Using multilateral hedonic methods to capture product relaunches

Ken Van Loon, Adrien Mierop & Dorien Roels

STATBEL
Belgium in figures

economie

.be

STAT**BEL**

- Shrinkflation or product relaunches where the price changes proportionally less than the packing size are problematic in scanner data

- Products tend to get a new unique product identifier

- Typically matched-methods are used with scanner data, these product relaunches are missed

Cat-astrophe? Whiskas pet food criticised over shrinking servings

Cat food brand caught in shrinkflation storm as owners say new pouches and recipe leaves pets hungry

(The Guardian, 2023)

STAT**BEL**

- Statbel has been using scanner data from supermarkets to compile the CPI since 2015

- Method: GEKS-Törnqvist/CCDI multilateral method, with half splice on published indices with a 25-month window

- Stock keeping units as product identifiers

  - They don't capture shrinkflation

  - Solution: semi-automatic procedures using text mining and manual verification by price collectors

  - Problem: time consuming, possibility of human errors, problematic cases when time period between new and discontinued product is too large

STAT**BEL**

- Since 2022 we use scanner data for consumer electronics and household appliances in the CPI

- Specific characteristics of these segments:

  - High attrition rate of products (short life cycle)

  - Products have high entry and low exit price

  - Different features of products leaving and entering the market

- Need to take the difference between disappeared and new products into account

- Method: ITGEKS with bilateral time dummy hedonic indices

- Idea: what if we could use these multilateral hedonic methods to also capture shrinkflation or product relaunches?
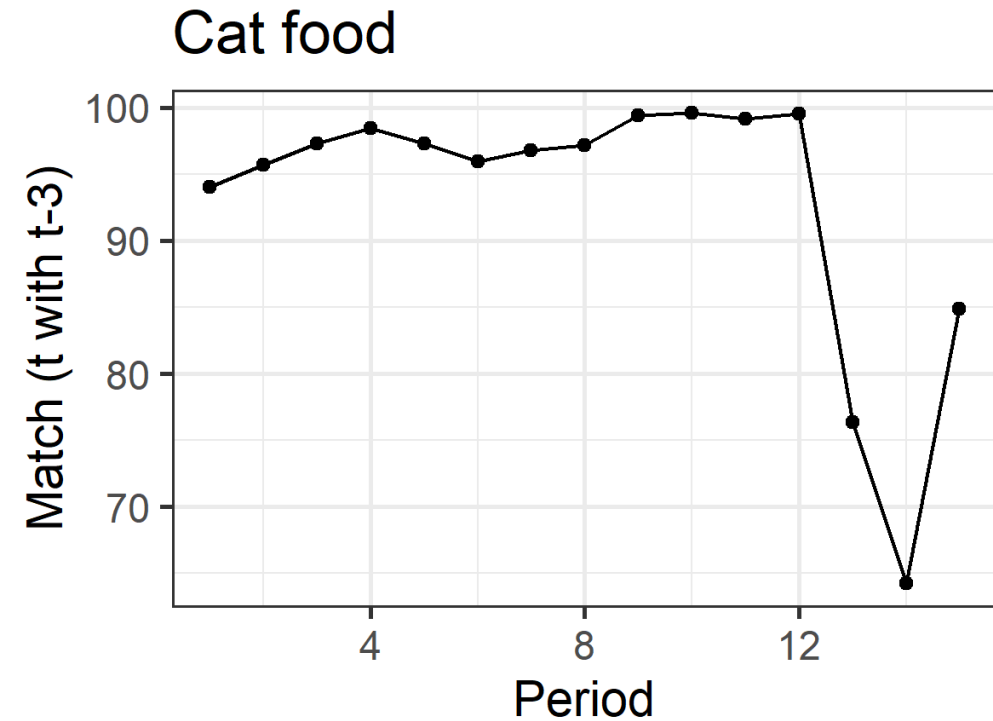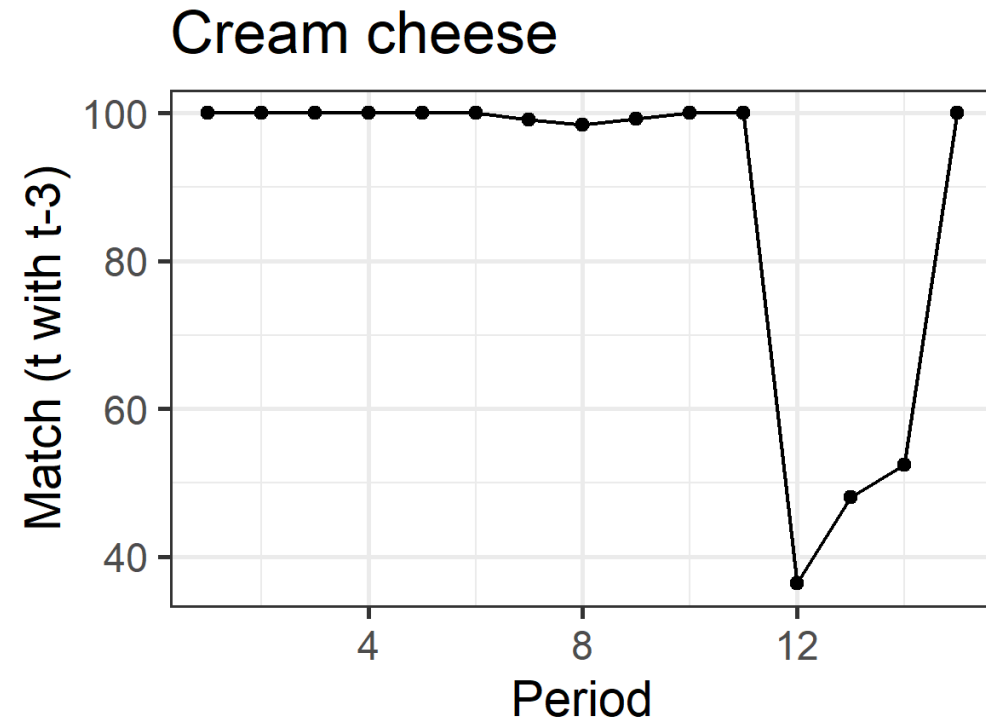
STAT**BEL**

