Stochastic Dominance

by

Russell Davidson

GREQAM
Centre de la Vieille Charité
2 rue de la Charité
13236 Marseille cedex 02, France

Department of Economics
McGill University
Montreal, Quebec, Canada
H3A 2T7

email: **Russell.Davidson@mcgill.ca**

JEL codes: I300, I320

March 2006

1. Introduction

Stochastic dominance is a term which refers to a set of relations that may hold between a pair of distributions. A very common application of stochastic dominance is to the analysis of income distributions and income inequality, the main focus in this article. The concept can, however, be applied in many other domains, in particular financial economics, where the distributions considered are usually those of the random returns to various financial assets. In what follows, there are often clear analogies between things expressed in terms of income distributions and financial counterparts.

2. Definition of stochastic dominance

In order to determine whether a relation of stochastic dominance holds between two distributions, the distributions are first characterized by their **cumulative distribution functions**, or CDFs. For a given set of incomes, the value of the CDF at income y is the proportion of incomes in the set that are no greater than y. In the context of a random variable Y, the value of the CDF of the distribution of Y at y is the probability that Y should be no greater than y.

Suppose that we consider two distributions A and B, characterized respectively by CDFs F_A and F_B. Then distribution B dominates distribution A stochastically at first order if, for any argument y, $F_A(y) \geq F_B(y)$. This definition often looks as though it is the wrong way round, but a moment's reflection shows that it is correct as stated. If y denotes an income level, then the inequality in the definition means that the proportion of individuals in distribution A with incomes no greater than y is no smaller than the proportion of such individuals in B. In other words, there is at least as high a proportion of poor people in A as in B, if poverty means an income smaller than y. If B dominates A at first order, then whatever poverty line we may choose, there is always more poverty in A than in B, which is why we say that A is the dominated distribution.

Higher orders of stochastic dominance can also be defined. To this end, we define repeated integrals of the CDF of each distribution. Formally, we define a sequence of functions by the recursive definition

$$D^1(y) = F(y), \quad D^{s+1}(y) = \int_0^y D^s(z)\,dz, \quad \text{for } s = 1, 2, 3. \ldots.$$

Thus the function D^1 is the CDF of the distribution under study, $D^2(y)$ is the integral of D^1 from 0 to y, $D^3(y)$ is the integral of D^2 from 0 to y, and so on. By definition, distribution B dominates A at order s if $D^s_A(y) \geq D^s_B(y)$ for all arguments y. The lower limit of 0 is used for clarity of exposition; in general it is the lowest income in the pooled distributions. The definition makes it clear that first-order dominance implies dominance at all higher orders, and more generally that dominance at order s implies dominance at all orders higher than s. Since the implications go in only one direction, it follows that higher-order dominance is a weaker condition than lower-order dominance. I will give a more detailed interpretation of the functions D^s shortly in the context of poverty indices.

In theoretical arguments, it is sometimes desirable to distinguish weak from strong stochastic dominance. The above definitions are of weak dominance. For strong dominance, it is required that the inequality should be strict for at least one value of the argument y. In empirical investigations, the distinction is of no interest, since no statistical test can detect a significant difference between weak and strong inequalities.

Some applications make use of the concept of **restricted stochastic dominance**. This means that the relevant inequality is required to hold over some restricted range of the argument y rather than for all possible values. In empirical work, it is often only restricted dominance that can usefully be studied, since with continuous distributions there are usually too few data in the tails of the distributions for statistically significant conclusions to be drawn. Again, for measures of poverty, it is only dominance over the range of incomes up to the poverty line that is of interest.

3. Relation between stochastic dominance and welfare

When studying either income inequality or poverty, one is automatically in a normative context. Most modern studies make explicit or implicit use of a **social welfare function** or SWF. In a paper by Blackorby and Donaldson (1980) (henceforth BD), various ethically desirable criteria are developed and the sorts of SWF that respect these criteria are characterized.

