Estimating Chain Drift in Price Indexes Based on Scanner Data

Jan de Haan
Heymerik van der Grient

Background

Extensive use of scanner data in Dutch CPI (re-design)

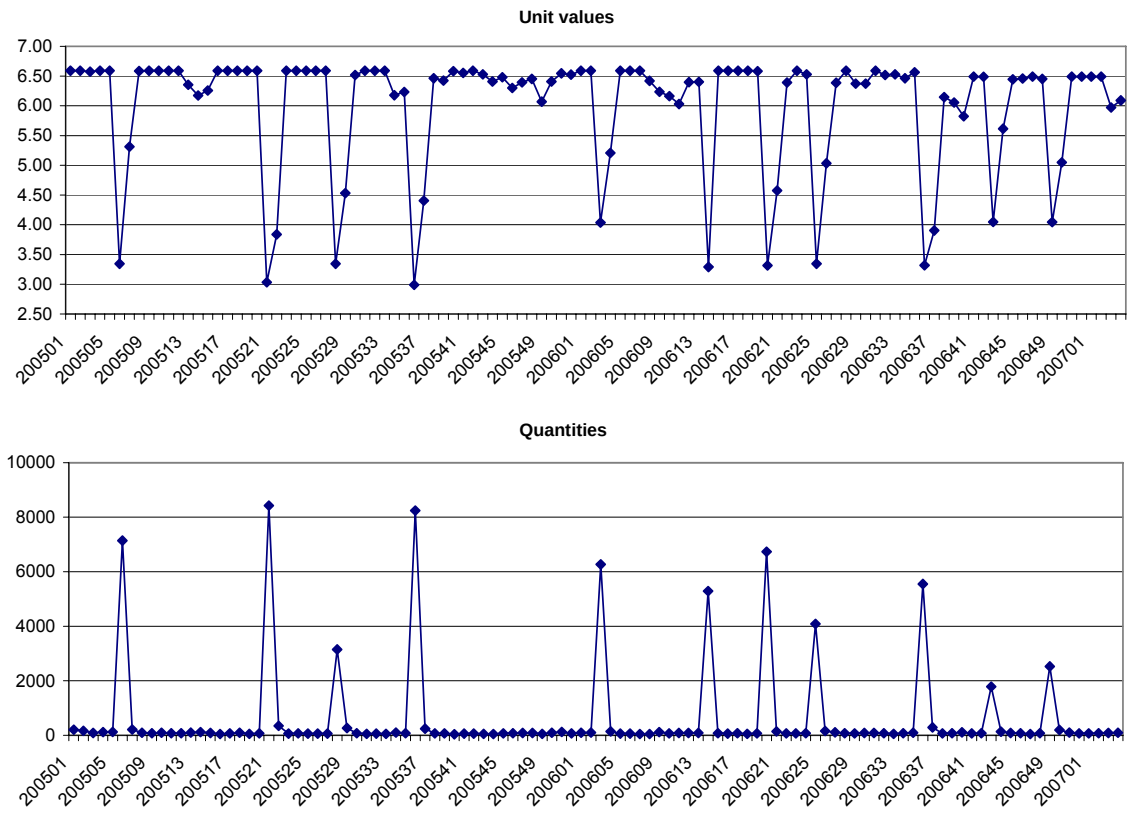
Standard index number theory breaks down for detailed high frequency data

Recent paper by Ivancic, Fox and Diewert (2009)

Features of scanner data

Price and quantity bouncing due to sales

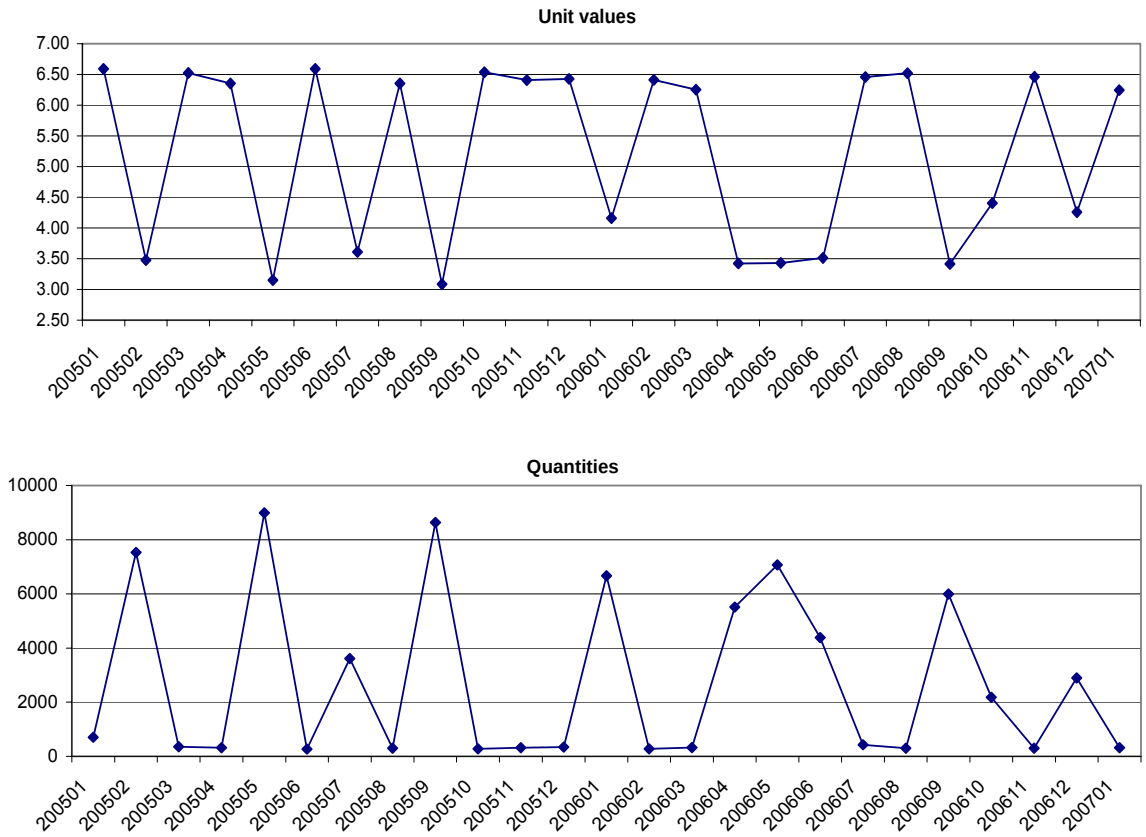
Weekly unit values and quantities; XXX tablets (detergent)



Features of scanner data (2)

