REPRESENTATIVENESS OF TOP EXPENDITURES IN ARAB REGION HOUSEHOLD SURVEYS

BY VLADIMIR HLASNY (EWHA WOMANS UNIVERSITY) AND VITO INTINI (UN CAPITAL DEVELOPMENT FUND)

IARIW General conference
Dresden, 21-27 August 2016

Discussant: Jorrit Zwijnenburg

Main topics of the paper

Aim of the paper:

Study the presence of measurement errors among highest expenditure households in five Arab countries (Egypt, Jordan, Palestine, Sudan and Tunisia).

Methodology:

Replacing top expenditures with values predicted under smooth parametric distributions and look at the impact on the Gini coefficient.

Subject:

Per capita household expenditure on the basis of household income and expenditure surveys.

Explaining the issue:
Top twenty per capita observations

Rank in sample	Palestine			Egypt	Jordan
	2007	**2010**	**2011**	**2010**	**2010**
1	71.1	306.5	115.0	42.2	216.5
2	49.7	131.2	89.1	36.7	30.4
3	41.2	83.1	80.5	29.2	25.9
4	40.4	74.5	62.7	24.9	25.0
5	31.4	69.0	59.1	24.6	24.7
6	29.0	64.4	52.2	23.1	23.5
7	28.4	63.9	50.7	21.9	21.3
8	28.0	62.7	49.8	21.6	20.7
9	27.1	56.2	47.6	21.1	20.7
10	25.	5	46	18	2
11	24.3	50.7	45.7	17.6	19.8
12	24.1	50.3	44.5	17.2	19.7
13	23.7	47.3	43.9	15.7	19.6
14	23.4	46.8	43.3	15.4	19.0
15	23.0	45.7	43.1	14.8	18.7
16	22.9	45.2	41.9	14.1	18.6
17	22.1	43.6	41.5	14.1	18.3
18	21.4	42.7	41.1	13.8	18.2
19	21.1	42.6	40.6	13.7	18.1
20	19.7	41.9	39.8	13.5	17.9

x 1.4   x 2.3   x 1.4   x 1.1   x 7.1

x 3.6   x 7.3   x 2.9   x 3.1   x 12.1

International dollars (x1000), purchasing power parity (UNSD 2015)

Replacement alternatives

The paper describes three alternatives to replace the top expenditures from the surveys:

1. Using values from a <u>Pareto distribution</u>:
   - Parametric distribution which is skewed and heavy-tailed
   - Motivated by an empirical regularity that top observations across countries and years follow a particular pattern.

2. Using values from a '<u>generalized beta type II' (GB2)</u> distribution:
   - Understood to better represent the entire expenditure distribution.

3. Using <u>randomly drawn values</u> instead of predicted values
   - Combines random values from the estimated distribution for top observations with actual lower-level values. Via repeating exercises an expected measure of inequality can be obtained.

Results

The correction for non-representative distribution of top expenditures with Pareto distributional results varies across the countries and surveys:

- Replacing top 0.1-5 percent:
  - In Egypt replacement leads to a small but systematic increase in the Gini.
  - In Jordan (2006), Palestine and Tunisia results are nearly identical.
  - In Jordan (2010) replacement leads to a large drop in the Gini.
  - In Sudan replacements leads to a small but systematic drop in the Gini.
- Replacing top 10-20 percent:
  - Replacement increases Gini coefficients in all countries.

Does the cut-off point matter?

