An Algorithm for Validating Basic Heading Prices

Robert J. Hill and Michael Scholz
University of Graz
Austria
robert.hill@uni-graz.at michael-scholz@uni-graz.at

20-21 May 2013

ICP TAG Meeting

Validation of the Basic Heading CPD Price Indexes

- Select a region.
- Select a country b in the region as the base.
- Let P^i_{bk} denote the CPD price index for basic heading i in country k with country b as the base.
- For country k find the maximum ratio of its CPD price indexes as follows:

$$MaxCPDRatio_k = \max_{i,j=1,\ldots,I} \left(\frac{P^i_{bk}}{P^j_{bk}} \right) = \frac{P^{i*}_{bk}}{P^{j*}_{bk}},$$

  where $i*$ and $j*$ are the headings that solve the maximization problem.
- If $MaxCPDRatio_k \leq Z$, then move on to the next country. (Note: Z is chosen as part of the validation process - for example it could be set to 40).

- If $MaxCPDRatio_k > Z$, then we conclude that there may be a problem with the price data of one of the following four country-basic headings:
  (i) Heading $i*$ in the base country b.
  (ii) Heading $i*$ in country k.
  (iii) Heading $j*$ in the base country b.
  (iv) Heading $j*$ in country k.
- The remainder of the algorithm is focused on establishing which of these four country-basic headings needs checking.

- Calculate the geometric mean of the basic heading price indexes for country k, excluding headings $i*$ and $j*$, as follows:

$$GM_k = \prod_{i \neq i*, j*}^{I} (P_{bk}^{i})^{1/(I-2)},$$

$$\text{Now let } \; A = \frac{P_{bk}^{max}}{GM_k}, \quad B = \frac{GM_K}{P_{bk}^{min}}.$$

- If $A > B$ we conclude that heading $i*$ needs checking [i.e., (i) or (ii)].
- If $A < B$ we conclude that heading $j*$ needs checking [i.e., (iii) or (iv)].
- In what follows we will assume that $A > B$. The algorithm proceeds in an analogous way if $A < B$.

- Calculate the geometric mean across all countries of the heading $i*$ price indexes deflated by market exchange rates (X), first with country b as the base and then with country k as the base.
- Alternatively, heading $i*$ price indexes could be deflated by the GEKS purchasing power parity (PPP) exchange rates. The advantage of using market exchange rates is that they are unaffected by errors in the data. The disadvantage of market exchange rates are that they can differ quite significantly from the underlying PPP-exchange rates.

$$GM_b^{i*} = \prod_{j=1}^{K} \left(\frac{P_{bj}^{i*}}{X_{bj}} \right)^{1/K}, \quad GM_k^{i*} = \prod_{j=1}^{K} \left(\frac{P_{kj}^{i*}}{X_{kj}} \right)^{1/K}.$$

- If $GM_b^{i*} > GM_k^{i*}$ we conclude that it is the price index of country b for heading $i*$ that is the outlier [i.e., case (iii)].

- Conversely, if $GM_b^{i*} < GM_k^{i*}$ we conclude that it is the price index of country k for heading $i*$ that is the outlier [i.e., case (iv)].
- For the minimum cases where $A < B$, we have that:

$$GM_b^{j*} = \prod_{j=1}^{K} \left(\frac{X_{bj}}{P_{bj}^{j*}} \right)^{1/K}, \quad GM_k^{j*} = \prod_{j=1}^{K} \left(\frac{X_{kj}}{P_{kj}^{j*}} \right)^{1/K}.$$

- If $GM_b^{j*} > GM_k^{j*}$ we conclude that it is the price index of country b for heading $j*$ that is the outlier [i.e., case (i)].
- If $GM_b^{j*} < GM_k^{j*}$ we conclude that it is the price index of country k for heading $j*$ that is the outlier [i.e., case (ii)].

- The selected country-basic heading is now deleted, and the algorithm is rerun. This process continues (i.e., if necessary more country-basic heading price indexes are deleted) until the inequality $MaxCPDRatio_k \leq Z$ is satisfied.
- Once the inequality is satisfied the algorithm moves on to the next country in the region. Once all countries have been checked, the algorithm changes the base country and repeats the process. The algorithm terminates once all countries in that region have been used as the base country.
- Move on to the next region.