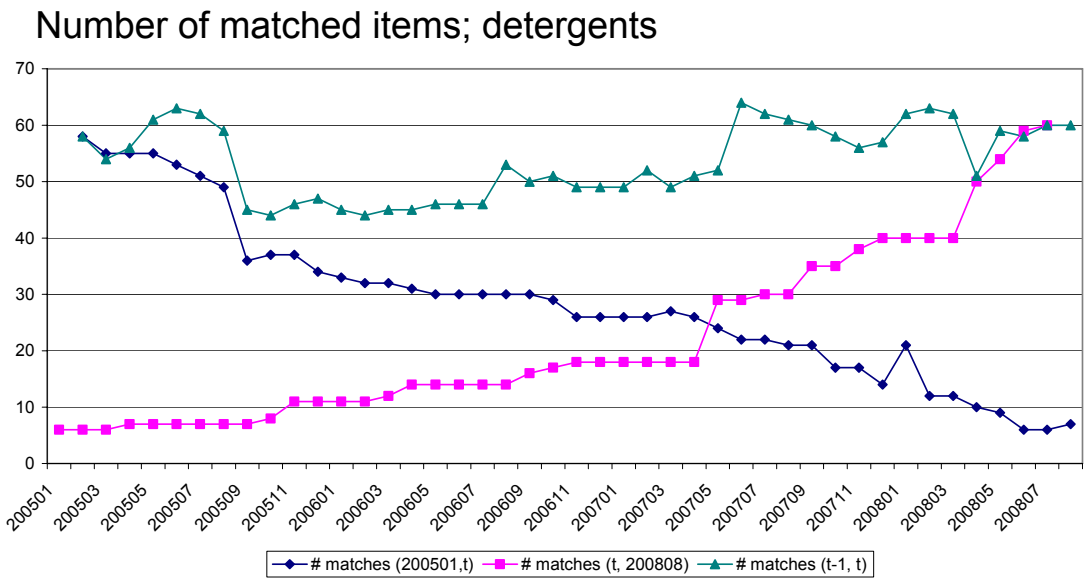
Price and quantity bouncing due to sales

Monthly unit values and quantities; XXX tablets (detergent)



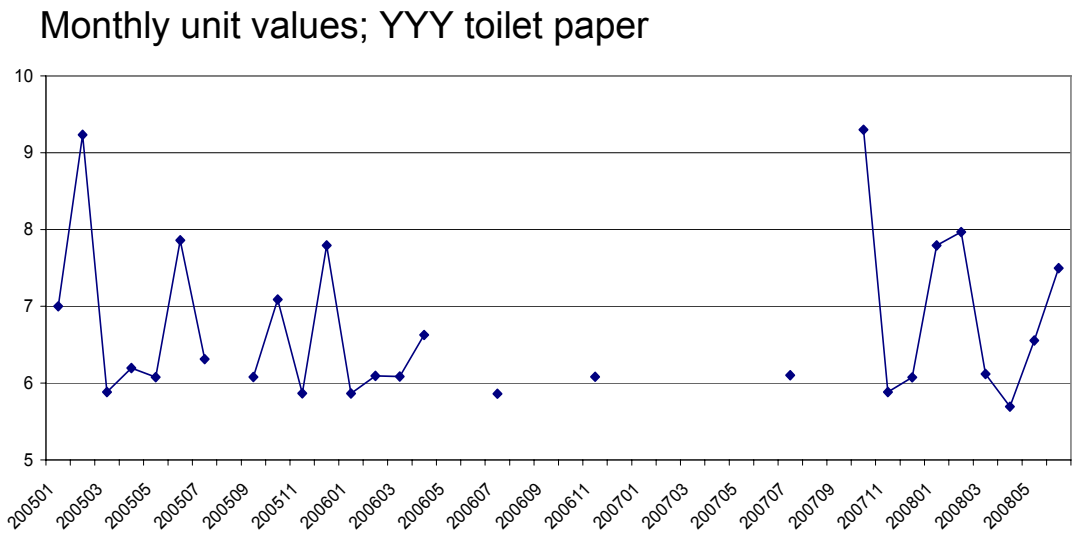
Features of scanner data (3)

High product turnover



Features of scanner data (4)

Temporarily unavailable items



Chain drift

Chain drift occurs if a chained index “does not return to unity when prices in the current period return to their levels in the base period” (CPI manual).

Standard micro-economic theory: given a set of prices, quantities are uniquely determined.

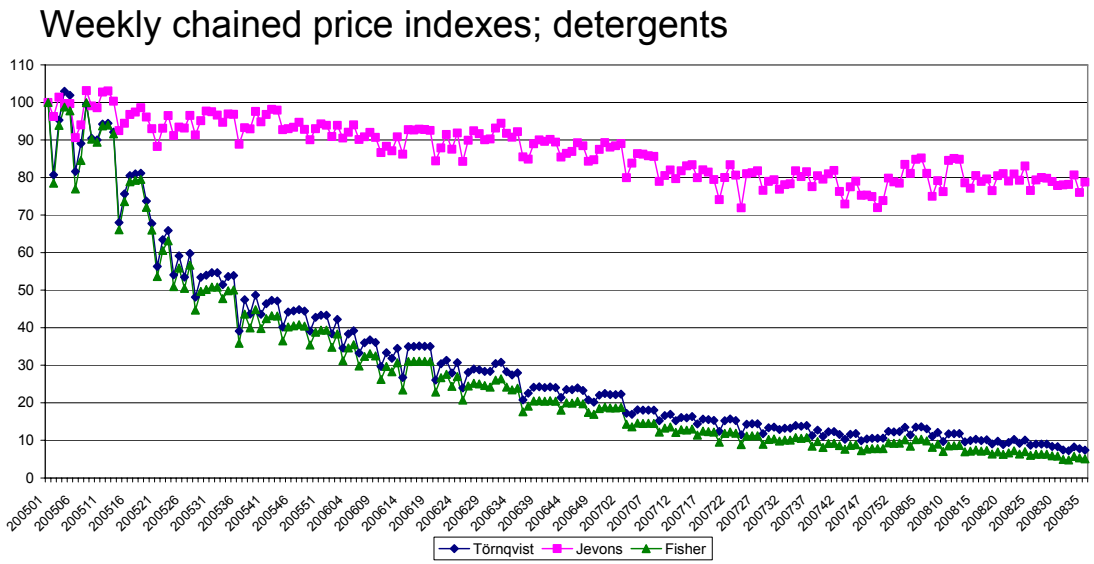
If prices bounce the quantities (hence, expenditures) should also return to their initial levels and superlative price indexes should be drift free.

In practice: ‘distortions’, e.g. due to consumers stocking up during sale periods.

