Jul 23rd, 9:00 AM - 11:00 AM

Well-being of female-headed Indian households: A multidimensional analysis using nonparametric stochastic dominance Presenters

Juan Tellez
Harvard University, juan_tellez@hks.harvard.edu

Jaya Krishnakumar
University of Geneva, jaya.krishnakumar@unige.ch

Well-being of female-headed Indian households: A multidimensional analysis using nonparametric stochastic dominance

Juan Tellez [1] and Jaya Krishnakumar [2]

Second International Conference of the Society of Economic Measurement (SEM), 22-24 July 2015, OECD, Paris, France

[1] Harvard University
[2] University of Geneva

- Basic premise: **Well being is multidimensional**, and needs to be grasped through multiple indicators.

- Basic premise: **Well being is multidimensional**, and needs to be grasped through multiple indicators.
- Many **approaches** advocate such a multidimensional vision: quality of life approach, capability approach, living conditions approach, basic needs approach etc.

- Basic premise: **Well being is multidimensional**, and needs to be grasped through multiple indicators.
- Many **approaches** advocate such a multidimensional vision: quality of life approach, capability approach, living conditions approach, basic needs approach etc.
- If we go by this premise, then the question arises as to **how to combine various indicators** for comparing well-being across individuals or over time?

- Possible solutions:
  - **Composite indices** (from simple averages to generalised means).

- Possible solutions:
  - **Composite indices** (from simple averages to generalised means).
  - **Parametric model based aggregates** : latent variable scores of well-being (FA, MIMIC, SEM, generalised SEM etc).

- Possible solutions:
  - **Composite indices** (from simple averages to generalised means).
  - **Parametric model based aggregates** : latent variable scores of well-being (FA, MIMIC, SEM, generalised SEM etc).
  - **Nonparametric comparisons** of multivariate distributions and testing for dominance.

- In this paper, we apply the third approach to **compare welfare distributions** across different sub-populations.

- In this paper, we apply the third approach to **compare welfare distributions** across different sub-populations.
- Although the use of this methodology to compare univariate distributions is common, extension of the nonparametric **dominance technique to the multivariate case** is recent (cf. Maasoumi and Racine 2013).

- In this paper, we apply the third approach to **compare welfare distributions** across different sub-populations.
- Although the use of this methodology to compare univariate distributions is common, extension of the nonparametric **dominance technique to the multivariate case** is recent (cf. Maasoumi and Racine 2013).
- This technique is particularly suitable in our context given the **multidimensional definition of well-being** that we adopt and allows us to make welfare comparisons without either aggregating over dimensions or over households.

- Our research question : To what extent do **social group** and **education levels** contribute to enhancing or reducing welfare of Indian households, and does education offset the effect of social discrimination?

- Our research question : To what extent do **social group** and **education levels** contribute to enhancing or reducing welfare of Indian households, and does education offset the effect of social discrimination?
- As there is a variety of **socio-economic characteristics** that define an Indian household, and as literature has pointed out large differences in **behaviour** between households headed by men and those headed by women, we decided to mainly focus on women.

- Our research question : To what extent do **social group** and **education levels** contribute to enhancing or reducing welfare of Indian households, and does education offset the effect of social discrimination?
- As there is a variety of **socio-economic characteristics** that define an Indian household, and as literature has pointed out large differences in **behaviour** between households headed by men and those headed by women, we decided to mainly focus on women.
- The reason for selecting **female-headed households** is that there are only a few studies on them as noted by Gangopadhyay and Wadhwa (2004) and we wanted to add to the scarce literature in this domain.

- Nonparametric techniques **do not require** a researcher to specify a **functional form** for an object being estimated. Let us note the density function of a continuous random variable X_j^c as $f(x_j^c)$.

- Nonparametric techniques **do not require** a researcher to specify a **functional form** for an object being estimated. Let us note the density function of a continuous random variable X_j^c as $f(x_j^c)$.

$$\hat{f}(x_j^c) = \frac{1}{nh_j} \sum_{i=1}^{n} k\left(\frac{x_j^c - X_{ij}^c}{h_j}\right),$$

- Nonparametric techniques **do not require** a researcher to specify a **functional form** for an object being estimated. Let us note the density function of a continuous random variable X_j^c as $f(x_j^c)$.

$$\hat{f}(x_j^c) = \frac{1}{nh_j} \sum_{i=1}^{n} k\left(\frac{x_j^c - X_{ij}^c}{h_j}\right),$$

- The following assumptions are made for $k(.)$:
  1. $k(u) = k(-u)$
  2. $\int k(u)du = 1$
  3. $\int uk(u)du = 0$
  4. $\int u^2 k(u)du = \upsilon_2 < \infty$