One of these criteria is the **anonymity** of individuals. If we take all the worldly goods of a rich man and give them to a poor man, and then give the few worldly goods of the poor man to the rich, then social welfare should be unchanged. Formally, a SWF that respects this requirement is symmetric with respect to its arguments, which are the incomes of the members of society.

Another requirement is the **Pareto principle**. According to it, we should rank situation B better than situation A if at least one individual is better off in B than in A, and no one is worse off. In order for a SWF to respect the Pareto principle, it must be increasing in all its arguments.

Another requirement, for measures of poverty only, is that a poverty index should not depend at all on the incomes of the nonpoor. BD show that this implies a separability condition on the SWF. If in addition we require that the poverty index should be defined for arbitrary poverty lines, then the separability condition becomes a requirement of additive separability. The SWF can therefore be written as

$$W(y_1, \ldots, y_N) = \sum_{i=1}^{N} u(y_i), \tag{1}$$

where the "utility" function u is increasing in its argument. Alternatively, the SWF can be any increasing transform of the function W. In all cases, the SWF is symmetric and increasing in its arguments, and so satisfies BD's ethical criteria.

It can be seen that first-order stochastic dominance of A by B means that B has higher social welfare than A for all SWFs of the form (1). This can be shown by a simple integration by parts, under the assumption that the function u is differentiable. In fact, this dominance is also a necessary condition for B to have higher welfare than A for all SWFs of the form (1).

It follows from the above argument that, if we use first-order stochastic dominance as a criterion for ranking distributions, then we need not restrict attention to a specific SWF, since any SWF of the form (1) gives the same ranking if one distribution dominates the other at first order.

A more restricted class of SWFs is given by functions of the form (1) where we impose the additional restriction that the second derivative of u is negative. It turns out that all the SWFs of this more restricted class give a unanimous ranking of two distributions if one dominates the other at second order. This sort of result can be extended to higher orders of dominance. As the dominance condition becomes progressively weaker, the class of SWFs that give unanimous rankings becomes progressively smaller, subject to more and more restrictive conditions on the function u and its derivatives.

4. Relation between stochastic dominance and poverty

The so-called **headcount ratio** is sometimes used as a measure of the amount of poverty in a given income distribution. This ratio is the proportion of individuals in the distribution with incomes below (or equal to) the poverty line. If this line is denoted by z, then the headcount ratio is the value of the CDF at z. If we have two populations, A and B, characterized by two CDFs, F_A and F_B, then, for poverty line z, the headcount ratio is higher in A than in B if and only if $F_A(z) > F_B(z)$. If the inequality $F_A(y) > F_B(y)$ holds for all values of y up to z, then we have restricted first-order stochastic dominance up to z.

Corresponding to any income y less than z, we define the **poverty gap** as $z - y$. When we restrict attention to the welfare of people with incomes less than z, it is convenient to use a function π that measures the disutility of the poverty gap, rather than a function u of the sort used in a SWF. Thus we have a class of **poverty indices**, defined as follows:

$$\Pi(z) = \int_0^z \pi(z-y)\,dF(y),$$

If $\pi' > 0$, which means that the disutility of poverty increases with the poverty gap, it can be shown that, for all poverty indices of the above form, there is more poverty in A than in B if B dominates A at first order over the range of incomes less than or equal to z.

Earlier, a sequence of functions D^s was introduced, these functions being repeated integrals of the CDF. A useful explicit representation of these functions is given by the formula

$$D^s(z) = \frac{1}{(s-1)!}\int_0^z (z-y)^{s-1}\,dF(y). \tag{2}$$

The formula clearly holds for $s = 1$, if we remember that $0! = 1$. It is not hard to show by induction that it also holds for integers greater than 1.

For $s = 2$, the formula becomes

$$D^2(z) = \int_0^z (z - y)\, dF(y),$$

from which we see that, for given z, $D^2(z)$ is the average poverty gap for poverty line z. If, for all $z \in [z_-, z_+]$, $D^2_A(z) > D^2_B(z)$, then it follows that the average poverty gap is greater in A than in B for all poverty lines in the interval $[z_-, z_+]$. But this condition is just restricted second-order stochastic dominance of A by B over that interval.

