

Brief introduction to multidimensional stochastic dominance

Gaston Yalonetzky

Oxford Poverty and Human Development Initiative, University of Oxford

OPHI-HDCA Summer School, Delft, 24 August - 3 September 2011.

We are grateful to the World Bank, two anonymous donors and OPHI for financial support

Table of contents

Introduction

Stochastic dominance conditions: for univariate to multivariate settings

Multidimensional stochastic dominance and complementarity/substitutability

Relevance for poverty measurement

Tests for univariate and multivariate stochastic dominance

The test of Barrett and Donald (2003)

Concluding remarks

Introduction: Multidimensional versus unidimensional dominance

- ▶ Stochastic dominance conditions provide an extreme form of robustness for ordinal comparisons: if they are fulfilled a comparison is robust to a broad range of parameter values and families of indices.

Introduction: Multidimensional versus unidimensional dominance

- ▶ Stochastic dominance conditions provide an extreme form of robustness for ordinal comparisons: if they are fulfilled a comparison is robust to a broad range of parameter values and families of indices.
- ▶ Multidimensional dominance is relevant for evaluation functions that map from a multivariate space.

Introduction: Multidimensional versus unidimensional dominance

- ▶ Stochastic dominance conditions provide an extreme form of robustness for ordinal comparisons: if they are fulfilled a comparison is robust to a broad range of parameter values and families of indices.
- ▶ Multidimensional dominance is relevant for evaluation functions that map from a multivariate space. E.g. An index of well-being that depends on several aspects of wellbeing.

Introduction: Multidimensional versus unidimensional dominance

- ▶ Stochastic dominance conditions provide an extreme form of robustness for ordinal comparisons: if they are fulfilled a comparison is robust to a broad range of parameter values and families of indices.
- ▶ Multidimensional dominance is relevant for evaluation functions that map from a multivariate space. E.g. An index of well-being that depends on several aspects of wellbeing.
- ▶ In unidimensional dominance second and even third orders may be interesting/relevant.

Introduction: Multidimensional versus unidimensional dominance

- ▶ Stochastic dominance conditions provide an extreme form of robustness for ordinal comparisons: if they are fulfilled a comparison is robust to a broad range of parameter values and families of indices.
- ▶ Multidimensional dominance is relevant for evaluation functions that map from a multivariate space. E.g. An index of well-being that depends on several aspects of wellbeing.
- ▶ In unidimensional dominance second and even third orders may be interesting/relevant. In multidimensional dominance second and higher orders are not that easy to interpret.

Introduction: Multidimensional versus unidimensional dominance

- ▶ Stochastic dominance conditions provide an extreme form of robustness for ordinal comparisons: if they are fulfilled a comparison is robust to a broad range of parameter values and families of indices.
- ▶ Multidimensional dominance is relevant for evaluation functions that map from a multivariate space. E.g. An index of well-being that depends on several aspects of wellbeing.
- ▶ In unidimensional dominance second and even third orders may be interesting/relevant. In multidimensional dominance second and higher orders are not that easy to interpret.
- ▶ By contrast, in multidimensional dominance other things matter: the joint distribution of the variables, as well as how they complement, or substitute, each other in their contributions toward the evaluation function (e.g. "increasing wellbeing").

Introduction: Multidimensional dominance conditions

- ▶ In this lecture we are going to review the derivation of dominance conditions and extend it to multivariate settings.

Introduction: Multidimensional dominance conditions

- ▶ In this lecture we are going to review the derivation of dominance conditions and extend it to multivariate settings.
- ▶ We will focus on first-order dominance and bivariate distributions and derive the respective conditions.

Introduction: Multidimensional dominance conditions

- ▶ In this lecture we are going to review the derivation of dominance conditions and extend it to multivariate settings.
- ▶ We will focus on first-order dominance and bivariate distributions and derive the respective conditions.
- ▶ We will discuss the logic behind some of the conditions and their connections to complementarity and substitutability between variables.