- Scanner data of supermarket chains used in production for the CPI/HICP:

  - A product identifier

  - Number of sales and turnover

  - Detailed product descriptions (i.e. separate variables for brand, variety and other information)

  - Separate variables for package size and the unit of measure (kilograms, litre, …)

  - Our classification to ECOICOP

- Merged with the internal classification datasets of supermarket chains

- Product groups with problematic product relaunches (in this case shrinkflation) were examined as well as "normal" product groups

STAT**BEL**

Cream cheese — Cat food
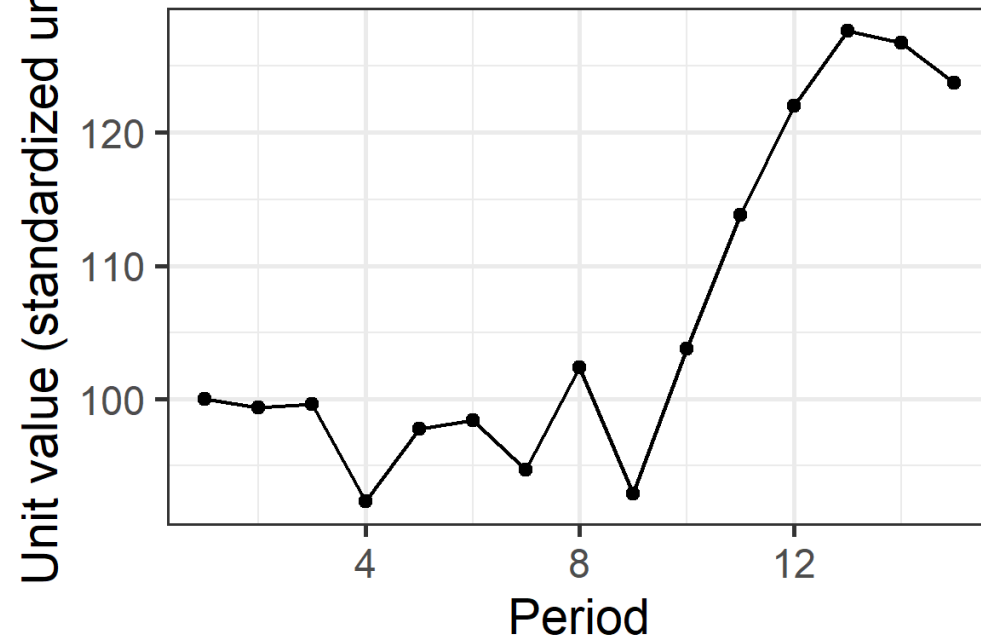
- Both segments experienced a significant product relaunch, e.g. for cream cheese in month 12, only ±40% of the number of sales of month 9 can be matched