As with welfare functions, this result can be extended. By progressively restricting the admissible class of poverty indices, in particular by imposing signs on the derivatives of π, it can be seen that all poverty indices in these more restricted classes unanimously see more poverty in A than in B if there is a progressively higher order of stochastic dominance; see Davidson and Duclos (2000) for more details.

An essential reference on poverty measurement is Atkinson (1987), in which the axiomatic approach of BD is extended to poverty measurement. See also three papers by Foster and Shorrocks (1988a, b, and c).

5. Relation between stochastic dominance and inequality

If a richer person in distribution A transfers some income to a poorer person in such a way that the richer person stays richer after the transfer, the post-transfer distribution B stochastically dominates A at second order. The Pigou-Dalton principle of transfers says that "Robin-Hood" transfers of the sort described should improve welfare. But it is easy to see that distribution B does not dominate A at first order, and indeed this is right and proper according to the Pareto principle, since the richer person is worse off after the transfer.

This example shows that, when we discuss inequality, we are not talking about the same thing as welfare. Any reasonable measure of inequality must declare that there is no inequality if everyone has the same income, even if everyone is in abject poverty.

The classical tool for studying inequality is the **Lorenz curve**. For any proportion p between zero and one, the ordinate of the corresponding point on the Lorenz curve for a given income distribution is the proportion of total income that accrues to the first $100p$ per cent of people when they are sorted in order of increasing income. By construction, the Lorenz curve fits into the unit square, lies below the 45-degree line that is the diagonal of that square, and is (weakly) convex. Figure 1 displays a typical Lorenz curve.

A distribution B is said to **Lorenz dominate** another distribution A if the Lorenz curve of B lies everywhere above that of A. We then say that there is less inequality in B than in A. But this comparison of A and B is not a welfare comparison, and, in particular, does not allow a comparison of poverty. This defect is remedied by