Possible solutions:

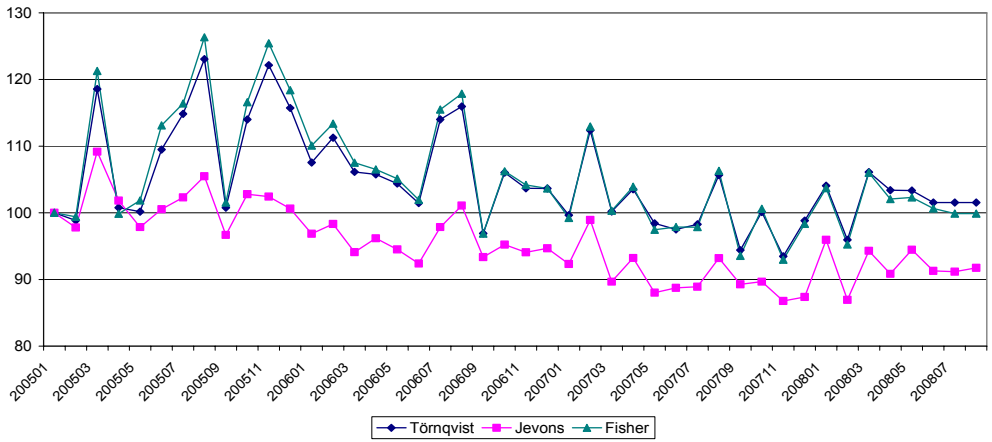
- no chaining (lack of matching over time)
- leaving out goods that are on sale (Stats. Norway)
- aggregation over time (reducing ‘noise’)
- other?

Chain drift (2)

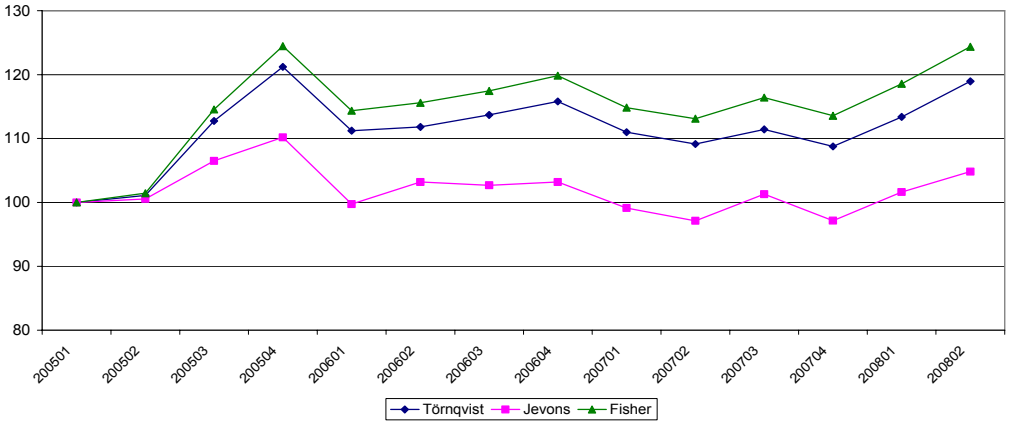


Chain drift (3)

Monthly chained price indexes; detergents



Quarterly chained price indexes; detergents



(Rolling year) GEKS: IFD method

GEKS index: geometric mean of the ratios of all bilateral indexes, computed with the same index number formula, between a number of entities (e.g. countries), where each entity is taken as the base.

GEKS index is transitive: same result is obtained if entities are compared with each other directly or via their relationships with other entities.

IFD: treat each time period as an entity. GEKS for current period *T* (with respect to base period 0):

$$P_{GEKS}^{0T} = \prod_{t=0}^{T} \left[P^{0t} / P^{Tt} \right]^{1/(T+1)} = \prod_{t=0}^{T} \left[P^{0t} \times P^{tT} \right]^{1/(T+1)}$$

P is (superlative) index which satisfies time reversal test.

Note: the direct (bilateral) index may be less representative than the indirect indexes but 'counts twice (for t=0,T)

(Rolling year) GEKS: IFD method (2)

GEKS index makes maximum use of all possible matches in the data between any two periods and should be free of (chain) drift.

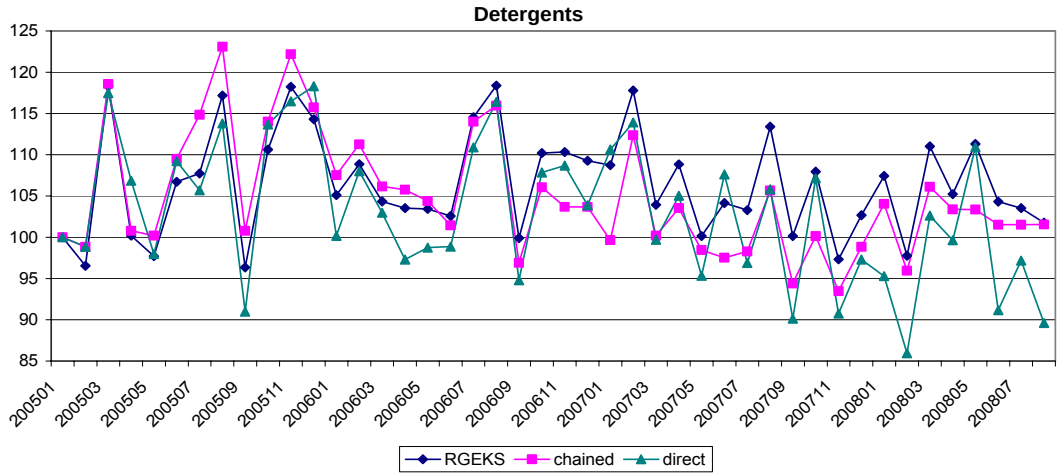
Disadvantages:

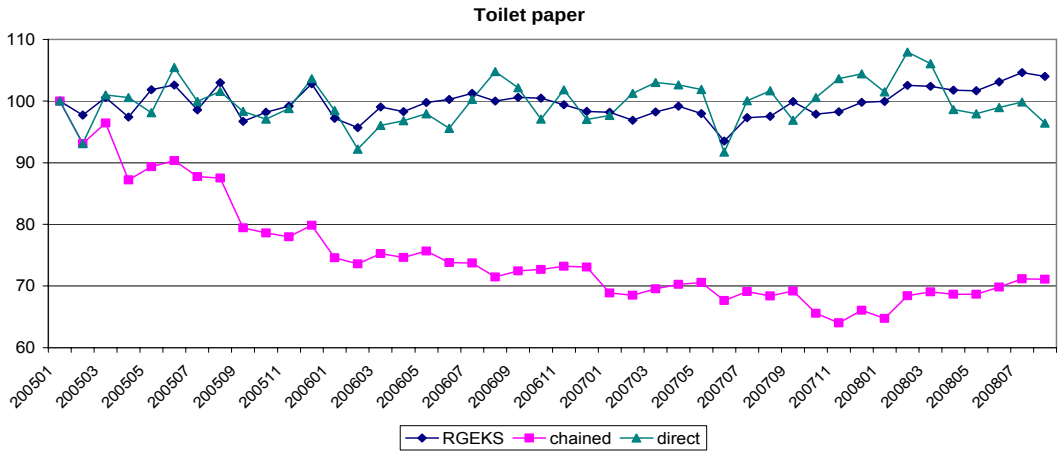
- lack of transparency
- loss of characteristicity (the period *T* index depends on the data of all earlier periods)
- revision: when data is added (*T*+1, ….), index number for period *T* usually changes

Rolling year GEKS: price and quantity data for the last 13 months; the most recent month-to-month change is chain linked to existing time series.