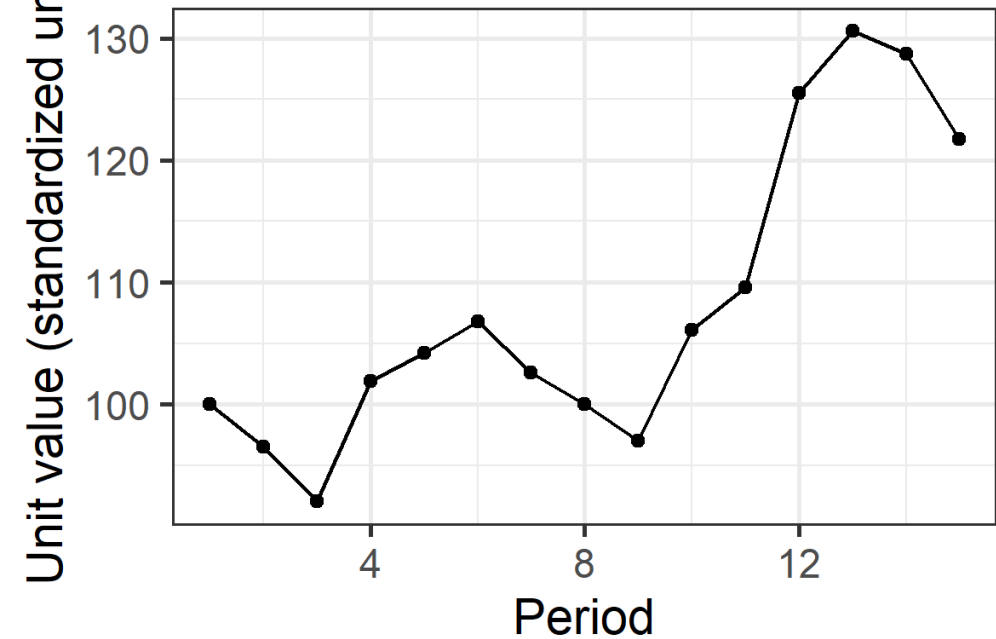STAT**BEL**

Cream cheese

Cat food

- Relaunches in both segments coincide with an increase in the standardised unit value price index (price/kg), indicating shrinkflation
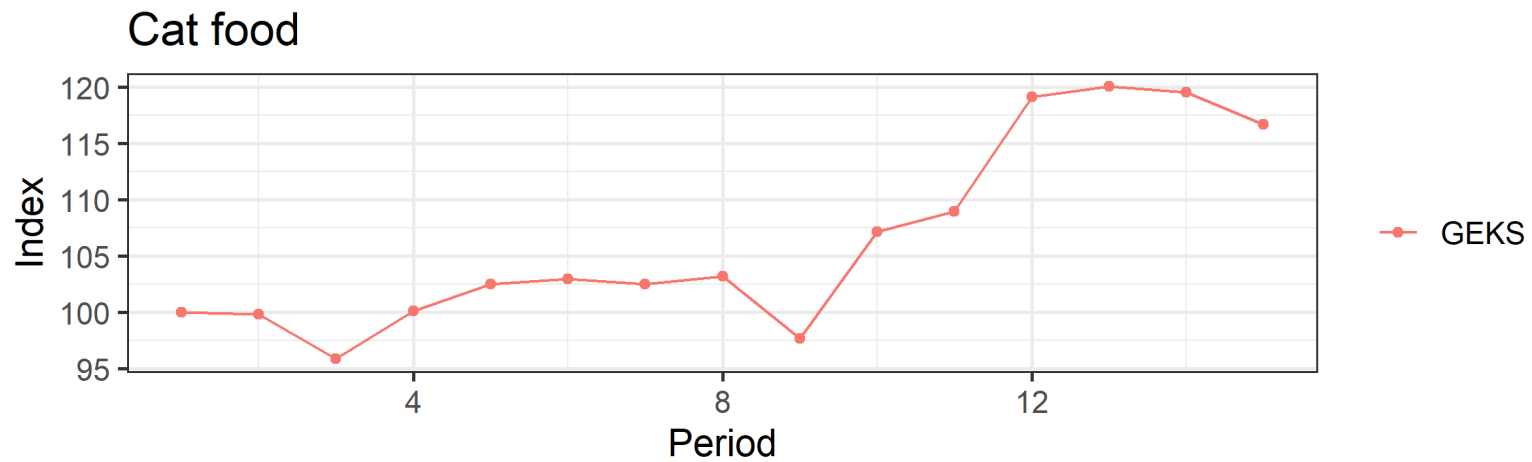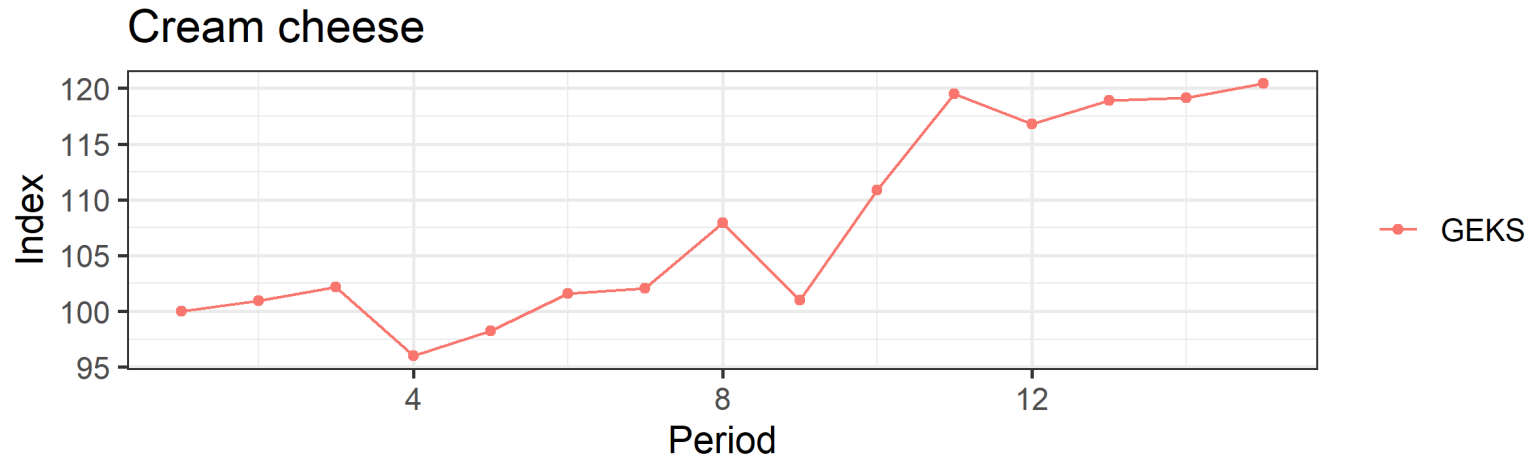
STAT**BEL**