Nonparametric vs. quasi-nonparametric Gini, Pales. '10
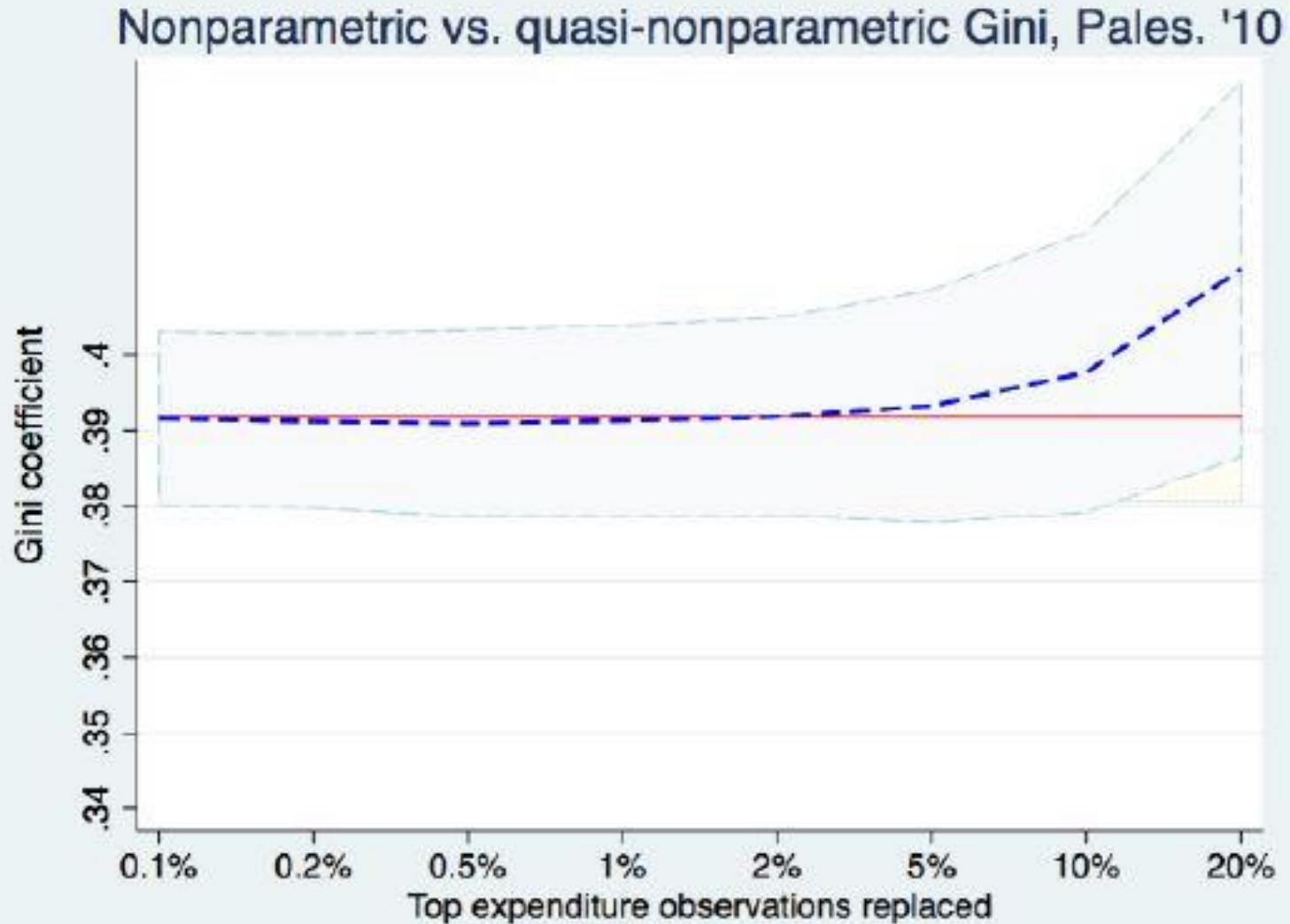
How appropriate is the Pareto distribution? – Hill plots

Hill plots show how the estimated Pareto parameter changes as one changes the delimitation of top incomes:

- Flat lines mean a stable Pareto distribution for the entire top.

- Non-stationary lines imply different parametric distributions for the subgroups within top incomes.