Introduction: Multidimensional dominance conditions

- ▶ In this lecture we are going to review the derivation of dominance conditions and extend it to multivariate settings.
- ▶ We will focus on first-order dominance and bivariate distributions and derive the respective conditions.
- ▶ We will discuss the logic behind some of the conditions and their connections to complementarity and substitutability between variables.
- ▶ We will discuss why these conditions are also relevant for poverty assessments.

Introduction: Multidimensional dominance conditions

- ▶ In this lecture we are going to review the derivation of dominance conditions and extend it to multivariate settings.
- ▶ We will focus on first-order dominance and bivariate distributions and derive the respective conditions.
- ▶ We will discuss the logic behind some of the conditions and their connections to complementarity and substitutability between variables.
- ▶ We will discuss why these conditions are also relevant for poverty assessments.
- ▶ We will briefly discuss how to test these conditions.

Traditional dominance conditions with one variable: an example

Consider the following wellbeing function: $W(x) = \int_{x_{min}}^{x_{max}} U(x)dF(x)$

Traditional dominance conditions with one variable: an example

Consider the following wellbeing function: $W(x) = \int_{x_{min}}^{x_{max}} U(x)dF(x)$
Traditional dominance conditions stem from expressing differences in social welfare functions as sums of products of derivatives of the individual welfare functions and functions of the cumulative (or survival) densities. Example:

Traditional dominance conditions with one variable: an example

Consider the following wellbeing function: $W(x) = \int_{x_{min}}^{x_{max}} U(x)dF(x)$
 Traditional dominance conditions stem from expressing differences in social welfare functions as sums of products of derivatives of the individual welfare functions and functions of the cumulative (or survival) densities. Example:

$$\Delta W = - \int_{x_{min}}^{x_{max}} \frac{\delta U}{\delta x}(x) \Delta F(x) dx$$

The condition is then: $\Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta x}(x) \geq 0 \leftrightarrow \Delta F(x) \leq 0 \forall x$

Traditional dominance conditions with one variable: an example

Consider the following wellbeing function: $W(x) = \int_{x_{min}}^{x_{max}} U(x)dF(x)$
 Traditional dominance conditions stem from expressing differences in social welfare functions as sums of products of derivatives of the individual welfare functions and functions of the cumulative (or survival) densities. Example:

$$\Delta W = - \int_{x_{min}}^{x_{max}} \frac{\delta U}{\delta x}(x) \Delta F(x) dx$$

The condition is then: $\Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta x}(x) \geq 0 \leftrightarrow \Delta F(x) \leq 0 \forall x$

The dominance condition is usually expressed in terms of distributions. E.g. if $\forall x : F^A(x) \leq F^B(x)$ then we say that "distribution A (first-order) dominates B".

First-order conditions for the bivariate case

With two variables we integrate by parts the following function:

$$W = \int_{y_{min}}^{y_{max}} \int_{x_{min}}^{x_{max}} U(x) f(x, y) dx dy$$

First-order conditions for the bivariate case

With two variables we integrate by parts the following function:

$$W = \int_{y_{min}}^{y_{max}} \int_{x_{min}}^{x_{max}} U(x) f(x, y) dx dy$$

Because we have a joint density ($f(x, y)$) we can integrate using either cumulative functions:

First-order conditions for the bivariate case

With two variables we integrate by parts the following function:

$$W = \int_{y_{\min}}^{y_{\max}} \int_{x_{\min}}^{x_{\max}} U(x) f(x, y) dx dy$$

Because we have a joint density ($f(x, y)$) we can integrate using either cumulative functions:

$$\begin{aligned} \Delta W = & - \int_{x_{\min}}^{x_{\max}} \frac{\delta U}{\delta x}(x, y_{\max}) \Delta F^x(x) dx \\ & - \int_{y_{\min}}^{y_{\max}} \frac{\delta U}{\delta y}(x_{\max}, y) \Delta F^y(y) dy + \int_{y_{\min}}^{y_{\max}} \int_{x_{\min}}^{x_{\max}} \frac{\delta^2 U}{\delta x \delta y}(x, y) \Delta F(x, y) dx dy \end{aligned}$$

First-order conditions for the bivariate case

...or survival functions:

First-order conditions for the bivariate case

...or survival functions:

$$\Delta W = \int_{x_{min}}^{x_{max}} \frac{\delta U}{\delta x}(x, y_{min}) \Delta \bar{F}^x(x) dx$$

$$\int_{y_{min}}^{y_{max}} \frac{\delta U}{\delta y}(x_{min}, y) \Delta \bar{F}^y(y) dy + \int_{y_{min}}^{y_{max}} \int_{x_{min}}^{x_{max}} \frac{\delta^2 U}{\delta x \delta y}(x, y) \Delta \bar{F}(x, y) dx dy$$

First-order conditions for the bivariate case

Notice the appearance of $\frac{\delta^2 U}{\delta x \delta y}(x, y)$.

First-order conditions for the bivariate case

Notice the appearance of $\frac{\delta^2 U}{\delta x \delta y}(x, y)$. This cross-partial derivative determines whether x and y are *ALEP complements* or *ALEP substitutes* in their contributions to U .

First-order conditions for the bivariate case

Notice the appearance of $\frac{\delta^2 U}{\delta x \delta y}(x, y)$. This cross-partial derivative determines whether x and y are *ALEP complements* or *ALEP substitutes* in their contributions to U . Four conditions stem from the two previous equations:

First-order conditions for the bivariate case

Notice the appearance of $\frac{\delta^2 U}{\delta x \delta y}(x, y)$. This cross-partial derivative determines whether x and y are *ALEP complements* or *ALEP substitutes* in their contributions to U . Four conditions stem from the two previous equations:

1. A condition for monotonically increasing functions with ALEP substitute arguments (e.g. $\frac{\delta^2 U}{\delta x \delta y}(x, y) \leq 0$):

$$\forall x, y : \Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta i}(x, y) \geq 0 \wedge \frac{\delta^2 U}{\delta x \delta y}(x, y) \leq 0 \leftrightarrow$$

$$\Delta F^x(x), \Delta F^y(y), \Delta F(x, y) \leq 0$$

First-order conditions for the bivariate case

Notice the appearance of $\frac{\delta^2 U}{\delta x \delta y}(x, y)$. This cross-partial derivative determines whether x and y are *ALEP complements* or *ALEP substitutes* in their contributions to U . Four conditions stem from the two previous equations:

1. A condition for monotonically increasing functions with ALEP substitute arguments (e.g. $\frac{\delta^2 U}{\delta x \delta y}(x, y) \leq 0$):

$$\forall x, y : \Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta i}(x, y) \geq 0 \wedge \frac{\delta^2 U}{\delta x \delta y}(x, y) \leq 0 \leftrightarrow$$

$$\Delta F^x(x), \Delta F^y(y), \Delta F(x, y) \leq 0$$

Notice that $\forall x, y : \Delta F(x, y) \leq 0$ suffices to ascertain $\Delta W \geq 0$.

First-order conditions for the bivariate case

2. A condition for monotonically increasing functions with ALEP complement arguments (e.g. $\frac{\delta^2 U}{\delta x \delta y}(x, y) \geq 0$):

$$\forall x, y : \Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta i}(x, y) \geq 0 \wedge \frac{\delta^2 U}{\delta x \delta y}(x, y) \geq 0 \leftrightarrow$$

$$\Delta \bar{F}^x(x), \Delta \bar{F}^y(y), \Delta \bar{F}(x, y) \geq 0$$

First-order conditions for the bivariate case

2. A condition for monotonically increasing functions with ALEP complement arguments (e.g. $\frac{\delta^2 U}{\delta x \delta y}(x, y) \geq 0$):

$$\forall x, y : \Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta i}(x, y) \geq 0 \wedge \frac{\delta^2 U}{\delta x \delta y}(x, y) \geq 0 \leftrightarrow$$

$$\Delta \bar{F}^x(x), \Delta \bar{F}^y(y), \Delta \bar{F}(x, y) \geq 0$$

Notice that $\forall x, y : \Delta \bar{F}(x, y) \geq 0$ suffices to ascertain $\Delta W \geq 0$.