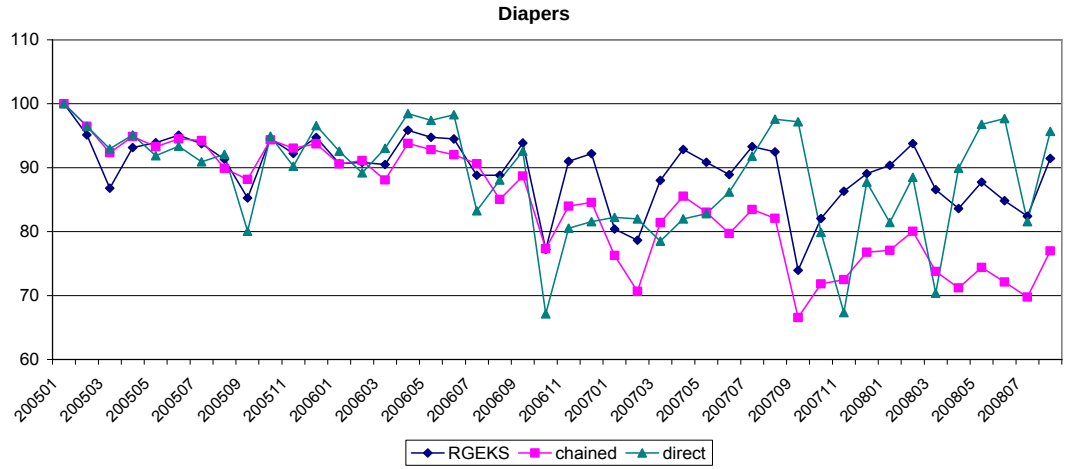
- no revision
- 13 month moving window is shortest one that allows a comparison of strongly seasonal goods

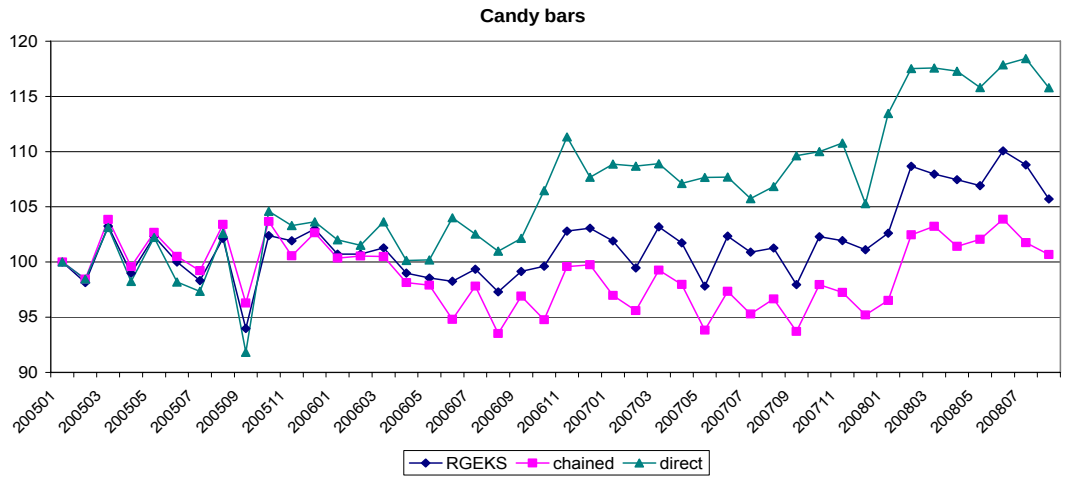
Monthly rolling year GEKS-Törnqvist, chained Törnqvist and direct Törnqvist price indexes



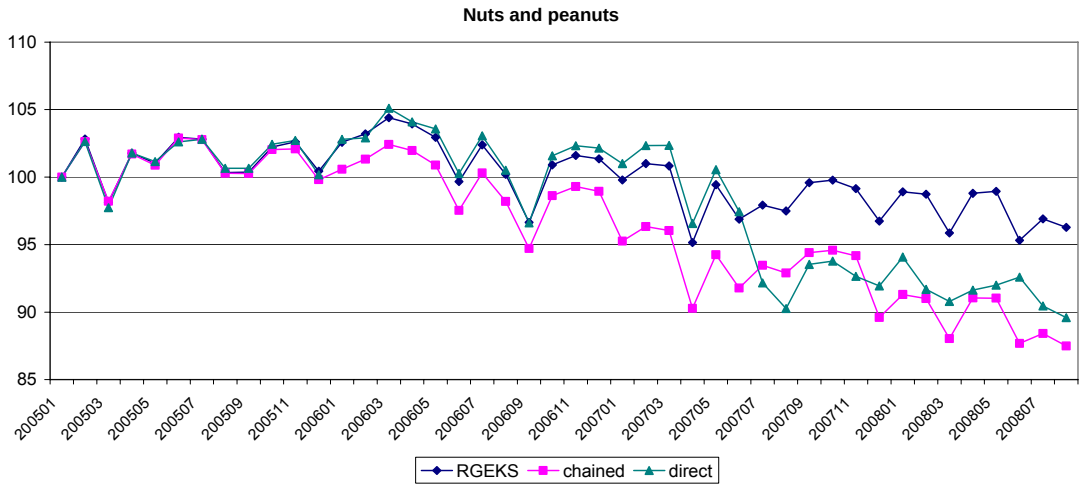


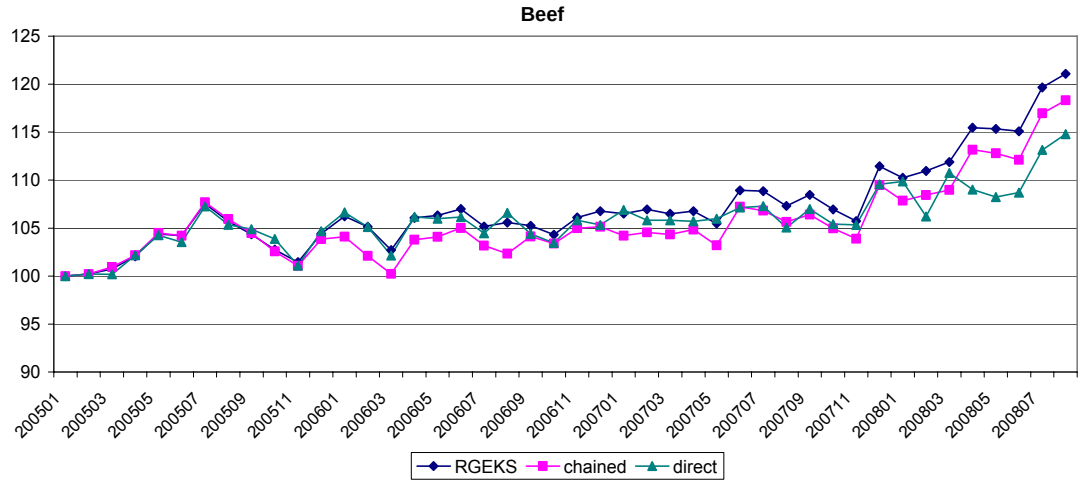
Monthly rolling year GEKS-Törnqvist, chained Törnqvist and direct Törnqvist price indexes