Conditional density functions and cumulative distribution functions are expressed as:

$$\hat{f}(y|x) = \frac{\hat{f}(y,x)}{\hat{f}(x)}$$

$$\hat{F}(y|x) = \frac{\int^{y} \hat{f}(s,x)ds}{\hat{f}(x)}$$

Conditional density functions and cumulative distribution functions are expressed as:

$$\hat{f}(y|x) = \frac{\hat{f}(y,x)}{\hat{f}(x)}$$

$$\hat{F}(y|x) = \frac{\int^{y} \hat{f}(s,x)ds}{\hat{f}(x)}$$

where

$$\hat{f}(y,x) = n^{-1} \sum_{i=1}^{n} K\gamma_y(Y_i, y) K\gamma_x(X_i, x),$$

$$\hat{f}(x) = n^{-1} \sum_{i=1}^{n} K\gamma_x(X_i, x),$$

and

$$\int^{y} \hat{f}(s,x)ds = n^{-1} \sum_{i=1}^{n} G\gamma_y(Y_i, y) K\gamma_x(X_i, x),$$

- The data comes from the third **National Family Health Survey (NFHS-3)**, which was conducted in two phases, the first from December 2005 to April 2006, and the second from April 2006 to August 2006 by the Ministry of Health and Family welfare with the collaboration of eighteen research organizations in a representative sample for India.

- The data comes from the third **National Family Health Survey (NFHS-3)**, which was conducted in two phases, the first from December 2005 to April 2006, and the second from April 2006 to August 2006 by the Ministry of Health and Family welfare with the collaboration of eighteen research organizations in a representative sample for India.
- To perform our analysis we selected only **women who were head of household** which reduced our sample of interest to 6'086 observations.

- The data comes from the third **National Family Health Survey (NFHS-3)**, which was conducted in two phases, the first from December 2005 to April 2006, and the second from April 2006 to August 2006 by the Ministry of Health and Family welfare with the collaboration of eighteen research organizations in a representative sample for India.
- To perform our analysis we selected only **women who were head of household** which reduced our sample of interest to 6'086 observations.
- Variables of interest: *Wealth Index*, *Level of hemoglobin*, *Level of education* and *Social group*.

- Let us note two distributions A and B, with respective CDFs F_A and F_B. We will say that B **dominates** A **stochastically at first order** if for any r:

$$F_A(r) \geq F_B(r)$$

- Let us note two distributions A and B, with respective CDFs F_A and F_B. We will say that B **dominates** A **stochastically at first order** if for any r:

$$F_A(r) \geq F_B(r)$$

- In order to test the multivariate stochastic dominance we will consider the following functions:

$$F_A = F(y|x = a)$$
$$F_B = F(y|x = b)$$

- Let us note two distributions A and B, with respective CDFs F_A and F_B. We will say that B **dominates** A **stochastically at first order** if for any r:

$$F_A(r) \geq F_B(r)$$

- In order to test the multivariate stochastic dominance we will consider the following functions:

$$F_A = F(y|x = a)$$
$$F_B = F(y|x = b)$$

- We will use the following **Kolmogorov-Smirnov (KS)** statistic as:

$$D = \min[\max(F_A - F_B), \max(F_B - F_A)]$$

- Let us note two distributions A and B, with respective CDFs F_A and F_B. We will say that B **dominates** A **stochastically at first order** if for any r:

$$F_A(r) \geq F_B(r)$$

- In order to test the multivariate stochastic dominance we will consider the following functions:

$$F_A = F(y|x = a)$$
$$F_B = F(y|x = b)$$

- We will use the following **Kolmogorov-Smirnov (KS)** statistic as:

$$D = \min[\max(F_A - F_B), \max(F_B - F_A)]$$

- Assumptions:

$$H_0 = D > 0$$
$$H_a = D \leq 0$$

Three dominance tests:

1.

$$F(wealth, health|education = low, group = j)$$

$$F(wealth, health|education = high, group = j)$$

Three dominance tests:

1.

$$F(wealth, health|education = low, group = j)$$

$$F(wealth, health|education = high, group = j)$$

2.

$$F(wealth, health|education = high, group = j)$$

$$F(wealth, health|education = high, group = i)$$

Three dominance tests:

1.
$$F(wealth, health|education = low, group = j)$$
$$F(wealth, health|education = high, group = j)$$

2.
$$F(wealth, health|education = high, group = j)$$
$$F(wealth, health|education = high, group = i)$$

3.
$$F(wealth, health|education = low, group = j)$$
$$F(wealth, health|education = low, group = i)$$

- For **all four social groups**, women hh with high level of education dominate stochastically women with low level of education at first order.