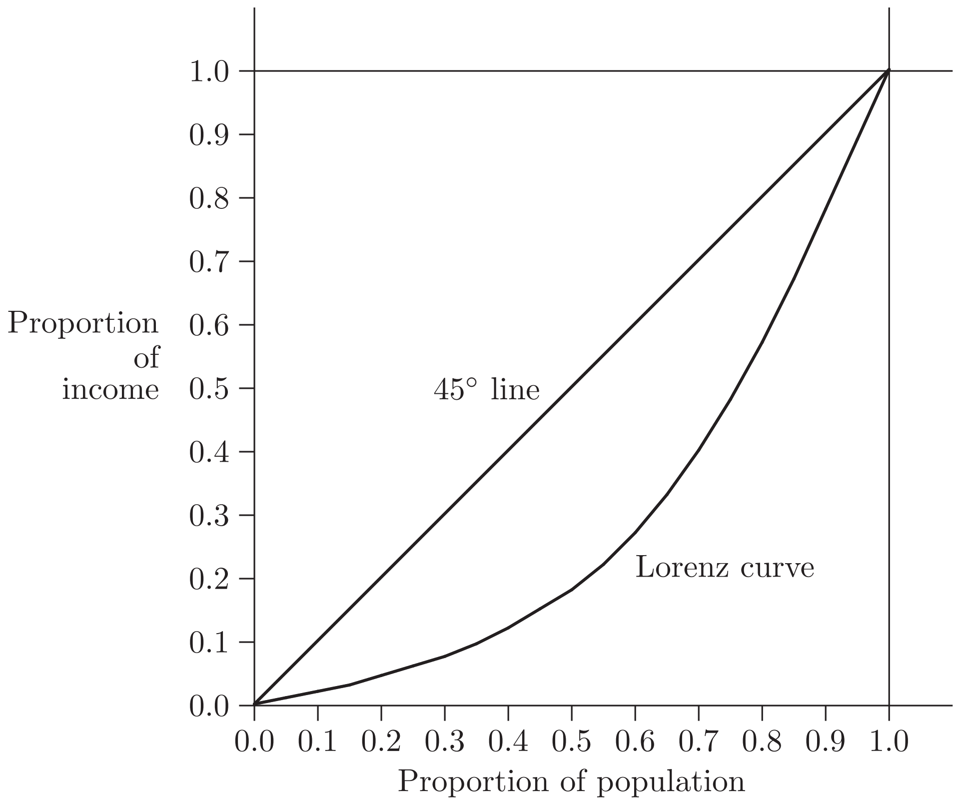

Figure 1. A Typical Lorenz Curve

the concept of **generalized Lorenz dominance**, based on the generalized Lorenz curve introduced by Shorrocks (1983). The ordinates of this curve are the Lorenz ordinates multiplied by the average income of the distribution. It turns out that generalized Lorenz dominance is the same thing as second-order stochastic dominance. Either one of these concepts implicitly mixes notions of welfare and inequality, as shown by the fact that the function u in a SWF of form (1) that respects second-order dominance has a negative second derivative, which implies diminishing marginal (social) utility of income. The discussion of the previous section shows that higher-order dominance criteria put more and more weight on the welfare of the poorest members of society.

Graphical representation and quantiles

Consider the setup in Figure 2, where the CDFs of two distributions A and B are plotted. The functions D^2 used for second-order dominance comparisons can be evaluated for a given argument, like z_1 in the figure, as the areas beneath the CDFs, by the usual geometric interpretation of the Riemann integral. We see that distribution B dominates A at second order because, although the CDFs cross, the areas between them are such that the condition for second-order dominance is always satisfied. Thus the vertical line MN marks off a large positive area between the graphs of the two CDFs up to the point at which they cross, and thereafter a small negative area bounded on the right by MN.