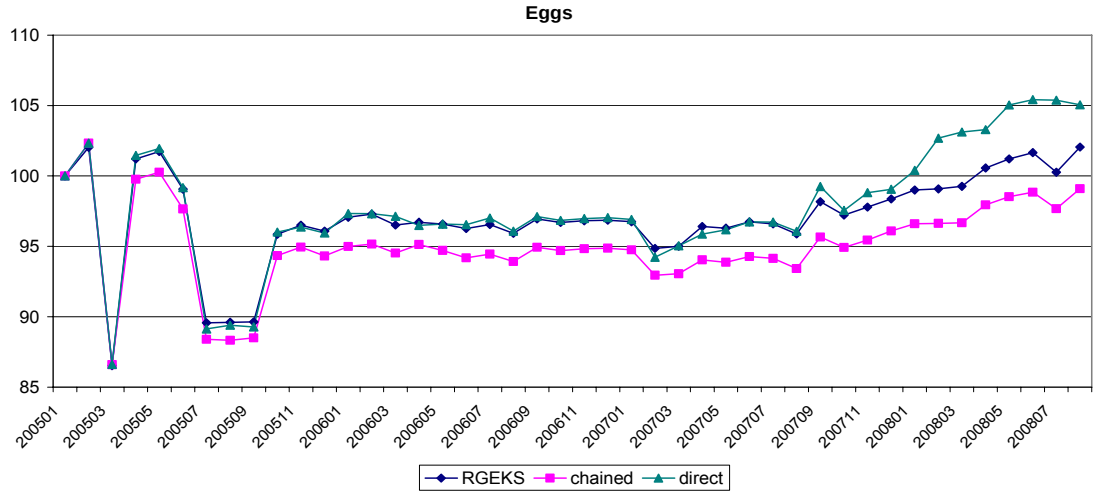


Monthly rolling year GEKS-Törnqvist, chained Törnqvist and direct Törnqvist price indexes

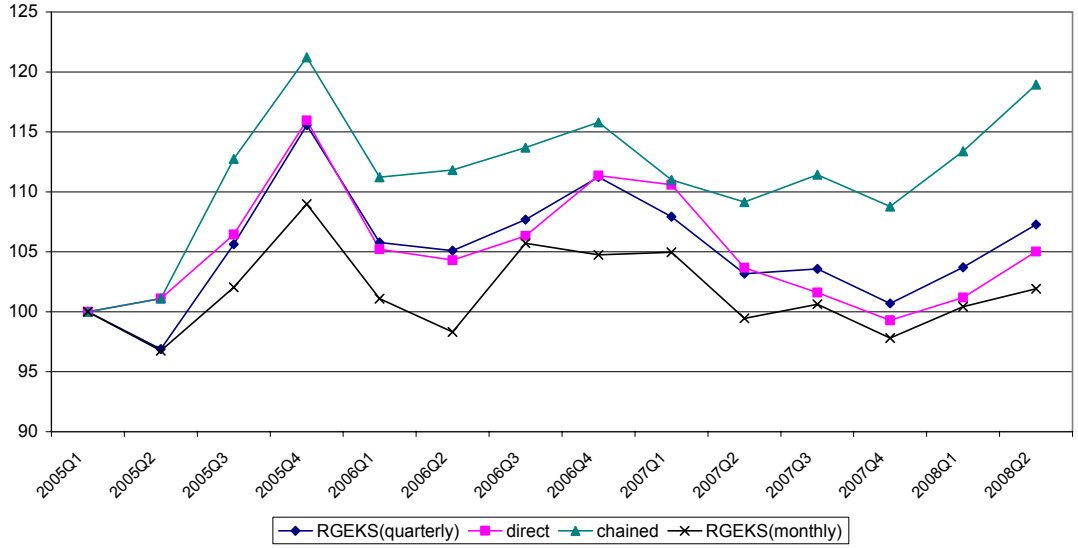




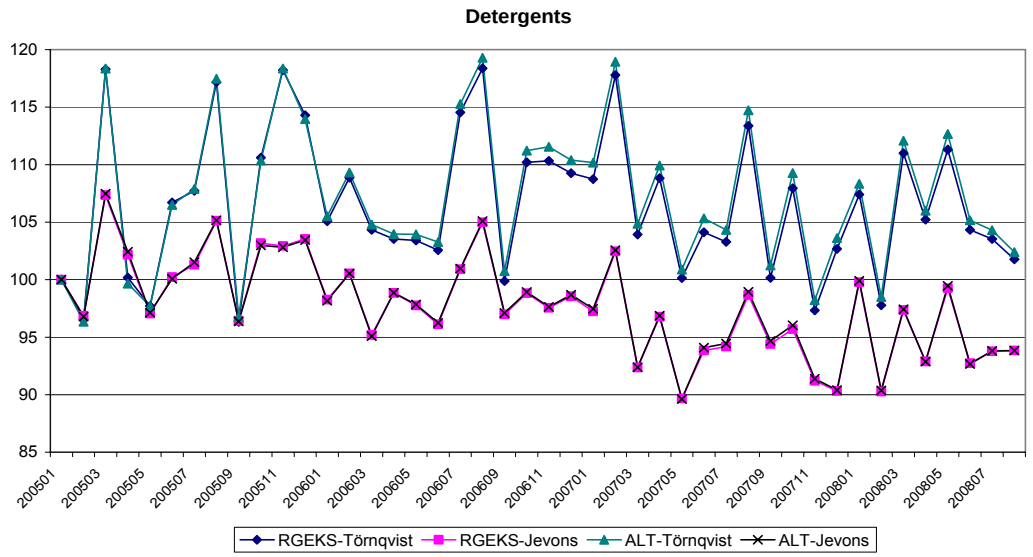
Monthly rolling year GEKS-Törnqvist, chained Törnqvist and direct Törnqvist price indexes

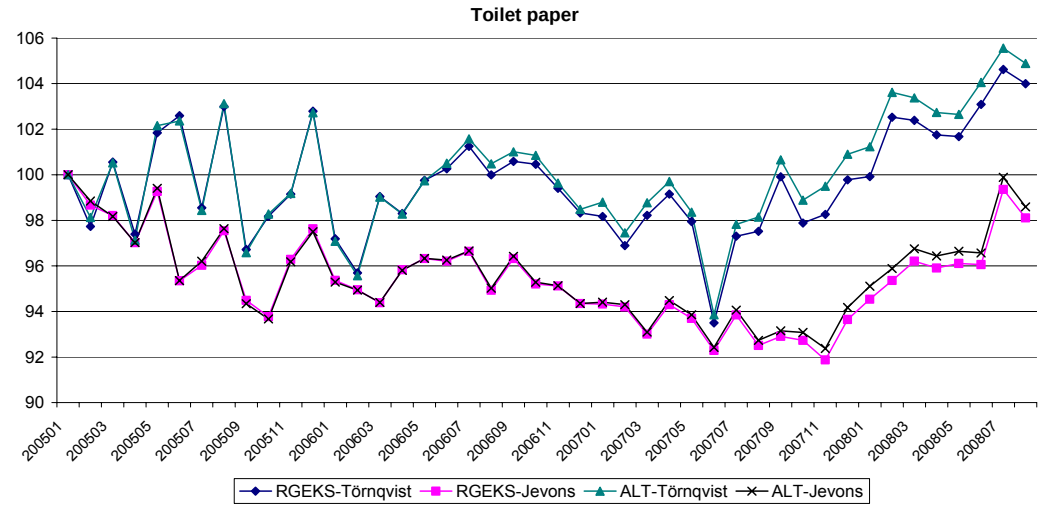


Monthly and quarterly rolling year GEKS-Törnqvist, quarterly chained Törnqvist and quarterly direct Törnqvist price indexes; detergents