First-order conditions for the bivariate case

3. A condition for monotonically increasing functions with ALEP neutral arguments (e.g. $\frac{\delta^2 U}{\delta x \delta y}(x, y) = 0$):

$$\forall x, y : \Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta i}(x, y) \geq 0 \wedge \frac{\delta^2 U}{\delta x \delta y}(x, y) = 0 \Leftrightarrow$$

$$\Delta F^x(x), \Delta F^y(y) \leq 0$$

First-order conditions for the bivariate case

3. A condition for monotonically increasing functions with ALEP neutral arguments (e.g. $\frac{\delta^2 U}{\delta x \delta y}(x, y) = 0$):

$$\forall x, y : \Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta i}(x, y) \geq 0 \wedge \frac{\delta^2 U}{\delta x \delta y}(x, y) = 0 \Leftrightarrow$$

$$\Delta F^x(x), \Delta F^y(y) \leq 0$$

Notice that, in this case, it is necessary to test $\Delta F^i(i) \leq 0$ for every variable (unlike in the previous cases).

First-order conditions for the bivariate case

4. A condition for ALL individual welfare functions that are monotonically increasing!

$$\forall x, y : \Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta i}(x, y) \geq 0 \leftrightarrow$$

$$\Delta F^x(x), \Delta F^y(y), \Delta F(x, y) \leq 0 \wedge \Delta \bar{F}^x(x), \Delta \bar{F}^y(y), \Delta \bar{F}(x, y) \geq 0$$

First-order conditions for the bivariate case

4. A condition for ALL individual welfare functions that are monotonically increasing!

$$\forall x, y : \Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta i}(x, y) \geq 0 \leftrightarrow$$

$$\Delta F^x(x), \Delta F^y(y), \Delta F(x, y) \leq 0 \wedge \Delta \bar{F}^x(x), \Delta \bar{F}^y(y), \Delta \bar{F}(x, y) \geq 0$$

Notice that, in this case, it is necessary and sufficient to test $\Delta F(x, y) \leq 0$ and $\Delta \bar{F}(x, y) \geq 0$.

First-order conditions for the general *multivariate* case

In the general multivariate case, first-order conditions are based on all cumulative and/or survival functions combining all variables.

First-order conditions for the general *multivariate* case

In the general multivariate case, first-order conditions are based on all cumulative and/or survival functions combining all variables. They include all marginal distributions, $F^i(j)$ and all cumulative and survival joints up to $F(x, y, \dots, w, z)$ and $\bar{F}(x, y, \dots, w, z)$.

First-order conditions for the general *multivariate* case

In the general multivariate case, first-order conditions are based on all cumulative and/or survival functions combining all variables. They include all marginal distributions, $F^i(i)$ and all cumulative and survival joints up to $F(x, y, \dots, w, z)$ and $\bar{F}(x, y, \dots, w, z)$.

In the case of cumulative distributions, the signs of the cross partial derivatives of the functions for which the conditions hold change the following way:
 $U_i > 0$, $U_{ij} < 0$ (ALEP substitution), $U_{ijk} > 0$ and so forth.

First-order conditions for the general *multivariate* case

In the general multivariate case, first-order conditions are based on all cumulative and/or survival functions combining all variables. They include all marginal distributions, $F^i(i)$ and all cumulative and survival joints up to $F(x, y, \dots, w, z)$ and $\bar{F}(x, y, \dots, w, z)$.

In the case of cumulative distributions, the signs of the cross partial derivatives of the functions for which the conditions hold change the following way:
 $U_i > 0$, $U_{ij} < 0$ (ALEP substitution), $U_{ijk} > 0$ and so forth.

In the case of survival functions, all the signs of the cross partial derivatives of the functions for which the conditions hold are positive (e.g. $U_i > 0$, $U_{ij} > 0, U_{ijk} > 0$).

First-order conditions for the general *multivariate* case

In the general multivariate case, first-order conditions are based on all cumulative and/or survival functions combining all variables. They include all marginal distributions, $F^i(i)$ and all cumulative and survival joints up to $F(x, y, \dots, w, z)$ and $\bar{F}(x, y, \dots, w, z)$.

