Measurement of Policy Outcomes: Homework 4

Please send your answers before November 4th, 2019 11:59 A.M. to
measurement.homework@gmail.com

Date of the tutorial session: November 6th, 2019

Important notes:

Please try to keep your answers within a reasonable length. 4 pages of written answers (excluding graphs and R scripts) should be enough.

As always, try to give complete and self-contained answers. Please make explicit references to your R script or comment it. The link between your answers and your R script should be elicited (or obvious).

Inequality in Côte d'Ivoire

You are provided with a dataset called civ_87_08.dta. It is compiled from a repeated survey which measured living standards in Côte d'Ivoire in 1987, 1998 and 2008. A new sample is drawn at each round, which means that households are not followed over time (this type of data is commonly known as *repeated cross-section*). The dataset also includes average prices of food and non-food products for a representative household of 4 members, male-headed and whose head is a blue-collar worker, either self-employed outside agriculture or a farmer.

The average food price is computed for an individual's daily meal which is composed of coffee and bread for breakfast, rice and peanut sauce for lunch and fried plantain and fish for dinner. Average price of non-food products includes daily rents, price of clothing, schooling fees, utilities, transportation, communication, information and leisure.

1 Sampling and outliers (3 points)

(a) In terms of sampling, what are the main limitations of survey data for measuring inequality?

(b) Why can it be important to exclude outliers for measuring inequality?

(c) Compute the level of consumption per capita per each household.

(d) It is proposed to computed the (non-weighted) median and standard deviation of the logarithm of consumption per capita for each year separately, and then to exclude those households whose logarithm consumption per capita is higher than the median by more than 3 standard deviations, or lower than the median by more than 3 standard deviations.

Table 1: Average Price in Côte d'Ivoire from 1987 to 2008[*].

	1987	1998	2008
Panel A: North.			
Food	1005	1805	2410
Non-food product	395	715	955
Panel B: South.			
Food	1215	2335	3105
Non-food product	705	1360	1805

[*] Price levels are in current local currency unit (*franc CFA*). They are computed for the purpose of this exercise and are to some extent fictitious.

– Why is it suggested to take the logarithm? Why can $\pm\, 3$ standard deviations be deemed rather safe?

– Why should you use the unweighted median rather than the weighted median to identify outliers?

– Withdraw these observations from your database. This should not drop more than 157 observations in total.

2 Consumer Price Index and Purchasing Power Parities (4 points)

The point of this section is to compute Consumer Price indices and purchasing power parity aggregates.

(a) Compute the budget shares of food and non-food consumption for the representative households in each region.

– Create a dummy variable which identifies the representative households.

– Generate the share of food expenditure in the household consumption, for each household.

– Build a weight equal to the number of people in the household multiplied by the weight of the household multiplied by the consumption per capita of the household.

– Calculate the weighted average of the share of food expenditure for the representative household in North and South Côte d'Ivoire using the weights that you just computed, and show mathematically why you should use these weights. Table 2 below sums up the results that you should be observing.[1]

(b) All information displayed in Table 1 and Table 2 is available in the dataset through the variables `avprice_food`, `avprice_other`, `bdgshr_food`, `bdgshr_other`. *Please use these variables from now on (do not use the variables that you have computed yourself above).* With 1987 as the base year, compute:

– the Laspeyres CPI for each region.

– the Paasche CPI for each region.

---

[1]It goes without saying that it is expected from you to display your own results and the way you got there.

Table 2: Budget Shares in Côte d'Ivoire from 1987 to 2008.

	1987	1998	2008
Panel A: North.			
Food	0.72	0.73	0.61
Non-food product	0.28	0.27	0.39
Panel B: South.			
Food	0.63	0.66	0.55
Non-food product	0.37	0.34	0.45

Comment on the differences between the two measures of inflation in a maximum of three sentences.

(c) Compute a PPP aggregate for the North region with respect to the South. For each year in your dataset, compute:

- the Laspeyres PPP of Northern Côte d'Ivoire.
- the Paasche PPP of Northern Côte d'Ivoire.
- the Fisher PPP of Northern Côte d'Ivoire.

Comment on the differences between the three measures in a maximum of three sentences.

(d) Table 3 below sums up the results that you should be observing.[2] *Please use these variables from now on (do not use the variables that you have computed yourself above).* Use the Laspeyres CPI and PPP available in Table 3 and convert the values of the consumption variable for all years and regions into comparable values, taking as a reference the South of Côte d'Ivoire in 1987.

Table 3: Laspeyres CPI and PPP Across Regions.

	CPI			PPP*		
	1987	1998	2008	1987	1998	2008
Regions:						
- North	100.00	180.00	240.36	72.91	68.91	66.39
- South	100.00	192.45	255.73	100.00	100.00	100.00

* A PPP of 72 in 1987 reads as follows: What you could buy with 100 CFA franc in the South was worth 73 CFA francs in the North.

---

[2]Again, it goes without saying that it is expected from you to display your own results and the way you got there.