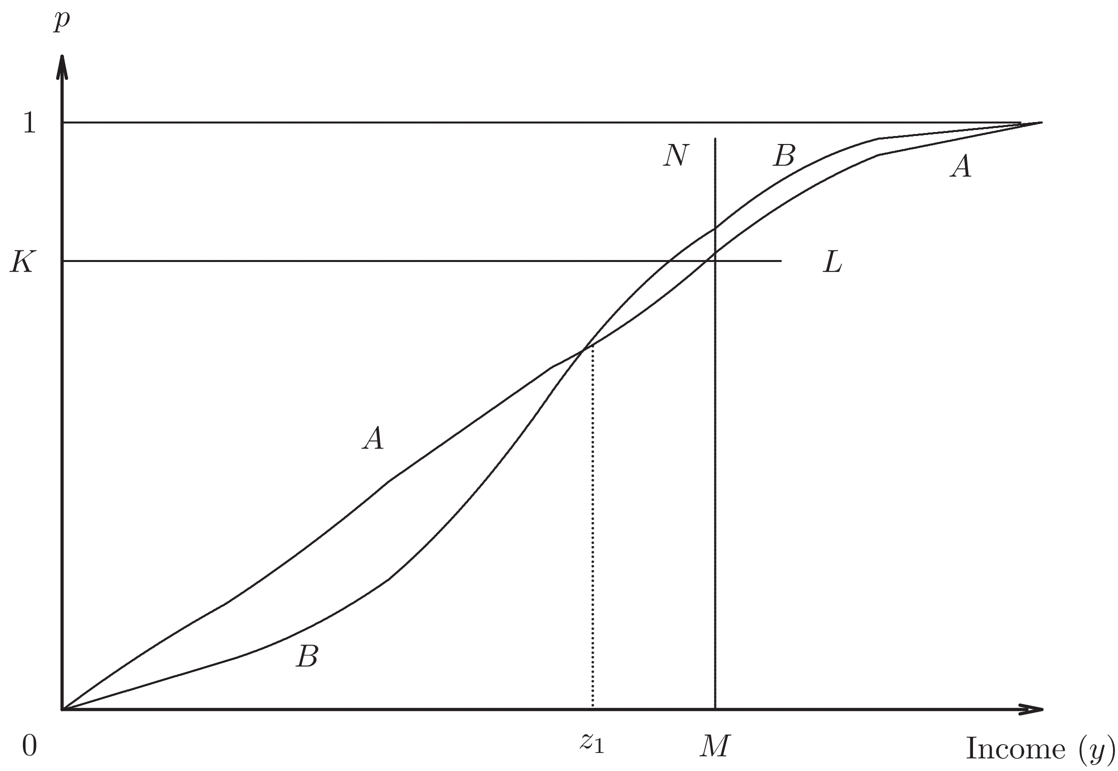

Figure 2. Generalized Lorenz and Second Order Dominance

For generalized Lorenz dominance, it can be shown that what must be non-negative everywhere is the area between the two curves, bounded not by a vertical line like MN, but rather by a horizontal line like KL. This area is the difference between the areas under two **quantile functions**, a quantile function being by definition the inverse of the CDF. Although it is tedious to demonstrate it algebraically, it is intuitively clear that if the areas bounded on the right by vertical lines like MN are always positive, then so are the areas bounded above by horizontal lines like KL. This is why generalized Lorenz dominance and second-order stochastic dominance are equivalent conditions. The whole theory of stochastic dominance can be developed using quantiles rather than incomes; this is called a **p-approach**. Such approaches are used to advantage in Jenkins and Lambert (1997, 1998), Shorrocks (1998), and also Spencer and Fisher (1992).

Another thing that emerges clearly from Figure 2 is that the threshold income z_1 up to which first-order stochastic dominance holds is always smaller than the threshold z_2 up to which we have second-order dominance. In the Figure, we have second-order dominance everywhere, and so we can set z_2 equal to the highest income in either distribution. More generally, we can define a threshold z_s as the greatest income up to which we have dominance at order s. The z_s constitute an increasing sequence.

A result shown in Davidson and Duclos (2000) is that, if the distribution B dominates A at first-order over a range $[0, z]$, with $z > 0$, then, no matter what happens for incomes above z, there is always *some* order s such that B dominates A at order s over the full range of the two distributions, provided only that that range is finite.

Estimation and Inference

Suppose that we have a random sample of N independent observations y_i, $i = 1, \ldots, N$, from a population. Then it follows from (2) that a natural estimator of $D^s(z)$ (for a nonstochastic z) is

$$\hat{D}^s(z) = \frac{1}{(s-1)!} \int_0^z (z-y)^{s-1}\, d\hat{F}(y) = \frac{1}{n(s-1)!} \sum_{i=1}^{N} (z-y_i)^{s-1} \mathrm{I}(y_i \le z), \qquad (3)$$

where $\hat{F}$ denotes the empirical distribution function of the sample and $\mathrm{I}(\cdot)$ is an indicator function equal to 1 when its argument is true and 0 otherwise. For $s = 1$, the formula (3) estimates the population CDF by the empirical distribution function. For arbitrary s, it has the convenient property of being a sum of independent and identically distributed (IID) variables, which makes it easy to show that (3) is consistent and asymptotically normal. The asymptotic variance is also easy to estimate in a distribution-free manner, by which is meant that no parametric assumptions need be made about the distributions under study.

When two distributions are compared for stochastic dominance, two kinds of situations typically arise. The first is when there are two independent populations, with random samples from each. The other arises when we have independent paired drawings from the same population. For instance, one variable could be before-tax income, and the other after-tax income for the same individual. Explicit expressions for the asymptotic variance of the difference between the estimates of $D^s(z)$ for the case of independent samples were given as early as 1989 in an unpublished thesis, Chow (1989). The sampling distribution of a related estimator for poverty indices and independent samples is found in Kakwani (1993), Bishop, Chow, and Zheng (1995) and Rongve (1997). For a different approach to inference on stochastic dominance, see Anderson (1996). A comprehensive approach to inference on stochastic dominance is found in Davidson and Duclos (2000).