▪ Standard formula for the GEKS

$$P_{GEKS}^{0,t} = \prod_{l=0}^{T} \left(P^{0l}/P^{tl} \right)^{(1/T+1)} = \prod_{l=0}^{T} \left(P^{0l} P^{lt} \right)^{(1/T+1)}$$

▪ With P^{0l} and P^{lt} being Törnqvist-indices:

$$P_T^{0,t} = \prod_{i=1}^{n} \left(\frac{p_i^t}{p_i^0} \right)^{0.5 \left(\frac{p_i^0 q_i^0}{\sum_{j=1}^{n} p_j^0 q_j^0} + \frac{p_i^t q_i^t}{\sum_{j=1}^{n} p_j^t q_j^t} \right)} = \prod_{i=1}^{n} \left(\frac{p_i^t}{p_i^0} \right)^{0.5 \left(s_i^0 + s_i^t \right)}$$

▪ This is a matched items index using unique product identifiers

STAT**BEL**

Cream cheese

Cat food

- GEKS index shows hardly any price increase at all from period 12, for cream cheese it even shows a small price decline at the moment of the product relaunch

9

STAT**BEL**

- Log-linear specification, with expenditure shares in each period serving as weights

$$\ln p_i^t = \alpha + \sum_{t=1}^{T} \delta^t D_i^t + \sum_{k=1}^{K} \beta_k z_{ik} + \varepsilon_i^t$$

- Prices of all items (i) for several periods (t) are pooled in the same regression, on their characteristics (z_{ik}) and on dummy variables for the periods (D_i^t)

- Advantage: its simplicity, since the index follows directly from the estimated time dummy parameters

- Disadvantage: 1) it forces parameter fixity for the whole window and 2) without product churn the index does not equal the GEKS

STAT**BEL**

Cream cheese
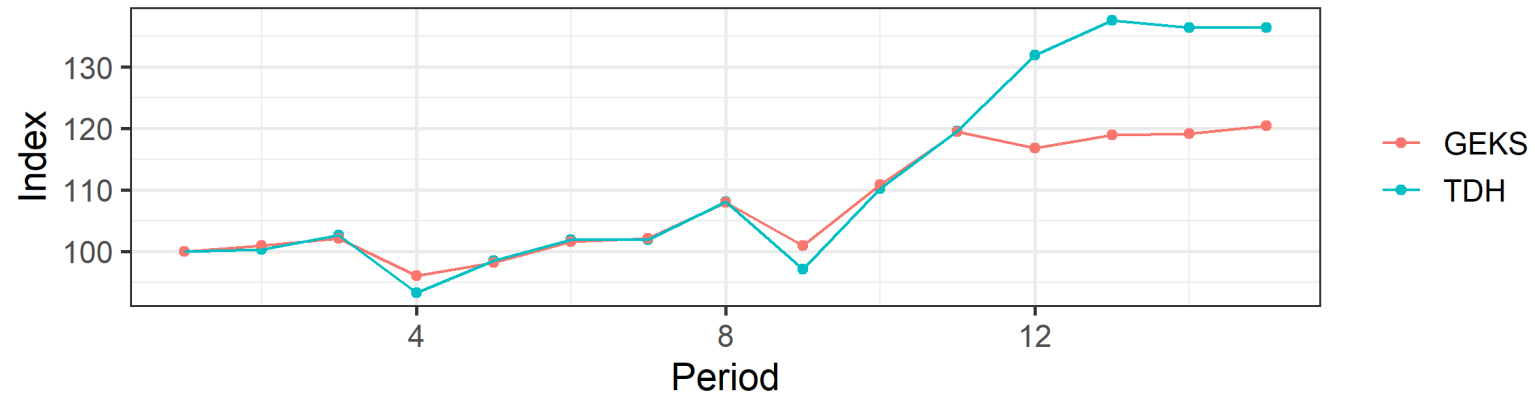
Cat food
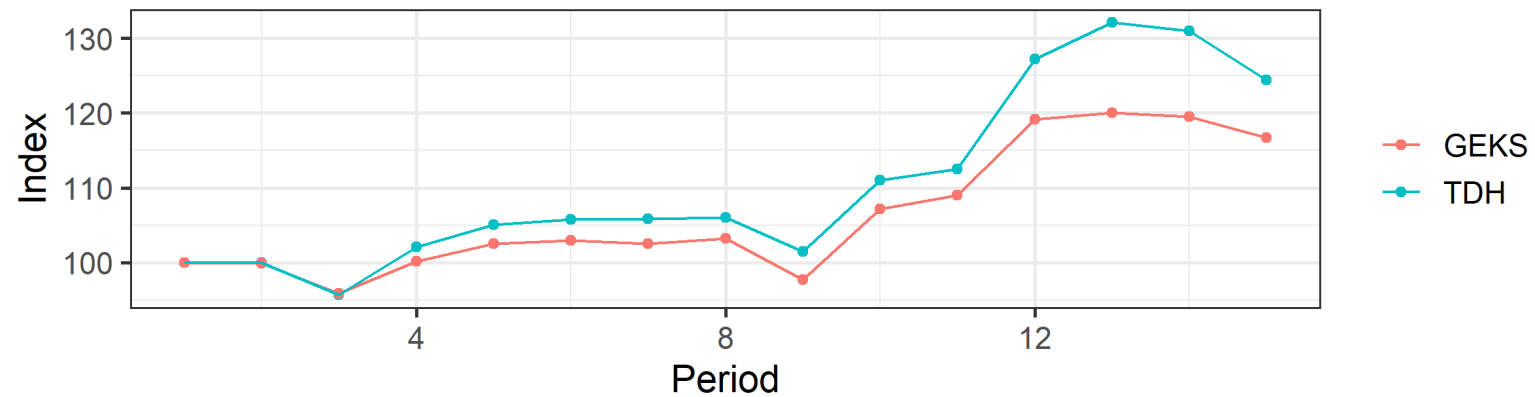
▪ TDH shows a correct price increase when the shrinkflation takes place in both product groups