- For **all four social groups**, women hh with high level of education dominate stochastically women with low level of education at first order.
- For women hh with **low level of education**, there is first order stochastic dominance from the group *None of above* over *Scheduled Castes* and *Other Backward Classes*.

- For **all four social groups**, women hh with high level of education dominate stochastically women with low level of education at first order.
- For women hh with **low level of education**, there is first order stochastic dominance from the group *None of above* over *Scheduled Castes* and *Other Backward Classes*.
- For women hh with **high level of education**, *Scheduled Castes* are stochastically dominated at fist order by all other three social groups.

- The same stochastic dominance tests performed above, were carried out for men head of households in order to be able to highlight the gender differences.

- The same stochastic dominance tests performed above, were carried out for men head of households in order to be able to highlight the gender differences.
- As for women, men with **higher education** dominate those with **low education** for each one of the four social groups.

- The same stochastic dominance tests performed above, were carried out for men head of households in order to be able to highlight the gender differences.
- As for women, men with **higher education** dominate those with **low education** for each one of the four social groups.
- When keeping a **high education level**, there is no dominance for any of the social groups this time, which was not the case for women.

- For a **low education level**, *Scheduled Tribes* are dominated by all the other three groups which is also a different scenario than for women.

- For a **low education level**, *Scheduled Tribes* are dominated by all the other three groups which is also a different scenario than for women.
- The stochastic dominance tests between men and women when both have the same level of education and belong to the same social group are all **in favour of** men.

- For a **low education level**, *Scheduled Tribes* are dominated by all the other three groups which is also a different scenario than for women.
- The stochastic dominance tests between men and women when both have the same level of education and belong to the same social group are all **in favour of** men.
- We extended the analysis but crossing the variables, which means that, we wanted to see if holding high education for women and low for men would be able to **break the dominance**, which is the case for all social groups.

- As a second check, we hold the same level of education (once high and once low) for both women and men, we keep only women belonging to *None of Above* and we compare it with men who are in the other three remmaining social groups. Once again, we observe that **the dominance disappears** in both cases.

- As a second check, we hold the same level of education (once high and once low) for both women and men, we keep only women belonging to *None of Above* and we compare it with men who are in the other three remmaining social groups. Once again, we observe that **the dominance disappears** in both cases.

- We run the tests once again for the trivariate function of Wealth, Health and Education; conditioned on Gender and the Social Group.

- We run the tests once again for the trivariate function of Wealth, Health and Education; conditioned on Gender and the Social Group.

$$F(wealth, health, education|gender, group)$$

- When keeping only women, there is **dominance** of *None of Above* over *Scheduled Castes* and *Other Backward Classes*.

- We run the tests once again for the trivariate function of Wealth, Health and Education; conditioned on Gender and the Social Group.

$$F(wealth, health, education|gender, group)$$

- When keeping only women, there is **dominance** of *None of Above* over *Scheduled Castes* and *Other Backward Classes*.
- For men, *None of Above* **dominate all other three groups** and *Other Backward Classes* **dominate** *Scheduled Castes* and *Scheduled Tribes*.

- The **nonparametric approach** allowed us to draw conclusions that were different from those obtained using **standard approaches** based on the frequency and empirical distributions.

- The **nonparametric approach** allowed us to draw conclusions that were different from those obtained using **standard approaches** based on the frequency and empirical distributions.
- From a methodological point of view, the clear **multivariate dominance** results show that one group is less well-off than another irrespective of where the poverty line may be set in both dimensions as our method compares entire distributions rather than certain derived measures like percentiles.

- **Education** is an important contributor to welfare improvement, but it does not remove all the negatives of social discrimination as even among people with a **high level of education**, some *lower castes* are still dominated, for women.

- **Education** is an important contributor to welfare improvement, but it does not remove all the negatives of social discrimination as even among people with a **high level of education**, some *lower castes* are still dominated, for women.
- There still exist **inequalities of opportunities** even if this social system has been abolished. Government policymakers should be on guard to combat these types of social injustice.

- Stochastic dominance results between the **social groups** of female-headed households and male-headed households are different, showing there is **heterogeneity** in this aspect.

- Stochastic dominance results between the **social groups** of female-headed households and male-headed households are different, showing there is **heterogeneity** in this aspect.
- This methodology has proven to be a powerful instrument and has provided useful and relevant insights that have led to enhanced understanding of **socio-economic characteristics** and **behaviour** of female headed-households in India.

- Stochastic dominance results between the **social groups** of female-headed households and male-headed households are different, showing there is **heterogeneity** in this aspect.
- This methodology has proven to be a powerful instrument and has provided useful and relevant insights that have led to enhanced understanding of **socio-economic characteristics** and **behaviour** of female headed-households in India.

Many thanks for your attention.