How appropriate is the Pareto distribution? – Hill plots

Hill plot, sampling-weighted sample
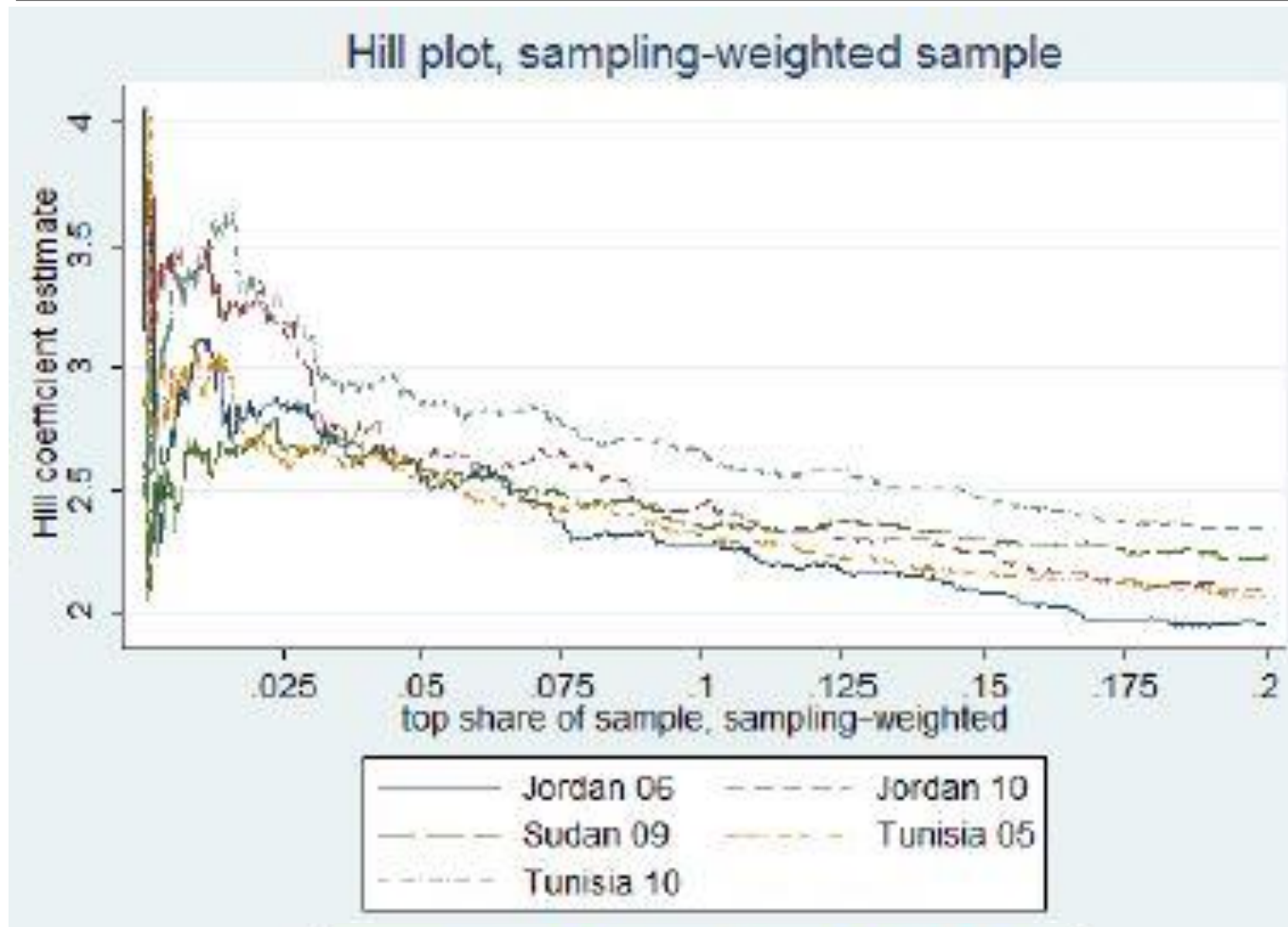
Core Jordan, Sudan and Tunisia different parametric distributions would characterize different subgroups of top incomes.

How appropriate is the Pareto distribution? – Comparison with GB2

- GB2 also leads to different results for countries when comparing with Gini on the basis of actual values.

- Corrections under GB2 distribution differ systematically from those under the Pareto distribution:

  - For Jordan and Palestine GB2 estimates exceed Pareto estimates.

  - For Sudan and Tunisia Pareto estimates exceed GB2 estimates.

  - For Egypt it depends on the reference year.

- However, in only few cases these differences appear to be significant (confidence intervals on the basis of random draw techniques).

Conclusion of the paper

- Similarity of the results of the various models suggests that none of the resulting Gini's can be rejected.

- Whether the actual (non-parametric) or parametric (either Pareto or GB2) Gini coefficients are closest to the true Gini's remains a question for further research.

- In the meantime, it is important to not rely on a single measure, but also consider alternative measures.

Discussion

- This paper adds relevant insights to the research on the tails of the distribution.

- There are several methods to check the plausibility of these results (for example compare results across data sources, compare results over time, confront with macro results) of which comparison with some parametric distributions is an important one.

- In order to draw conclusions on the representativeness of the tails in the surveys on the basis of these comparisons one has to assume these models are applicable for the relevant countries.

- The impact on Gini seems small, although the impact on more specific indicators may still be large.

Discussion

Some questions:

- Is there any information on the general applicability of these parametric models for the countries in the study?

- Do you have any other information that may point to issues in the micro data (such as gaps between micro and macro aggregates)?

- Do you plan to apply the same techniques to income results? If so, are you planning to also confront income and consumption data on a micro-level?

- Do you plan to analyse the impact on other measures which may be more sensitive to extreme observations?

Thank you for your attention!