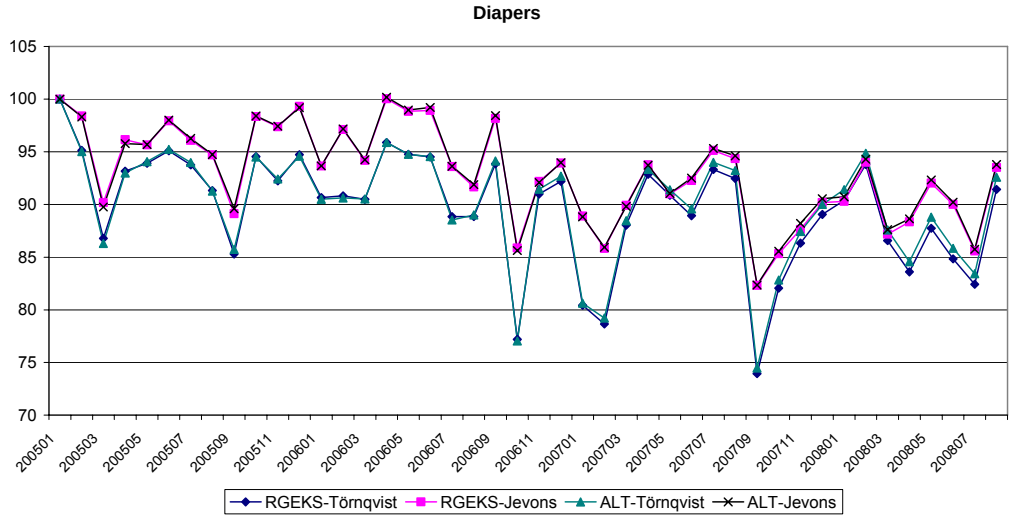


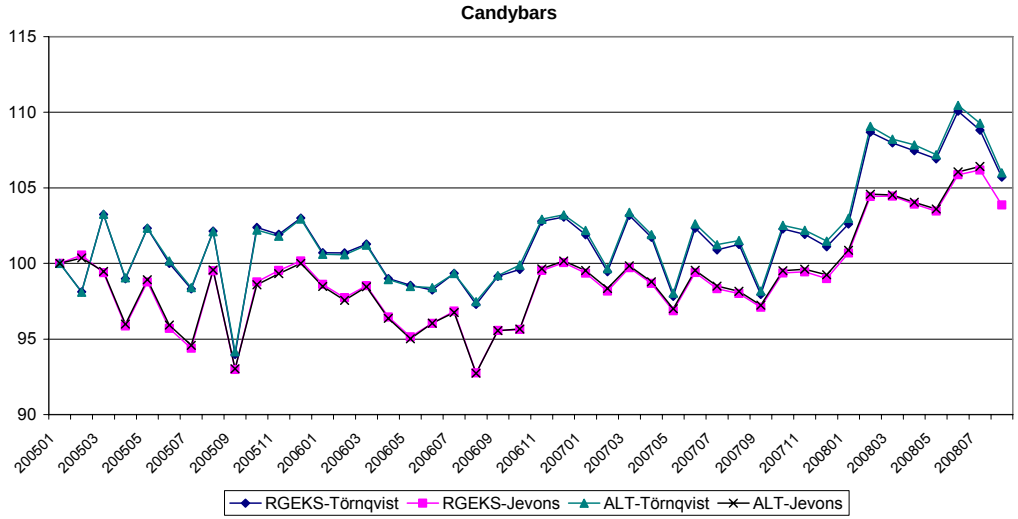
Rolling year GEKS-Törnqvist and GEKS-Jevons indexes



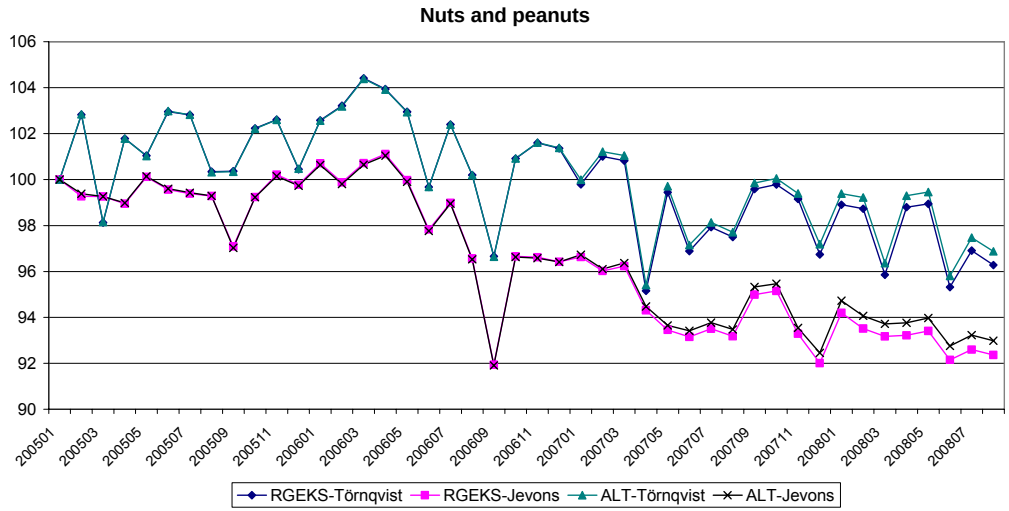


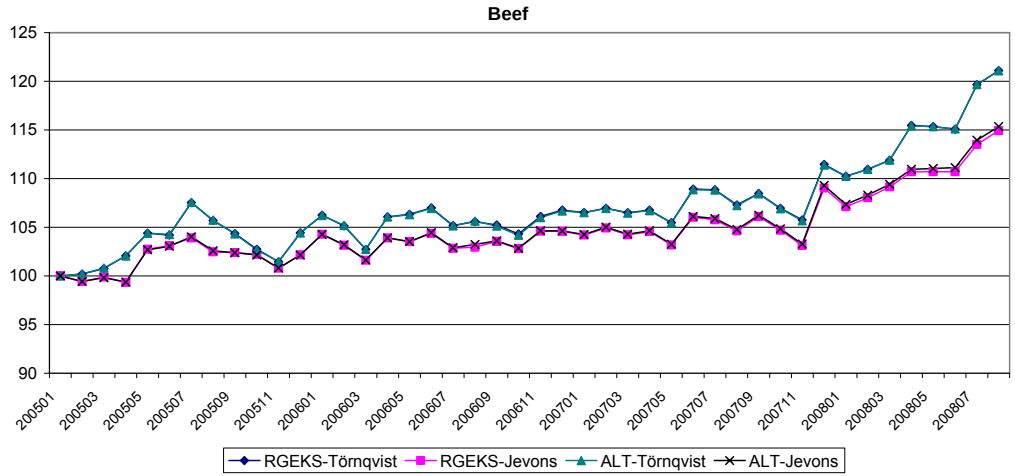
Rolling year GEKS-Törnqvist and GEKS-Jevons indexes



