Short-term movements of the RY-GEKS price index: is the failure at the identity test really a problem?

Ottawa Group meeting
Copenhagen, May 2013

Claude Lamboray
claude.lamboray@statec.etat.lu
How to use scanner data?

1) Sample of products & fixed basket
   - One-to-one replacements that can be done with the help of an automatic pre-selection, possibly followed by manual intervention

2) “Census” of products & chaining
   - Risk of chain drift
     → RY-GEKS has been proposed as a solution to this problem.
RY-GEKS

The RY-GEKS is based on an underlying price index (Fisher, Tornqvist).

The method compares various periods over a 13-month moving window.

\[ RY-GEKS_{t-1,t} = \left( \frac{P_{t-12,t}}{P_{t-12,t-1}} \cdot \frac{P_{t-11,t}}{P_{t-11,t-1}} \cdot \ldots \cdot \frac{P_{t-2,t}}{P_{t-2,t-1}} \cdot (P_{t-1,t})^2 \right)^{1/13} \]
Identity test

The **identity test** states that if prices stay constant in two comparison periods, then the price index should show no movement whatever the quantities observed.

The RY-GEKS violates the identity test.

→ need to assess the risk of possibly too unnatural behaviour of the index, due to the failure on the identity test (M. Ribe 2012, *Some properties of the RGEKS index for scanner data*).
Identity test

Focus on short-term movements: only adjacent periods

Better understanding of the behaviour of the RY-GEKS

An index that fails the identity test and that passes the continuity test can yield results that go in the opposite direction of the prices.
Identity test

<table>
<thead>
<tr>
<th>Period</th>
<th>Price 1</th>
<th>Price 2</th>
<th>Quant. 1</th>
<th>Quant. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>10</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-2</td>
<td>100</td>
<td>10</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>t-1</td>
<td>10</td>
<td>100</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>t</td>
<td>10</td>
<td>100</td>
<td>100</td>
<td>10</td>
</tr>
</tbody>
</table>

Prices are identical in \( t-1 \) and in \( t \).

The RY-GEKS decreases by 75% between \( t-1 \) and \( t \).
Change of the RY-GEKS if prices are identical

We now assume that prices are identical in \( t-1 \) and in \( t \).

The change between \( t-1 \) and \( t \) of the RY-GEKS is driven by a \textbf{weighted relative covariance} (Bortkiewicz decomposition):

\[
RYGEKS_{t-1,t} = \prod_{m=2}^{12} \left( 1 + \text{RelCov}_{wm} \left( \frac{p_t}{p_{t-m}}, \frac{x_t}{x_{t-1}} \right) \right)^{1/26}
\]
Change of the RY-GEKS if prices are identical

We now assume that prices are identical in $t-1$ and in $t$.

The change between $t-1$ and $t$ of the RY-GEKS is driven by a **weighted relative covariance** (Bortkiewicz decomposition):

$$RYGEKS_{t-1,t} = \prod_{m=2}^{12} \left( 1 + RelCov_{wm} \left( \frac{p_t}{p_{t-m}}, \frac{x_t}{x_{t-1}} \right) \right)^{1/26}$$

The failure at the identity test can be attenuated if

- the variance of price relatives is low
We now assume that prices are identical in \( t-1 \) and in \( t \).

The change between \( t-1 \) and \( t \) of the RY-GEKS is driven by a **weighted relative covariance** (Bortkiewicz decomposition):

\[
RYGEKS_{t-1,t} = \prod_{m=2}^{12} \left( 1 + RelCov_{wm} \left( \frac{p_t}{p_{t-m}}, \frac{x_t}{x_{t-1}} \right) \right)^{1/26}
\]

The failure at the identity test can be attenuated if
- the variance of price relatives is low
- the variance of quantity relatives is low
Change of the RY-GEKS if prices are identical

We now assume that prices are identical in $t-1$ and in $t$.

The change between $t-1$ and $t$ of the RY-GEKS is driven by a \textbf{weighted relative covariance} (Bortkiewicz decomposition):

$$RYGEKS_{t-1,t} = \prod_{m=2}^{12} \left( 1 + \text{RelCov}_{wm} \left( \frac{p_t}{p_{t-m}}, \frac{x_t}{x_{t-1}} \right) \right)^{1/26}$$

Compare this to the unit value index:

$$UV_{t-1,t} = 1 + \text{RelCov}_{wm} \left( p_t, \frac{x_t}{x_{t-1}} \right)$$
Simulations

Real scanner data set available at STATEC (one shop of a major retail chain).

Prices in $t\!-\!1$ are replaced by prices in $t$ and quantities are left unchanged.

<table>
<thead>
<tr>
<th></th>
<th>Prices</th>
<th>Quantities</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t!-!1$ (September 2012)</td>
<td>$p_t$</td>
<td>$x_{t!-!1}$</td>
</tr>
<tr>
<td>$t$ (October 2012)</td>
<td>$p_t$</td>
<td>$x_t$</td>
</tr>
</tbody>
</table>
Simulations

Change of the RY-GEKS if prices in t-1 are replaced by prices in t.
Simulations

Relative variance of price relatives.
Simulations

Relative variance of quantity relatives.
Simulations

On average, 69% of the prices are identical in period $t$ and in period $t-1$.

We only keep those items with identical prices in $t$ and in $t-1$.

<table>
<thead>
<tr>
<th>Prices</th>
<th>Quantities</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{t-1}$</td>
<td>$x_{t-1}$</td>
</tr>
<tr>
<td>$p_t = p_{t-1}$</td>
<td>$x_t$</td>
</tr>
</tbody>
</table>
Simulations

Change of the RY-GEKS if only those items are kept with identical prices in $t$ and in $t-1$. 

![Graph showing the change of the RY-GEKS for different items.](image-url)
Conclusion

If prices are identical in two adjacent periods, then the change of the RY-GEKS depends on a weighted covariance between price relatives (period t compared to period t-m) and quantity relatives (period t compared to period t-1).

To what extent can we accept such short-term movements when compiling price indices based on scanner data?

What can be done to mitigate the problem?

How does this insight contribute to the discussion on the different ways of using scanner data?