In the case of cumulative distributions, the signs of the cross partial derivatives of the functions for which the conditions hold change the following way:
 $U_i > 0$, $U_{ij} < 0$ (ALEP substitution), $U_{ijk} > 0$ and so forth.

In the case of survival functions, all the signs of the cross partial derivatives of the functions for which the conditions hold are positive (e.g. $U_i > 0$, $U_{ij} > 0$, $U_{ijk} > 0$).

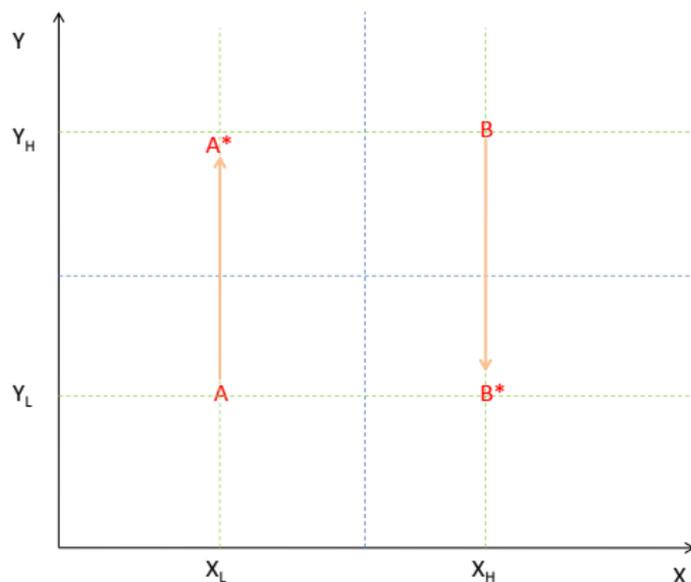
Hence when conditions on both the cumulative and the survival functions are fulfilled:

$$\forall x, y, \dots, z : \Delta W \geq 0 \forall U \mid \frac{\delta U}{\delta i}, \frac{\delta^3 U}{\delta i \delta j \delta k}, \dots \geq 0$$

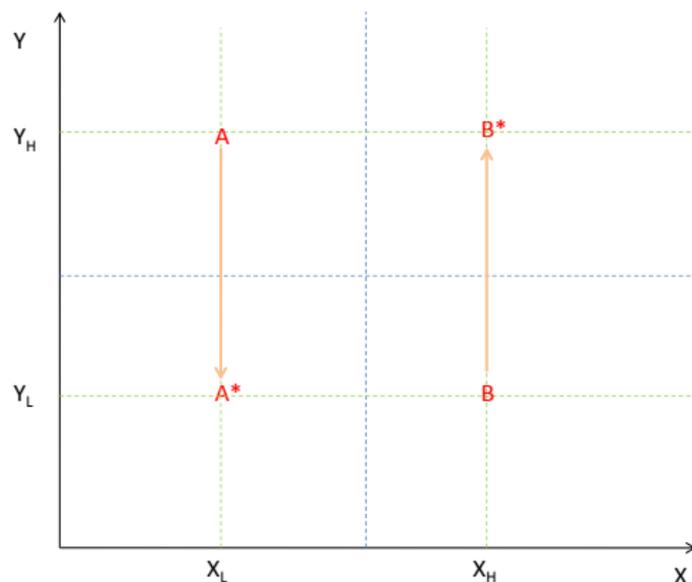
ALEP relationships and multidimensional dominance conditions

Why the conditions on cumulative distributions are associated with ALEP substitutability and the conditions on survival functions are associated with ALEP complementarity?

ALEP substitution and $\Delta F(x, y) \leq 0$



ALEP complementarity and $\Delta \bar{F}(x, y) \geq 0$



Applying dominance conditions to poverty measurement

Social poverty functions are characterized by:

Applying dominance conditions to poverty measurement

Social poverty functions are characterized by:

- ▶ Being additive and symmetric across individuals:

$$P(X; z) = \frac{1}{N} \sum_{n=1}^N P_n(x; z).$$

Applying dominance conditions to poverty measurement

Social poverty functions are characterized by:

- ▶ Being additive and symmetric across individuals:

$$P(X; z) = \frac{1}{N} \sum_{n=1}^N P_n(x; z).$$

- ▶ Higher poverty is worse.