3 Absolute poverty measures (5 points)

The variable `conspc_87` represents the daily per capita consumption measured in PPP 1987 franc CFA. To measure poverty over time and between countries, Martin Ravallion, a famous poverty economist, suggested in 1990 an absolute poverty line of 1\$ PPP a day.[3]

(a) The poverty threshold of one dollar a day is equivalent to 237 franc CFA 1987.

 – Build a dummy variable equal to 1 if daily per capita consumption, `conspc_87`, is below the threshold and 0 if it is above the threshold.

 – Build a weight equal to the number of people in the household times the weight of the household.

 – Calculate the poverty headcount rate over time, across regions and at the national level.

 – Comment on the evolution of poverty rates across regions and over time.

(b) Now, the poverty threshold is set to two dollars a day, or 474 franc CFA in 1987.

 – Calculate the poverty headcount rate over time, across regions and at the national level.

 – Compare your results to those observed with a threshold of one dollar a day. What do you observe?

(c) Assume that price levels and preferences do not differ across regions. This means that the CPI observed in the South is also valid in the North. Likewise, the PPP factor is 1 across regions.

 – Create a new `cpi` variable which expands the CPI observed in the South to both regions.

 – Use the new `cpi` variable and daily per capita consumption at current prices (`conspc`) and generate daily per capita consumption in constant 1987 franc CFA, (name it `newconspc_87`).

 – Build a dummy variable equal to 1 if `newconspc_87` is below the threshold of 237 franc CFA a day, and 0 otherwise.

 – Calculate the poverty headcount rate over time, across regions and at the national level.

 – Compare your results to those observed when you accounted for differences of prices and choices across regions. What do you observe?

4 Growth incidence curves (4 points)

The growth incidence curve (GIC) plots the growth rate at each quantile of per capita consumption.
The GIC of daily per capita consumption measured in constant 1987 franc CFA in Côte d'Ivoire puts:

 – on the horizontal axis the percentiles **p** of `conspc_87`

 – on the vertical axis the average annual growth rate **g(p)** at each percentile of `conspc_87`

And:

$$g_{t_0 \to t_n}(p) = \sqrt[n]{\frac{y_{t_n}(p)}{y_{t_0}(p)}} - 1$$

---

[3]This definition has then been adopted by the World Bank, then re-evaluated multiple times, which always stirs debate. One of today's widespread poverty lines is fixed at 1.90 \$2011 a day, and is very much in use in some poverty assessment works, notably by the World Bank; it is the close parent of the previous 1.25 \$2005. Imagine the consequences of re-evaluating the history of prices in a country!

(a) For each year of your dataset, associate to each percentile its corresponding per capita consumption measure using the variable `conspc_87`.[4] [5]

(b) Compute the average growth for each percentile over the periods 1987-98, 1998-08.

(c) Draw the two curves on the same graph.

(d) Comment the growth incidence curves in 3 sentences at most.

(e) Given the GIC between 1998 and 2008, is inequality likely to fall or rise in 2008 ? Justify your answer in 2-3 sentences.

5   Lorenz curves (4 points)

The Lorenz curve is a consensual tool for analyzing inequality. For a given variable X, it puts:

- on the horizontal axis the individuals ranked by ascending order of X, or percentile $p_{X,i}$ (where, for instance, $1 \leqslant i \leqslant 100$)

- on the vertical axis the cumulative shares of X, $l(p_{X,i})$ And:

$$l(p_i) = \frac{\sum_{j \leqslant i} w_j x_j}{\sum_j w_j x_j} \quad \text{with} \quad p_i = \frac{\sum_{j \leqslant i} w_j}{\sum_j w_j}$$

(a) For each of the 3 years, for the North, the South, and Côte d'Ivoire as a whole, compute the Lorenz curves of daily per capita consumption in PPP 1987 franc CFA and draw them on the same graph.[6] [7]

(b) Do these Lorenz curves cross? If yes, where?[8] Comment on the change of inequality between 1987, 1998 and 2008 in Côte d'Ivoire (max. 200 words)

> *a word of caution from your TA:* To the best of my knowledge, available user-written R commands will *not* allow you to answer this question precisely (and thus, correctly). You are therefore *strongly* encouraged to work out the curves manually.

---

[4]Hint: Use the command ewcdf() evaluated on the conspc87 variable to compute the percentiles.

[5]Hint: For this question it is particularly crucial not to forget to use weights at all times

[6]Hint: The easiest way to draw the Lorenz curve for a given variable is to compute its percentiles, and, for each percentile, compute the cumulative share of the variable until this percentile (using the cumsum function), and then plot, in a graph, the shares of each percentile on the y axis against the percentiles on the x axis.

[7]Hint: For this question it is particularly crucial not to forget to use weights at all times

[8]Hint: To spot crossings accurately, you can for instance compute pairwise differences between the Lorenz curves that you draw, and examine whether these differences change sign at some point.