Choice of window length and linking method for extending multilateral index series

Antonio Chessa, Florin Barb and Stefan Boumans
Team CPI, Statistics Netherlands

Ottawa Group meeting
7-10 June 2022, Rome, Italy
Outline

• Motivation of study
• Problem statement and characterisation
• Empirical study
• Results
• Main findings and next steps
Motivation of study

• Current extension method (FBEW) will be replaced in our CPI in 2023

• A set of components and prerequisites are being defined as part of our implementation plan

• New methodological questions:
  • How do different extension methods behave during lockdown?
  • Same question with regard to high inflation in 2022
Past work on extension problem

- **Initial work (< 2020, ref’s at the end):**
  - Focus on window length \( W = 13 \) months
  - Studies differ in size and scope (methods, product groups)
  - Results are difficult to compare, do not point towards a preferred method

- **More recent work:**
  - Big differences for seasonal items (Chessa, 2021)
  - Methods show drift for \( W = 13 \)
  - Community now seems to converge towards \( W = 25 \)

- **Behaviour of methods still needs to be better understood** (see also previous slide)
Extension problem: (1) Compute 1st index series

Compute indices time window 1

\[ P_0^{[1]} \quad P_1^{[1]} \quad P_2^{[1]} \quad P_3^{[1]} \quad \ldots \quad P_{11}^{[1]} \quad P_{12}^{[1]} \]
(2) Shift window, calculate 2nd series

COMPUTE indices

- time window 1

\[ P_0^{[1]} \quad P_1^{[1]} \quad P_2^{[1]} \quad P_3^{[1]} \quad \ldots \quad P_{11}^{[1]} \quad P_{12}^{[1]} \]

- time window 2

\[ P_1^{[2]} \quad P_2^{[2]} \quad P_3^{[2]} \quad \ldots \quad P_{11}^{[2]} \quad P_{12}^{[2]} \quad P_{13}^{[2]} \]
(3) Link and extend series 1

COMPUTE indices
time window 1

\[ P_0^{[1]} \quad P_1^{[1]} \quad P_2^{[1]} \quad P_3^{[1]} \quad \ldots \quad P_{11}^{[1]} \quad P_{12}^{[1]} \]

Link to

\[ P_1^{[2]} \quad P_2^{[2]} \quad P_3^{[2]} \quad \ldots \quad P_{11}^{[2]} \quad P_{12}^{[2]} \quad P_{13}^{[2]} \]

and/or

COMPUTE indices
time window 2

EXTEND 1st index series

\[ P_0^{[1]} \quad P_1^{[1]} \quad P_2^{[1]} \quad P_3^{[1]} \quad \ldots \quad P_{11}^{[1]} \quad P_{12}^{[1]} \quad ? \]
(3) Link and extend series 1

- **Non-revisable indices**
  - \( P_0^{[1]} \) \( P_1^{[1]} \) \( P_2^{[1]} \) \( P_3^{[1]} \) \( \ldots \) \( P_{11}^{[1]} \) \( P_{12}^{[1]} \)

- **Link to**
  - \( P_0^{[2]} \) \( P_1^{[2]} \) \( P_2^{[2]} \) \( P_3^{[2]} \) \( \ldots \) \( P_{11}^{[2]} \) \( P_{12}^{[2]} \) \( P_{13}^{[2]} \)

- **Transitive**

- **Generally not transitive, result depends on linking**

- **Extend 1st index series**
  - \( P_0^{[1]} \) \( P_1^{[1]} \) \( P_2^{[1]} \) \( P_3^{[1]} \) \( \ldots \) \( P_{11}^{[1]} \) \( P_{12}^{[1]} \) ?
Implications

• Characterisation of extension methods needed

• Loss of transitivity ⇒
  • How can “drift” be controlled in extended (i.e. published) indices?
  • Benchmark needed
Characterisation of extension methods

• **Time window variables:**
  • Window length \( W \)
  • Temporal window adjustment

• **Linking variables:**
  • Linking period(s) or interval \( L \)
  • Index in linking period(s)
### Examples of extension methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Window length</th>
<th>Window adjustment</th>
<th>Linking month</th>
<th>Linking index</th>
<th>Extension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window splice</td>
<td>13 months</td>
<td>Rolling</td>
<td>First month window</td>
<td>Last recalculated</td>
<td>Year on year</td>
</tr>
<tr>
<td>Movement splice</td>
<td>Arbitrary</td>
<td>Rolling</td>
<td>Penultimate month</td>
<td>Published</td>
<td>Month on month</td>
</tr>
<tr>
<td>FBRW</td>
<td>≥ 13 months</td>
<td>Rolling</td>
<td>Base month</td>
<td>Published</td>
<td>Fixed base</td>
</tr>
<tr>
<td>FBEW</td>
<td>Starts with 2 mths</td>
<td>Expanding</td>
<td>Base month</td>
<td>Published</td>
<td>Fixed base</td>
</tr>
<tr>
<td>WISP</td>
<td>13 months</td>
<td>Rolling</td>
<td>First month</td>
<td>Published</td>
<td>Year on year</td>
</tr>
<tr>
<td>HASP</td>
<td>25 months</td>
<td>Rolling</td>
<td>Central month</td>
<td>Published</td>
<td>Year on year</td>
</tr>
<tr>
<td>Mean splice</td>
<td>Any</td>
<td>Rolling</td>
<td>All possibilities</td>
<td>Recalc./publ.</td>
<td>Average index</td>
</tr>
</tbody>
</table>

