I(x_i^p \leq t)\right]^2 + \frac{1}{n_p}\sum_{i=1}^{n_p}\left[p_i \cdot I(y_i^p \leq t)\right]^2 - \left[\widehat{ED}_p(t_k)\right]^2.
\end{aligned}
$$

The UI rule says that we infer a SEHP dominance of P over Q based on the sample estimates of expected poverty if none of the poverty inequalities that define the criterion is statistically positive and at least one is statistically negative. In order to define this rule formally, we denote by C_α the critical value for a level of significance of α as determined from the Studentized Maximum Modulus distribution provided by Stoline & Ury (1979). The degree of freedom of this statistics is $2 \cdot K$, which corresponds to the number of equalities that we want to test simultaneously (K inequalities for expected poverty in the bad state, and K inequalities for expected poverty in either state). Given these definitions, the UI rule says that:

- If $\max\left(T_1^b, ..., T_K^b, T_1, ..., T_K\right) < C_\alpha$ and $\min\left(T_1^b, ..., T_K^b, T_1, ..., T_K\right) < -C_\alpha$, we infer that p dominates q.
- If $\min\left(T_1^b, ..., T_K^b, T_1, ..., T_K\right) > -C_\alpha$ and $\max\left(T_1^b, ..., T_K^b, T_1, ..., T_K\right) > C_\alpha$, we infer that q dominates p.
- If $\min\left(T_1^b, ..., T_K^b, T_1, ..., T_K\right) > -C_\alpha$ and $\max\left(T_1^b, ..., T_K^b, T_1, ..., T_K\right) < C_\alpha$, we infer that p and q are not different.

- we infer that p and q are not-comparable otherwise.

The comparisons that are reported herein are performed at the 95 % confidence level.

3.3. Results

Figures 1 and 2 provide the expected headcount poverty curves (expressed in fraction of the whole population) in situation of death and in all states, respectively, for the two years (1995 and 2002). The curves are depicted for the bottom part (below 7500 rupees per person) of the distribution to make them more visible. As is clear from figure 1, no matter what is the poverty line, there is a much lower fraction of the population that is likely to be poor and dead in 2002 than in 1995. This largely reflects the fact that the probability of dying has gone down in India between the two period, as is made clear when comparing the curves for very high poverty lines at which everyone is poor.

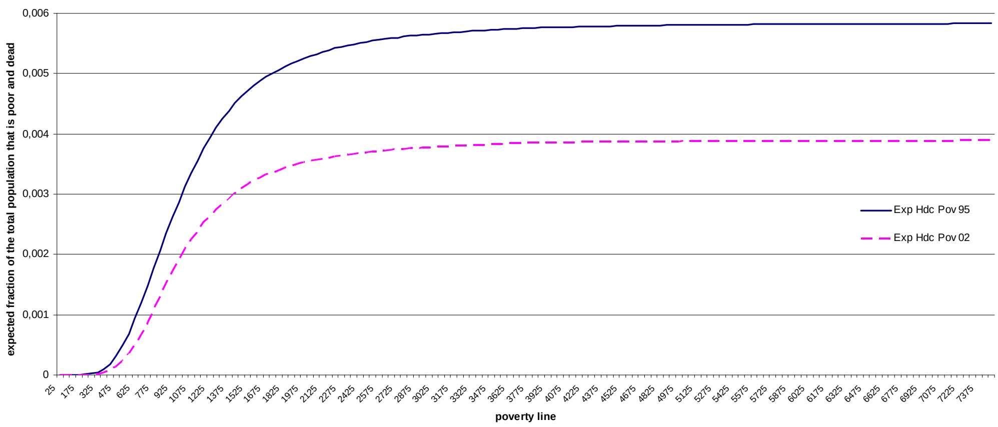

Figure 1. Expected headcount poverty in death.