Applying dominance conditions to poverty measurement

Social poverty functions are characterized by:

- ▶ Being additive and symmetric across individuals:

$$P(X; z) = \frac{1}{N} \sum_{n=1}^N P_n(x; z).$$

- ▶ Higher poverty is worse.

- ▶ Monotonicity: Higher x should not increase poverty:

$$\frac{\delta P_n(x; z)}{\delta x} \leq 0$$

Applying dominance conditions to poverty measurement

Social poverty functions are characterized by:

- ▶ Being additive and symmetric across individuals:

$$P(X; z) = \frac{1}{N} \sum_{n=1}^N P_n(x; z).$$

- ▶ Higher poverty is worse.

- ▶ Monotonicity: Higher x should not increase poverty:

$$\frac{\delta P_n(x; z)}{\delta x} \leq 0$$

So if we define $W(X; z) = -P(X; Z)$ then we can apply the above mentioned conditions to poverty comparisons as well!

The test of Barrett and Donald (2003)

- ▶ This test is a type of Kolmogorov-Smirnov test of homogeneity of distributions.

The test of Barrett and Donald (2003)

- ▶ This test is a type of Kolmogorov-Smirnov test of homogeneity of distributions.
- ▶ The logic is very intuitive: dominance conditions are based on comparing integrals of $F^A(z)$ with $F^B(z)$ for a range of z , depending on the order of dominance (e.g. for first order, just compare cumulative or survival distributions).

The test of Barrett and Donald (2003)

There are four possible outcomes (illustrated with first-order):

1. A dominates B: $F^A(z) \leq F^B(z) \forall z \wedge \exists z | F^A < F^B(z)$

The test of Barrett and Donald (2003)

There are four possible outcomes (illustrated with first-order):

1. A dominates B: $F^A(z) \leq F^B(z) \forall z \wedge \exists z | F^A < F^B(z)$
2. B dominates A: $F^B(z) \leq F^A(z) \forall z \wedge \exists z | F^B < F^A(z)$

The test of Barrett and Donald (2003)

There are four possible outcomes (illustrated with first-order):

1. A dominates B: $F^A(z) \leq F^B(z) \forall z \wedge \exists z | F^A < F^B(z)$
2. B dominates A: $F^B(z) \leq F^A(z) \forall z \wedge \exists z | F^B < F^A(z)$
3. No dominance because A=B: $F^A(z) = F^B(z) \forall z$

The test of Barrett and Donald (2003)

There are four possible outcomes (illustrated with first-order):

1. A dominates B: $F^A(z) \leq F^B(z) \forall z \wedge \exists z | F^A < F^B(z)$
2. B dominates A: $F^B(z) \leq F^A(z) \forall z \wedge \exists z | F^B < F^A(z)$
3. No dominance because A=B: $F^A(z) = F^B(z) \forall z$
4. No dominance because A and B cross:
 $\exists z_1 | F^A(z_1) < F^B(z_1) \wedge \exists z_2 | F^B(z_2) < F^A(z_2)$

The test of Barrett and Donald (2003)

The generic hypothesis subject to test is the following (illustrated with first-order):

$$H_0 : F^A(z) \leq F^B(z) \forall z \in [z_{min}, z_{max}]$$

The test of Barrett and Donald (2003)

The generic hypothesis subject to test is the following (illustrated with first-order):

$$H_0 : F^A(z) \leq F^B(z) \forall z \in [z_{min}, z_{max}]$$

$$H_1 : \exists z \in [z_{min}, z_{max}] | F^A(z) > F^B(z)$$

The test of Barrett and Donald (2003)

The generic hypothesis subject to test is the following (illustrated with first-order):