FBRW = Fixed Base Rolling Window  
FBEW = Fixed Base Expanding Window  
WISP = Window Splice on Published indices  
HASP = Half Splice on Published indices  

Note: Combination of variables (blue columns) leads to type of index extension (green column)
Example 1: Movement splice (W = 13)

COMPUTE indices time window 1

COMPUTE indices time window 2

EXTEND 1st index series

\[ \begin{align*}
    P_0^{[1]} & \quad P_1^{[1]} & \quad P_2^{[1]} & \quad P_3^{[1]} & \quad P_4^{[1]} & \quad P_5^{[1]} & \quad P_6^{[1]} & \quad P_7^{[1]} & \quad P_8^{[1]} & \quad P_9^{[1]} & \quad P_{10}^{[1]} & \quad P_{11}^{[1]} & \quad P_{12}^{[1]} \\

    P_1^{[2]} & \quad P_2^{[2]} & \quad P_3^{[2]} & \quad P_4^{[2]} & \quad P_5^{[2]} & \quad P_6^{[2]} & \quad P_7^{[2]} & \quad P_8^{[2]} & \quad P_9^{[2]} & \quad P_{10}^{[2]} & \quad P_{11}^{[2]} & \quad P_{12}^{[2]} & \quad P_{13}^{[2]} \\

    \text{Month on month index month 13 in time window 2} \\
\end{align*} \]

\[ P_{13}^{[1]} = \frac{P_{13}^{[2]}}{P_{12}^{[2]}} \]
Example 2: WISP (W = 13)
Example 3: HASP (W = 13)

1. **Compute indices**
   - Time window 1:
     - \( P_0^{[1]} \) to \( P_{12}^{[1]} \)
   - Time window 2:
     - \( P_1^{[2]} \) to \( P_{13}^{[2]} \)

2. **Half year index month 13 in window 2**

3. **Extend 1st index series**

   \[
   = P_7^{[1]} \frac{P_{13}^{[2]}}{P_7^{[2]}}
   \]
Empirical study: (1) Data sets

• Main study:
  • Transaction data sample covering 11.6% of Dutch CPI
  • Product groups in almost all COICOP divisions (2-digit level)
  • Period: 4-5 years

• Additional study:
  • Focus on high inflation in 2022
  • Selection of product groups (e.g. oils and fats, cereals)
Empirical study: (2) Methods

• Window length:
  • $W = 13, 25, 37$ months
  • Full period (used to compare extended index series)

• Extension methods:
  • WISP, HASP, MOSP (movement splice), MESP (mean splice)
  • Compared with CPI (FBEW method is used for most T-data sets)
  • Index method: Geary-Khamis (Dutch CPI since Jan. 2016)

• Implementation:
  • Data are stored on SQL Server
  • Transact SQL procedure used for index calculations
Empirical study: (3) Assessment

• yoy indices extended series minus yoy full window series

• Impact of yoy differences in CPI:
  • Measured at all-items level
  • Absolute differences with benchmark < 0.05 pp
  • See also Article 2(21) of EU Regulation 2016/792
Results: (1) Full sample, W = 13

Left axis: yoy diff's full sample
Right axis: Impact all-items CPI
Results: (2) Full sample, W = 25 and 37

Left axis: yoy diff’s full sample
Right axis: Impact all-items CPI
Results: (3) Drift and volatility

Left axis: yoy diff’s full sample
Right axis: Impact all-items CPI
Results: (4) yoy deviations 2-digit level, W = 25
Results: (5) High inflation 2022, W = 25

Graphs show yoy extended series – yoy full window (pp)
Trimmed mean splice: Linking between 6 and 18m ago
Deviation in yoy indices

• **Drift:**
  - W = 13 is often too short, not only for seasonal items
  - MOSP is particularly affected (mom chaining), irrespective of W

• **Volatility:**
  - First months of window not suited as linking months either
  - Lead to data censoring problems, e.g.:
    - When products leave after linking month at discount prices
    - Also in other situations with changing price and sales dynamics around linking month
  - Problem in WISP, irrespective of W, but not in HASP

• **Above effects do not simply average out in MESP**
Message: Centrality in linking month

First month time window  Linking month  Current month

Time lag for controlling volatility  Time lag for controlling drift
Next steps in Dutch CPI

• Switch to HASP-25 for all product types (T-data)

• Monitoring options in production:
  • Compare with HASP(W=37, L=12), e.g. in cases with sparse data
  • Compare with most recent 25-month transitive series, which is useful for detecting new entries with high introduction prices
Thank you!
Questions?

Contact: ag.chessa@cbs.nl
Research funded by Eurostat grant IMREVA 101034180 – 2020-NL-PRICE
References (1)


References (2)