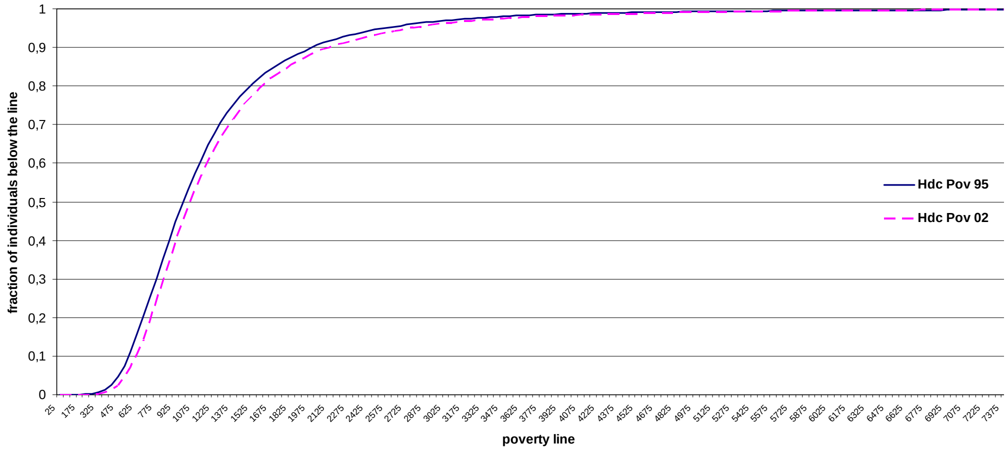

Figure 2. Expected headcount dominance irrespective of the state

Dominance is not so clear when one looks at figure 2 and examines total expected headcount poverty curves irrespective of the state. As it happens, there is crossing of the two curves in their upper part so that the verdict obtained from performing a proper inference is that 2002 does not dominate 1995 for the SEHP criterion. Hence, there are individual VNM preferences in the class U_1 which, when aggregated by a Paretian and VNM social planner, consider that the distribution of risk of death was better in 1995 than it was in 2002. It is clear from looking at figure 2 that this non-dominance comes from the upper part of the distribution of individual's consumption. That is, high consumption individuals happen to be slightly worse off consumption-wise in 2002 than 1995. As this effect is statistically significant, it prevents 2002 from dominating 1995-1996 for the SEHP criterion (remember that condition (2.6) must hold for all poverty lines, including very high ones, in order for SEHP dominance to be observed). We emphasize however that the "crossing" of headcount poverty curve takes place only at the highest income levels and that removing from the sample 5% of the richest individuals would be more than enough to obtain dominance. Hence we can say that, except for a few very rich individuals, the distribution of exposure to risks of death among Indian is better in 2002 than 1995 for the SEHP dominance criterion.

We must say that we are slightly uncomfortable with this conclusion -

that high income groups have financially suffered from growth in India - that goes against the widely acknowledged feeling of the "unequal" feature of Indian's growth that is supposed - to the contrary - to have benefited more to high income people than to low-income ones.[3] We suspect that the well-known inadequacy of NSS data to reflect the consumption patterns of high income groups (as documented among others by Banerjee & Piketty (2005)) is at work here and that the non-dominance of the upper tail of the distribution of consumption is an artefact of the NSS consumption data.

Yet this artifact is not sufficiently important to prevent 2002 from dominating 1995 when one uses the more discriminatory, but less robust, SEHP criterion. Figures 3 and 4 show the expected poverty gap curves in situation of death and the expected poverty gap curve irrespective of the states. While there seem to be a clear dominance of 2002 over 1995 insofar as the expected poverty gap in death is concerned, the impression is probably not as striking when looking at the poverty gap curves irrespective of the state. Yet the small difference that appears on Figure 4 happens to be statistically significant.

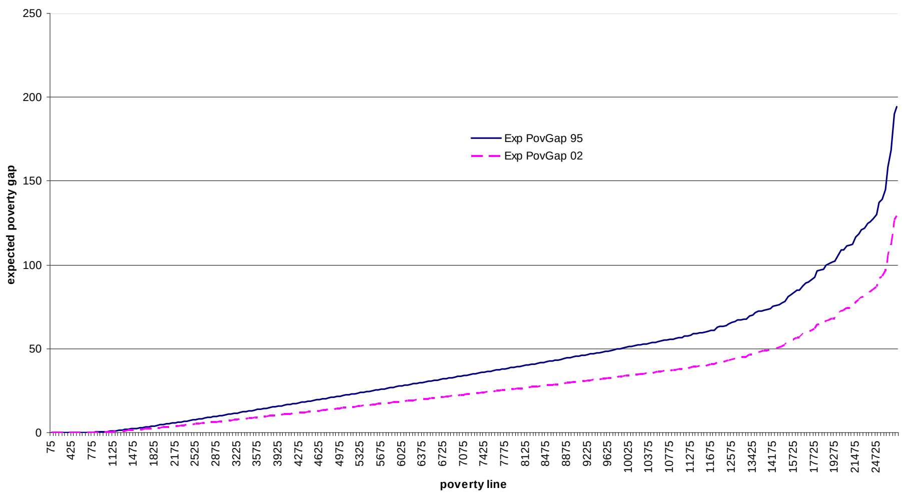

Figure 3. Expected poverty gap in death

[3]This conclusion was also obtained in our companion piece Gravel & Mukhopadhyay (2007).