STAT**BEL**

- Method we use for consumer electronics and household appliances scanner data

- ITGEKS with bilateral time dummy hedonic indices as inputs for the GEKS Törnqvist

- In the GEKS formula the Törnqvist-indices (P^{0l} and P^{lt}) are replaced with weighted bilateral time dummy hedonic indices

$$\ln p_i^t = \alpha + \delta^t D_i^t + \sum_{k=1}^{K} \beta_k z_{ik} + \varepsilon_i^t$$

- Advantage: fixity of the parameters only for the two periods being compared

- Disadvantage: many bilateral regressions must be run → window period of 25 months requires 300 bilateral time dummy hedonic regressions.
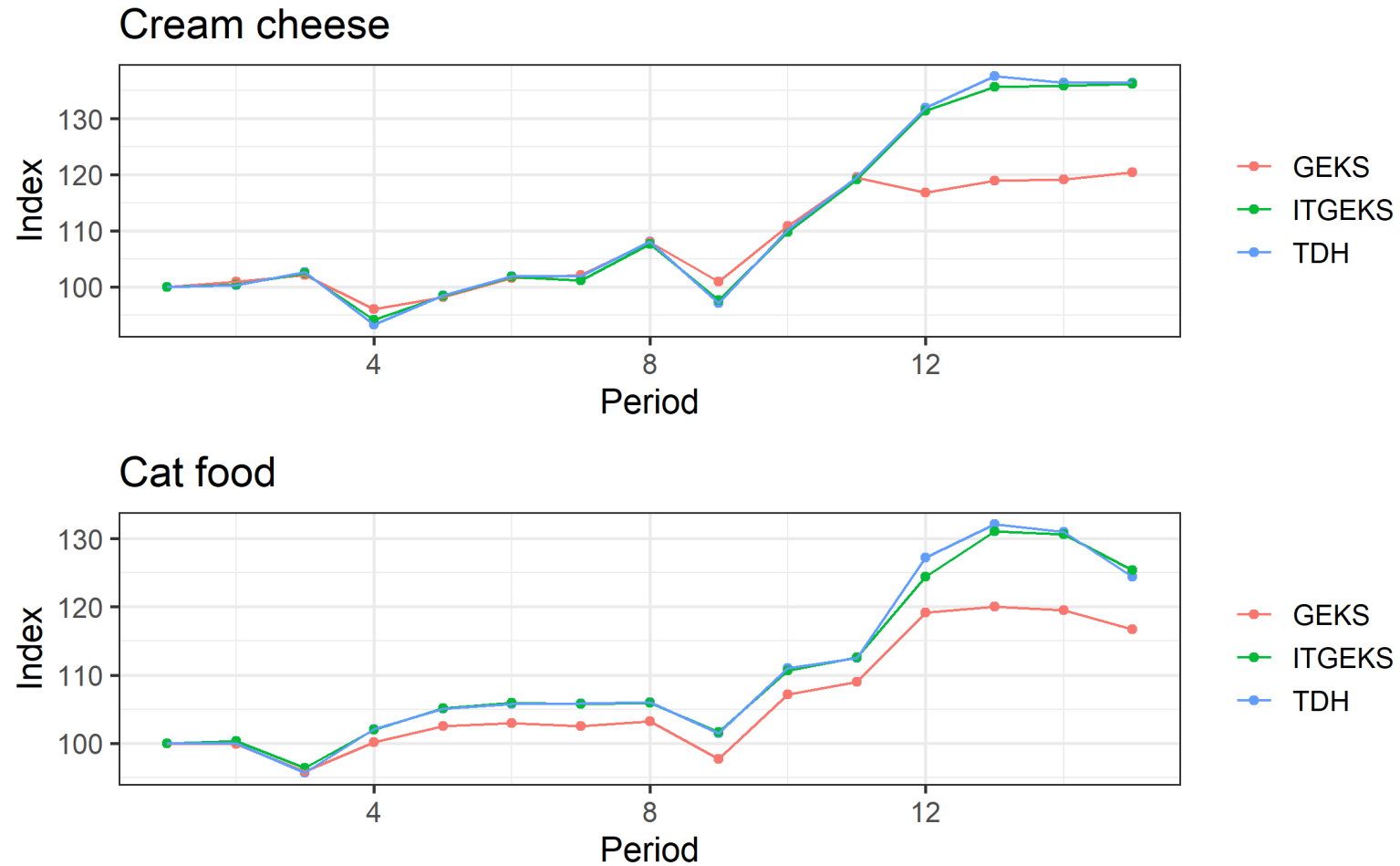
STAT**BEL**