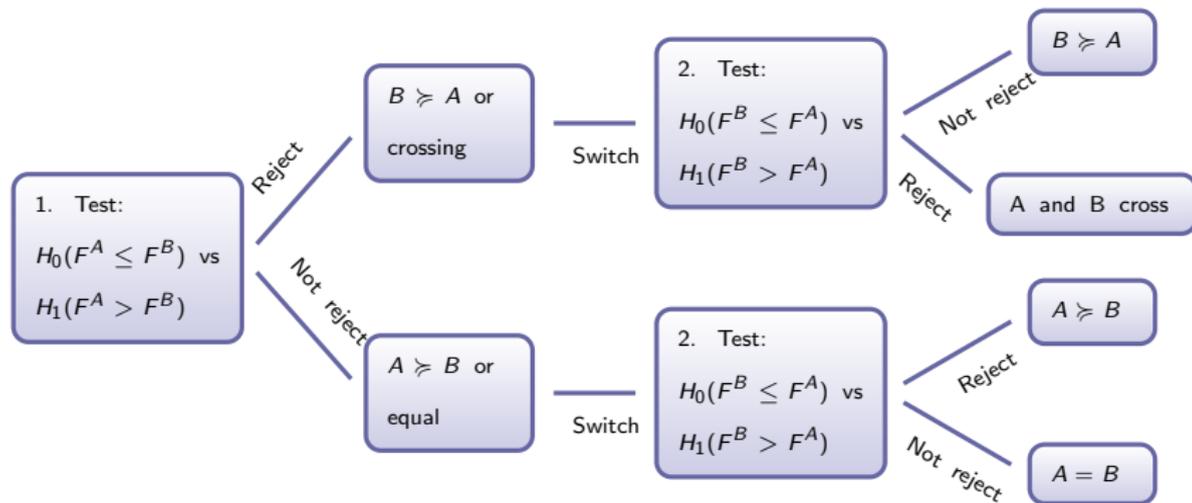
$$H_0 : F^A(z) \leq F^B(z) \forall z \in [z_{min}, z_{max}]$$

$$H_1 : \exists z \in [z_{min}, z_{max}] | F^A(z) > F^B(z)$$

Results of these tests require interpretation in order to ascertain any of the four possibilities.

- Brief introduction to multidimensional stochastic dominance
 - Tests for univariate and multivariate stochastic dominance
 - The test of Barrett and Donald (2003)

The test of Barrett and Donald (2003)



- Brief introduction to multidimensional stochastic dominance
 - └ Tests for univariate and multivariate stochastic dominance
 - └ The test of Barrett and Donald (2003)

The test of Barrett and Donald (2003)

The test proceeds as follows (illustrated with first-order):

The test of Barrett and Donald (2003)

The test proceeds as follows (illustrated with first-order):

- ▶ For a range of z , estimate: $F^A(z) - F^B(z)$

The test of Barrett and Donald (2003)

The test proceeds as follows (illustrated with first-order):

- ▶ For a range of z , estimate: $F^A(z) - F^B(z)$
- ▶ Estimate the supremum and multiply by a function of the sample sizes: $\hat{S} = \left(\frac{NM}{N+M}\right)^{1/2} \sup_z [F^A(z) - F^B(z)]$

The test of Barrett and Donald (2003)

The test proceeds as follows (illustrated with first-order):

- ▶ For a range of z , estimate: $F^A(z) - F^B(z)$
- ▶ Estimate the supremum and multiply by a function of the sample sizes: $\hat{S} = \left(\frac{NM}{N+M}\right)^{1/2} \sup_z [F^A(z) - F^B(z)]$
- ▶ \hat{S} is the statistic we need. We now need to know how likely it is for this value to appear under the null hypothesis.

The test of Barrett and Donald (2003)

The test proceeds as follows (illustrated with first-order):

- ▶ For a range of z , estimate: $F^A(z) - F^B(z)$
- ▶ Estimate the supremum and multiply by a function of the sample sizes: $\hat{S} = (\frac{NM}{N+M})^{1/2} \sup_z [F^A(z) - F^B(z)]$
- ▶ \hat{S} is the statistic we need. We now need to know how likely it is for this value to appear under the null hypothesis.
- ▶ There are different procedures to derive the distribution under the null hypothesis. Barrett and Donald (2003) develop two types. We are going to show one of them: a bootstrap method.