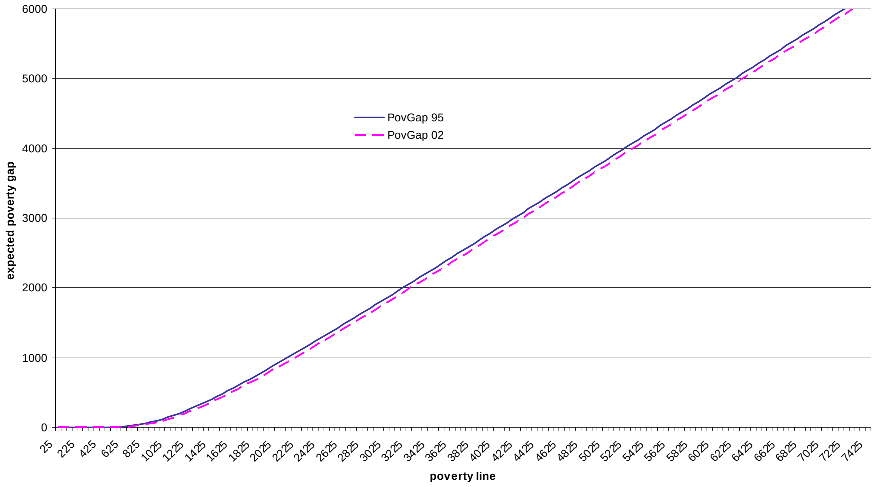

Figure 4. Expected poverty gap irrespective of the state

The table below summarizes the results of the inference procedure. The main conclusion obtained therefore is that there is unanimity of all VNM and Paretian social planners to consider that the distribution of risks of death in India is better in 2002 than in 1995 if these social planners assume that individuals' expected utility functions belong to the class $\mathbb{U}_2$. Moreover, despite the "ambiguity" of the verdict obtained for the even wider unanimity of judgements that correspond to the SEHP criterion, we reemphasize that it would be easy to obtain dominance of 2002 over 1995 for the SEHP criterion as well by simply dropping the 5% richest individuals from our sample.

	2002 vs 1995	min $(T_1^b, ..., T_K)$	max $(T_1^b, ..., T_K)$
Exp. HP in death	2002	-44.0737	-1.6387
HP	ambiguous	-45.3984	8.3139
SEHP	ambiguous	-45.3984	8.3139
EPG in death	2002	-45.5006	-1.7343
PG	2002	-48.9424	-2.2509
SEPG	2002	-44.9424	-1.7343

4. Conclusion

As it comes out from this study, it happens that the distribution of risks of death in India has improved significantly and robustly on the 1995-2002 period. Any ethics which evaluates, at the social level, risky prospects with VNM preferences and which respect individuals VNM preferences will agree with this statement if it assumes that individuals VNM preferences are in the class $\mathbb{U}_2$. This is certainly a strong result which should lead one toward some optimism with respect to the impact of the growth that India has experimented in the last 15 years on an important aspect of human life: exposure to life threatening risks. Of course, we did not get dominance for the even more robust SEHP criterion but, as argued in the text, we believe that this failure to achieve dominance, caused by the fact that high expenditure consumers have observed a slight drop in their consumption level between 1995-2002, is largely due to the inadequacy of NSS data to account for high consumption levels. Strong sequential expected headcount dominance would obtain if one would focus on the 95% poorest part of the population.

Two broad lessons can be derived from this paper. First, our analysis illustrates the workability of the dominance methodology to perform normative appraisal, even when the appraisal concerns the distributions of several attributes or, as examined in this paper, exposures to risk of death. It is sometimes held that dominance methodology is not workable because it is too complex to use and does not always lead to clear-cut conclusions (the ranking of social states that it provides is incomplete). What this paper and others show is that such a doubt is unwarranted. Dominance criteria are easy to implement and do more than often give rise to a clear conclusion. When they do, the robustness of the conclusion is worth keeping in mind, especially when the dominance verdict is obtained in the comparisons of distributions of several attributes.