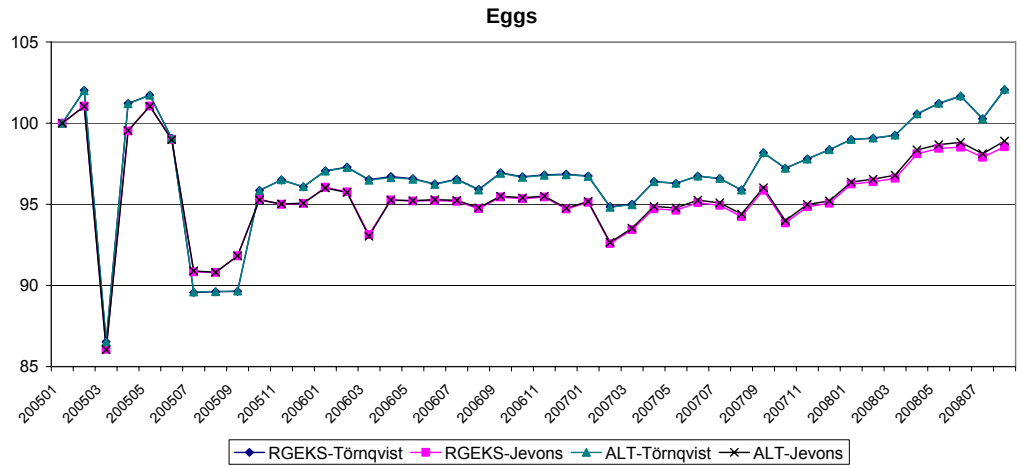


Rolling year GEKS-Törnqvist and GEKS-Jevons indexes





Rolling year GEKS-Törnqvist and GEKS-Jevons indexes



Chained Jevons: Dutch method

July 2009: scanner data for a large number of supermarket chains in CPI using a monthly chained Jevons index

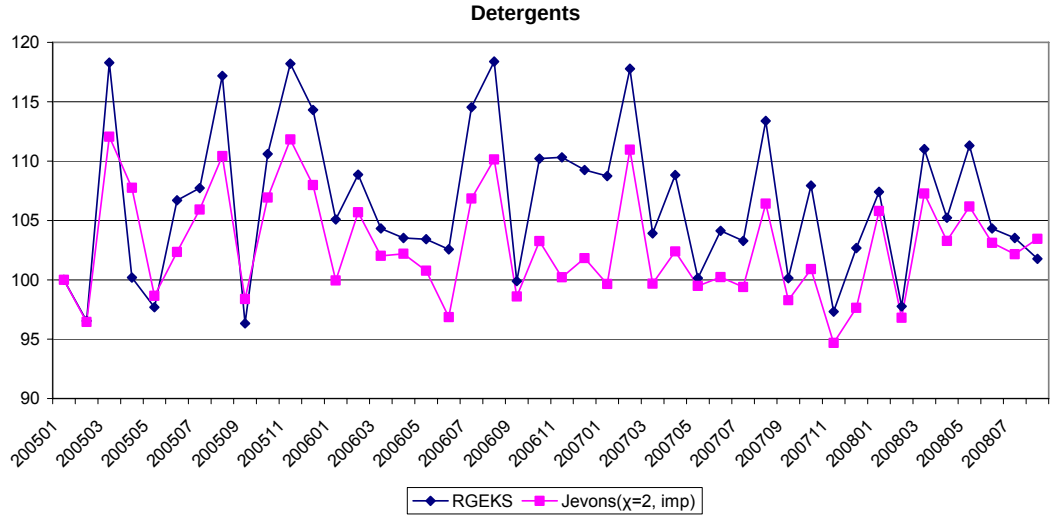
Disadvantages:

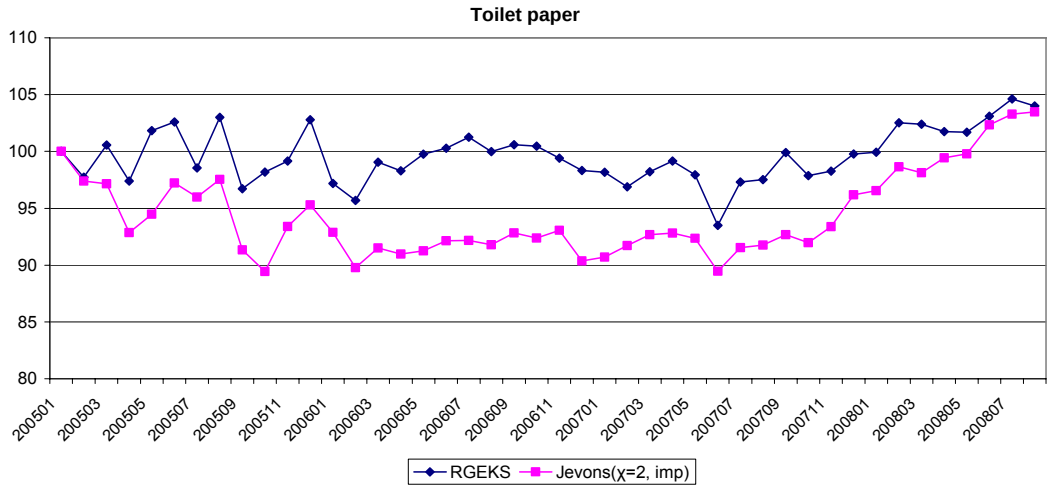
- unweighted index; relatively unimportant items would have the same impact as more important items
- temporarily missing items would be excluded from the computation
- no explicit quality adjustments

'Solutions':