- Which weights are used to estimate the bilateral time dummy hedonic indices?

- Mean expenditure shares for matched items (U_M^{0t}) and half expenditure shares for new (U_N^{0t}) and disappeared items (U_D^{0t}) → makes it is algebraically equivalent to:

$$P_{ITGEKS}^{0,t} = \prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0}\right)^{0.5\,(s_i^0 + s_i^t)} \prod_{i \in U_D^{0t}} \left(\frac{\hat{p}_i^t}{p_i^0}\right)^{0.5\,(s_i^0)} \prod_{i \in U_N^{0t}} \left(\frac{p_i^t}{\hat{p}_i^0}\right)^{0.5\,(s_i^t)}$$

- Advantage compared to TDH: without product churn the index equals the GEKS

STAT**BEL**

Cream cheese

Cat food

▪ Difference between the ITGEKS and the TDH is limited. Explained by the high R squared values (0.903 - 0.963), also indicates that parameter fixity is not that problematic.

14

STAT**BEL**

- Extended the analysis to 4 randomly selected product groups where our "semi-automatic" had not found any problematic product relaunches: coffee, chocolate, soft drinks and breakfast cereals.

- These segments have product attrition, but no product relaunches that could result in a biased index

- Traditional GEKS Törnqvist could serve as a benchmark

- 40 months period, requires 780 bilateral time dummy hedonic regressions in the ITGEKS