The approach proposed by McFadden (1989) is based on the supremum of the difference between the estimates (3) for two independent populations. For $s = 1$, this turns out to be a variant of the Kolmogorov-Smirnov test, with known properties. For higher values of s, although it is easy to compute the statistic, its asymptotic properties under the null are not analytically tractable. However, simulation-based methods can surmount this difficulty; see Barrett and Donald (2003).

A somewhat vexed question in testing for dominance is whether to test the null hypothesis of dominance or that of non-dominance. The latter has the advantage that, if the null is rejected, all that remains is dominance. More generally, the former approach rejects the null of dominance only when there is clear evidence against it, and the latter accepts the alternative of dominance only when there is clear evidence in its favour. The former approach is more common in the literature; see for instance

Richmond (1982), Beach and Richmond (1985), Wolak (1989), and Bishop, Formby, and Thistle (1992). The latter is discussed in an unpublished paper, Howes (1993), and in Kaur, Prakasa-Rao, and Singh (1994).

Bibliography

Anderson, G. (1996). Nonparametric Tests of Stochastic Dominance In Income Distributions, *Econometrica*, 64, 1183-1193.

Atkinson, A.B. (1987). On the Measurement of Poverty, *Econometrica*, 55, 749–764.

Barrett, G. and S.G. Donald (2003). Consistent Tests for Stochastic Dominance *Econometrica*, 71, 71-103.

Beach, C.M. and J. Richmond (1985). Joint Confidence Intervals for Income Shares and Lorenz Curves, *International Economic Review*, 26, 439–50.

Bishop, J.A., K.V. Chow and B. Zheng (1995). Statistical Inference and Decomposable Poverty Measures, *Bulletin of Economic Research*, 47, 329–340.

Bishop, J.A, J.P. Formby, and P. Thistle (1992). Convergence of the South and Non-South Income Distributions, 1969–1979, *American Economic Review*, 82, 262–272.

Blackorby, C. and D. Donaldson (1980). Ethical Indices for the Measurement of Poverty, *Econometrica*, 48, 1053–1062.

Chow, K.V. (1989). *Statistical Inference for Stochastic Dominance: a Distribution Free Approach*, Ph.D. thesis, University of Alabama.

Davidson, R. and J.-Y. Duclos (2000). Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality, *Econometrica*, 68, 1435-1464.

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

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

Foster, J.E. and A.F. Shorrocks (1988c). Inequality and Poverty Orderings, *European Economic Review*, 32, 654–662.

Howes, S. (1993). Asymptotic Properties of Four Fundamental Curves of Distributional Analysis, Unpublished paper, STICERD, London School of Economics.

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.

Jenkins, S.P. and P.J. Lambert (1998). Three ‘I’s of Poverty Curves and Poverty Dominance: TIPs for Poverty Analysis, *Research on Economic Inequality*, 8, 39–56.

Kakwani, N. (1993). Statistical Inference in the Measurement of Poverty, *Review of Economics and Statistics*, 75, 632–639.

Kaur, A., B.L. Prakasa Rao, and H. Singh (1994). Testing for Second-Order Stochastic Dominance of Two Distributions, *Econometric Theory*, 10, 849–866.

Richmond, J. (1982). A General Method for Constructing Simultaneous Confidence Intervals, *Journal of the Americal Statistical Association*, 77, 455–460.

Rongve, I. (1997). Statistical Inference for Poverty Indices with Fixed Poverty Lines, *Applied Economics*, 29, 387–392.

Shorrocks, A.F. (1983). Ranking Income Distributions, *Economica*, 50(197), 3-17.

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.

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

Wolak, F.A. (1989). Testing Inequality Constraints in Linear Econometric Models, *Journal of Econometrics*, 41, 205–235.