The bootstrap method in the test of Barrett and Donald (2003)

1. We pool the two samples of A and B (a total of $N+M$ observations) and draw several subsamples with replacement (R):

The bootstrap method in the test of Barrett and Donald (2003)

1. We pool the two samples of A and B (a total of $N+M$ observations) and draw several subsamples with replacement (R):
2. For each subsample (say 1000) we estimate:
$$\overline{S}_r = \left(\frac{NM}{N+M}\right)^{1/2} \sup_z [F^A(z; r) - F^B(z; r)]$$

The bootstrap method in the test of Barrett and Donald (2003)

1. We pool the two samples of A and B (a total of $N+M$ observations) and draw several subsamples with replacement (R):
2. For each subsample (say 1000) we estimate:
$$\bar{S}_r = \left(\frac{NM}{N+M}\right)^{1/2} \sup_z [F^A(z; r) - F^B(z; r)]$$
3. The p-value is then approximated by:
$$p^{A,B} \simeq \frac{1}{R} \sum_{r=1}^R 1(\bar{S}_r > \hat{S})$$

The bootstrap method in the test of Barrett and Donald (2003)

1. We pool the two samples of A and B (a total of $N+M$ observations) and draw several subsamples with replacement (R):
2. For each subsample (say 1000) we estimate:
$$\bar{S}_r = \left(\frac{NM}{N+M}\right)^{1/2} \sup_z [F^A(z; r) - F^B(z; r)]$$
3. The p-value is then approximated by:
$$p^{A,B} \simeq \frac{1}{R} \sum_{r=1}^R 1(\bar{S}_r > \hat{S})$$
4. If, say, $p^{A,B} < 0.01$ we reject the null hypothesis: under the hypothesis a value like \hat{S} is very unlikely.

Concluding remarks

- ▶ Stochastic dominance conditions are useful to ascertain whether an *ordinal* comparison is robust to changes in the parameters or members of a family of evaluation functions.

Concluding remarks

- ▶ Stochastic dominance conditions are useful to ascertain whether an *ordinal* comparison is robust to changes in the parameters or members of a family of evaluation functions.
- ▶ When dominance conditions are not fulfilled then the comparison depends on the choice of parameters (e.g. poverty lines, risk/inequality aversion, etc.).

Concluding remarks

- ▶ Stochastic dominance conditions are useful to ascertain whether an *ordinal* comparison is robust to changes in the parameters or members of a family of evaluation functions.
- ▶ When dominance conditions are not fulfilled then the comparison depends on the choice of parameters (e.g. poverty lines, risk/inequality aversion, etc.).
- ▶ One could restrict the dominance analysis to smaller sets of parameters (or members of families of indices), but this needs to be done carefully, lest significant parts of the domain of interest may be left out.

Concluding remarks

- ▶ Stochastic dominance conditions are useful to ascertain whether an *ordinal* comparison is robust to changes in the parameters or members of a family of evaluation functions.
- ▶ When dominance conditions are not fulfilled then the comparison depends on the choice of parameters (e.g. poverty lines, risk/inequality aversion, etc.).
- ▶ One could restrict the dominance analysis to smaller sets of parameters (or members of families of indices), but this needs to be done carefully, lest significant parts of the domain of interest may be left out.
- ▶ When one is concerned for cardinal comparisons, e.g. between countries or across time, then dominance conditions are not that useful. Sensitivity analysis is required.

Concluding remarks

- ▶ Stochastic dominance conditions are useful to ascertain whether an *ordinal* comparison is robust to changes in the parameters or members of a family of evaluation functions.
- ▶ When dominance conditions are not fulfilled then the comparison depends on the choice of parameters (e.g. poverty lines, risk/inequality aversion, etc.).
- ▶ One could restrict the dominance analysis to smaller sets of parameters (or members of families of indices), but this needs to be done carefully, lest significant parts of the domain of interest may be left out.
- ▶ When one is concerned for cardinal comparisons, e.g. between countries or across time, then dominance conditions are not that useful. Sensitivity analysis is required.
- ▶ We have assumed the variables are continuous. But these results can be easily extended to ordinal variables (e.g. see Yalonetzky, 2011). Further extensions to combinations of continuous and ordinal variables are possible.