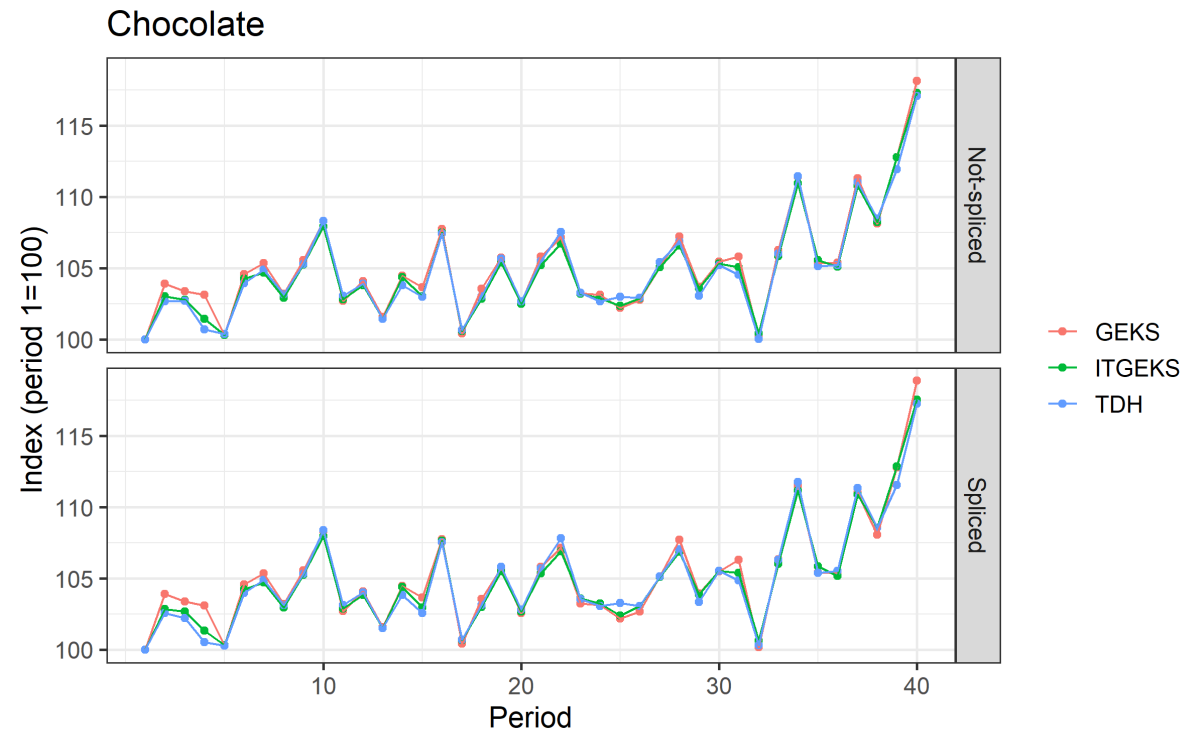- Compare index using entire window to an index with splicing (HASP-25)
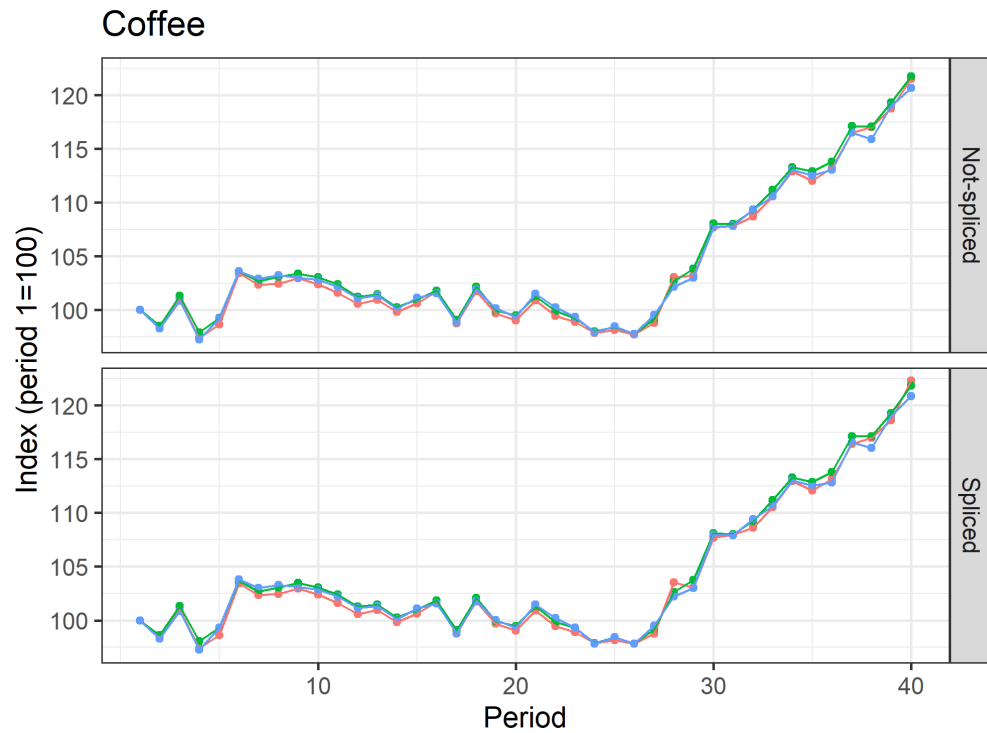
STAT**BEL**