A second lesson that can be drawn from this exercise is empirical. We have shown that there is a strong sense by which one can say that the distribution of risk of death has improved in India over the period 1995-2002. While we interpret this result as pushing toward some optimism with respect to the impact of growth on Indians' well-being, we emphasis that this optimism is of a moderate variety. It is a minimal requirement that growth should benefit all. What this study, as well as that conducted in Gravel & Mukhopadhyay (2007), shows is that this minimal requirement seems to be met, at least when one looks at exposure to life-threatening risks or at the other district-levels attributes considered in Gravel & Mukhopadhyay (2007). But this conclusion is certainly not antagonistic to the one that would like Indian growth to satisfy more ambitious normative requirements with respect to poverty alleviation.

References

Atkinson, A. B., & Bourguignon, F. 1982. The comparison of multi-dimensioned distribution of economic status. *Review of economic studies*, **49**, 183–201.

Atkinson, A. B., & Bourguignon, F. 1987. Income distributions and differences in needs. *In:* Feiwel, G. R. (ed), *Arrow and the foundation of the theory of economic policy*. London: Macmillan.

Banerjee, A., & Piketty, T. 2005. Are the rich growing richer? evidence from indian tax data. *Pages 520–529 of:* Deaton, A., & Kozel, V. (eds), *The great indian poverty debate*. Macmillan India, Delhi.

Bhalla, S. S. 2005. In(equality) we believe. *Business standard*, **August 21**.

Bishop, C. M., & Formby, J. P. 1999. Test of significance for lorenz partial orders. *In:* Silber, J. (ed), *Handbook of inequality measurement*. Kluwer Academic Press.

Bishop, C. M., Formby, J. P., & Thistle, P. D. 1989. Statistical inference, income distributions and social welfare. *In:* Slotje, D. J. (ed), *Research on economic inequality*. JAI Press.

Crawford, I. 2005. *A nonparametric test of stochastic dominance for multivariate distributions*. University of Surrey, Institute for Fiscal Studies.

Datt, G., & Ravallion, M. 2002. Is india economic growth leaving the poors behind ? *The journal of economic perspectives*, **16**, p.89–108.

Davidson, R., & Duclos, J. Y. 2000. Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, **58**, p.1435–1465.

DEATON, A. 2005. Prices and poverty in india. *Pages 412–427 of:* DEATON, A., & KOZEL, V. (eds), *The great indian poverty debate*. Delhi: Macmillan India.

DEATON, A., & DRÈZE, J. 2005. Poverty and inequality in india: A reexamination. *Pages 428–465 of:* DEATON, A., & KOZEL, V. (eds), *The great indian poverty debate*. Delhi: Macmillan India.

DEATON, A., & KOZEL, V. 2005. *The great indian poverty debate*. Delhi: Macmillan India.

DUCLOS, P. Y., SAHN, D., & YOUNGER, S. D. 2006. Robust multidimensional poverty comparisons. *The economic journal*, **116**, 943–968.

GRAVEL, N., & MUKHOPADHYAY, A. 2007. *Is india better off now than fifteen years ago ? a robust multidimensional answer*. CSH occasionnal paper no 20.

GRAVEL, N., & TARROUX, B. 2007. *Robust normative comparisons of socially risky situations*. IDEP working paper no 08-02.

GRAVEL, N., MOYES, P., & TARROUX, B. 2007. Robust international comparisons of distributions of disposable income and access to regional public goods. *Economica*. forthcoming.

HARSANYI, J. 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of political economy*, **63**, 309–321.

HIMANSHU, & SEN, A. 2005. Poverty and inequality in india. *Pages 500–515 of:* DEATON, A., & KOZEL, V. (eds), *The great indian poverty debate*. Macmillan India, Delhi.

KOLM, S. C. 1977. Multidimensional egalitarianisms. *Quarterly journal of economics*, **91**, 1–13.

RAWLS, JOHN. 1971. *A theory of justice*. Cambridge, Massachusett: Berknap Press of Harvard University Press.

SEN, A. K. 1987. *The standard of living*. Cambridge, UK: Cambridge University Press.

SEN, A. K. 1992. *Inequality reexamined*. Cambridge, MA: Harvard University Press.

STOLINE, M. R., & URY, H. K. 1979. Tables of the studentized maximum modulus distributions and an application to multiple comparisons among means. *Technometrics*, **21**, 87–93.

WEYMARK, J. A. 1993. Harsanyi's social aggregation theorem and the weak pareto principle. *Social choice and welfare*, **10**, 209–221.