- All 4 segments have similar conclusions and splicing doesn't make a difference

STAT**BEL**

▪ All 4 segments have similar conclusions and splicing doesn't make a difference

Soft drinks

Breakfast cereal
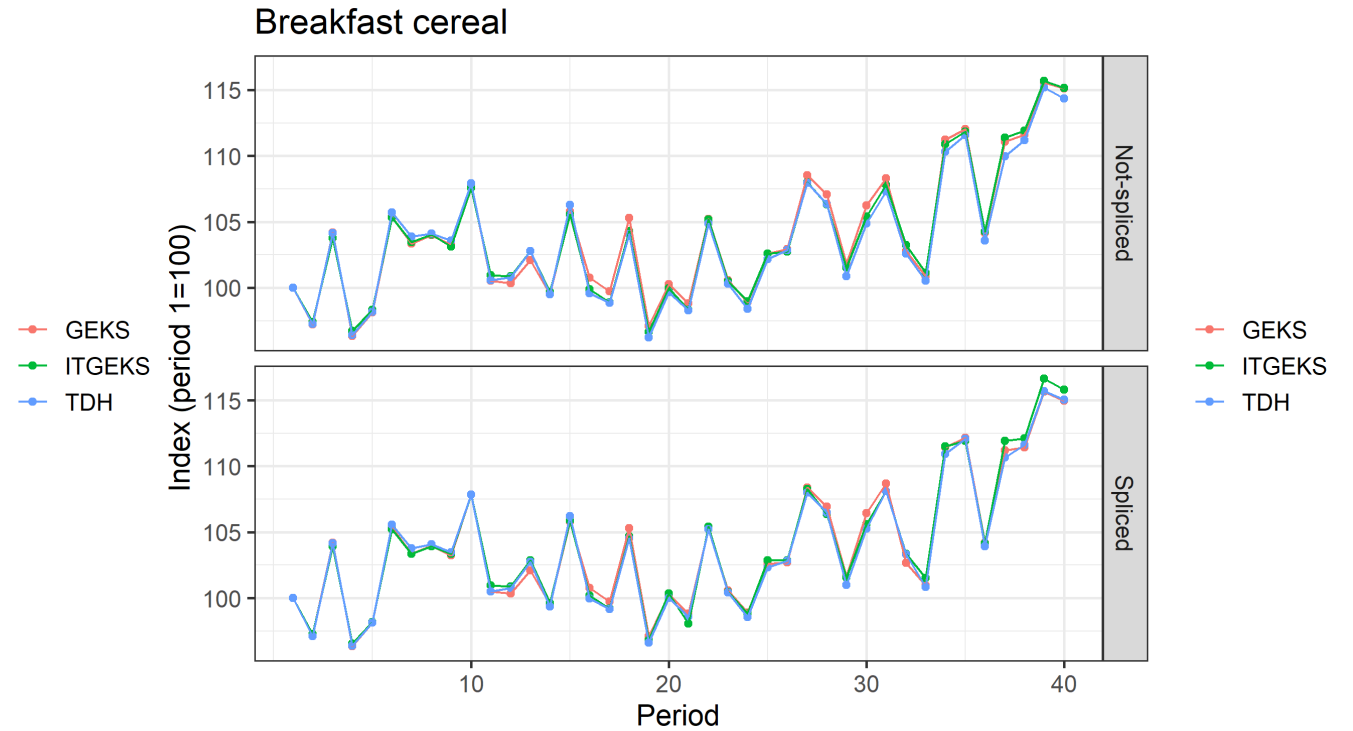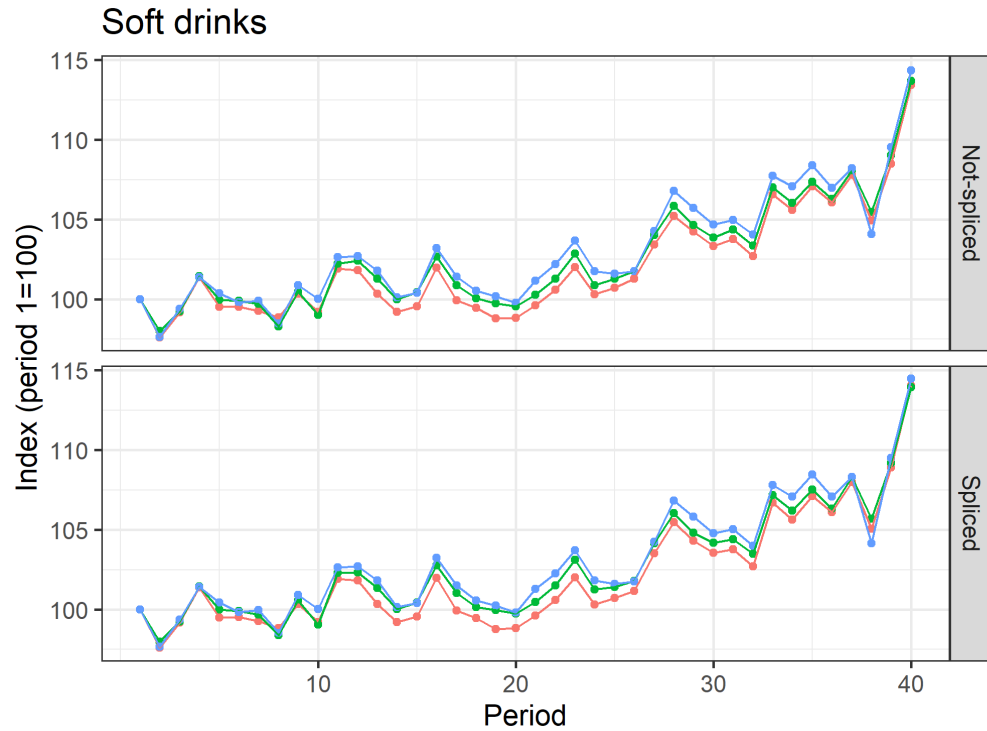
STAT**BEL**

- Conclusion

  - using unique product identifiers to compile the index can lead to a (downward) bias

  - multilateral methods which use hedonics are applicable to supermarket scanner data

  - these methods can capture (certain) product relaunches or shrinkflation

  - difference between ITGEKS and the much easier to compile TDH was limited

  - splicing does not change the conclusions

  - even if countries lack sufficient metadata to do hedonics: still useful to compile a matching sales indicator and a standardised unit value price index to identify potential problematic groups

- Next steps:

  - extend our analysis to other product groups

  - compare with methods that use stratification or product clustering