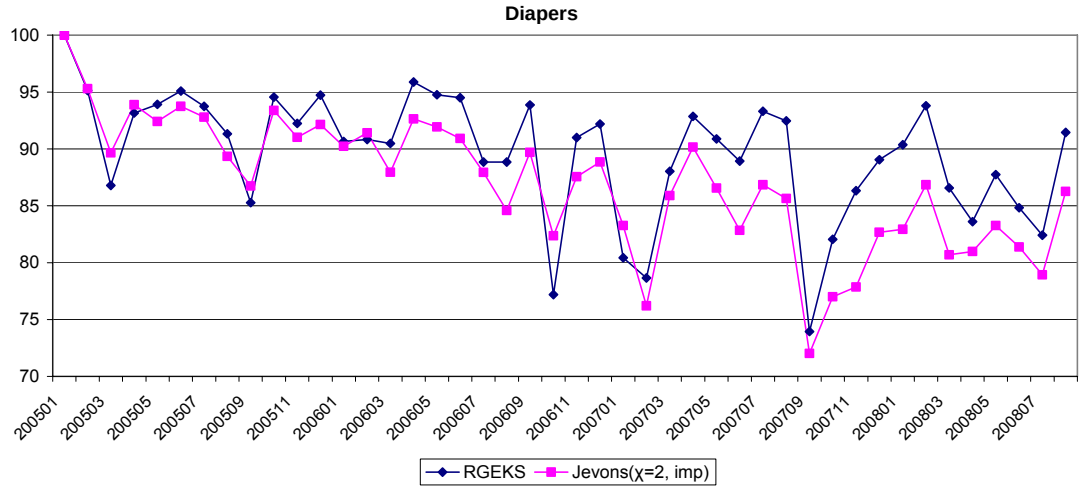
- implicit weighting through cut-off sampling
- imputation of 'missing prices', as usual
- computer system allows for making explicit adjustments (e.g. for changes in package size)

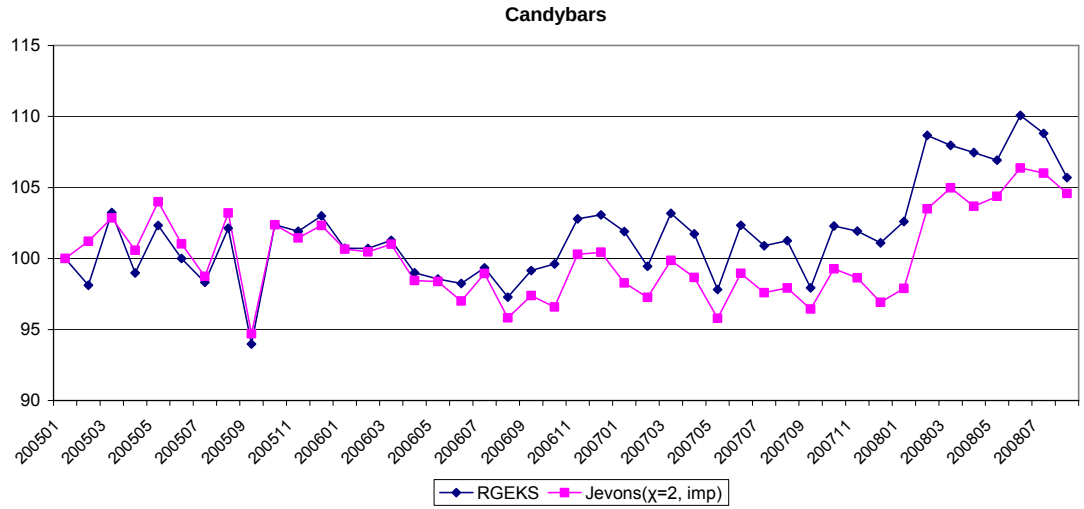
Rolling year GEKS-Törnqvist indexes and chained Jevons indexes (imputations and cut-off sampling)



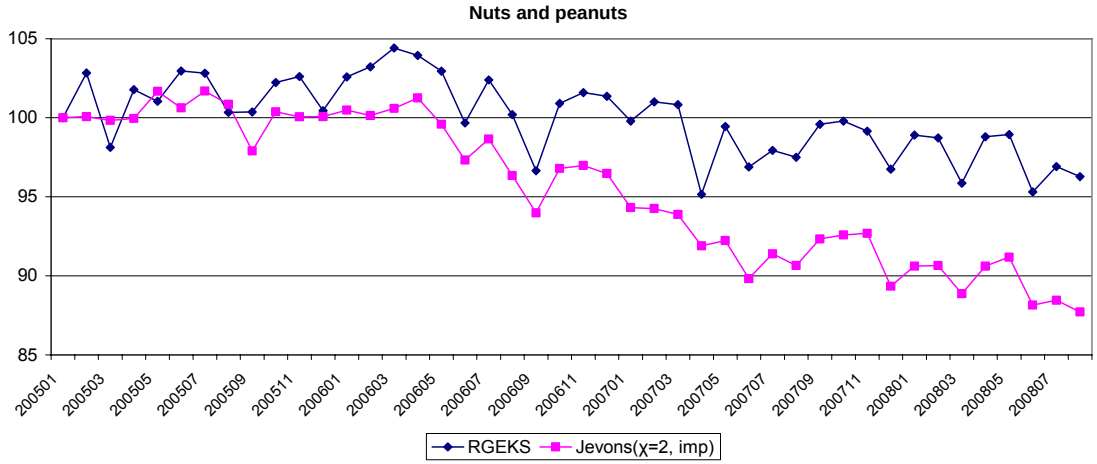


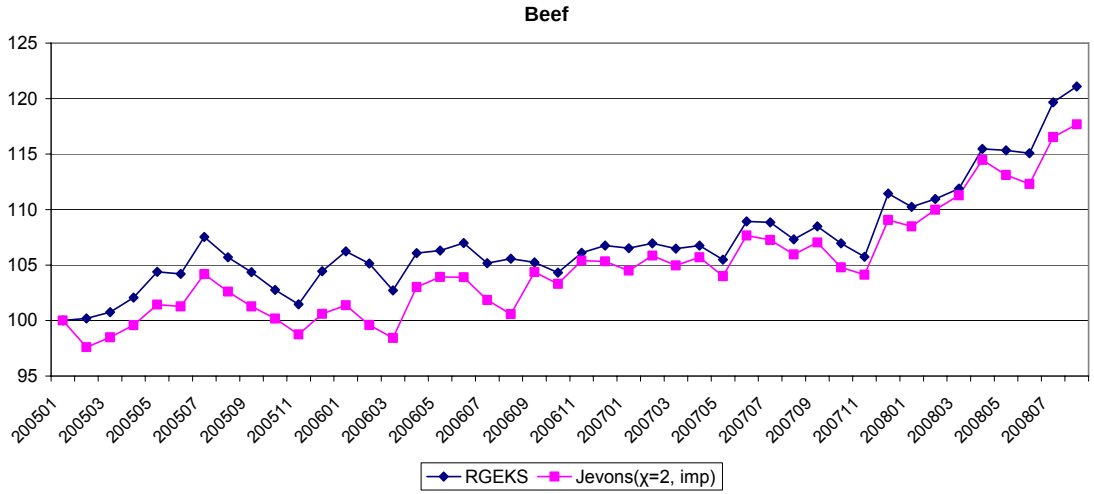
Rolling year GEKS-Törnqvist indexes and chained Jevons indexes (imputations and cut-off sampling)



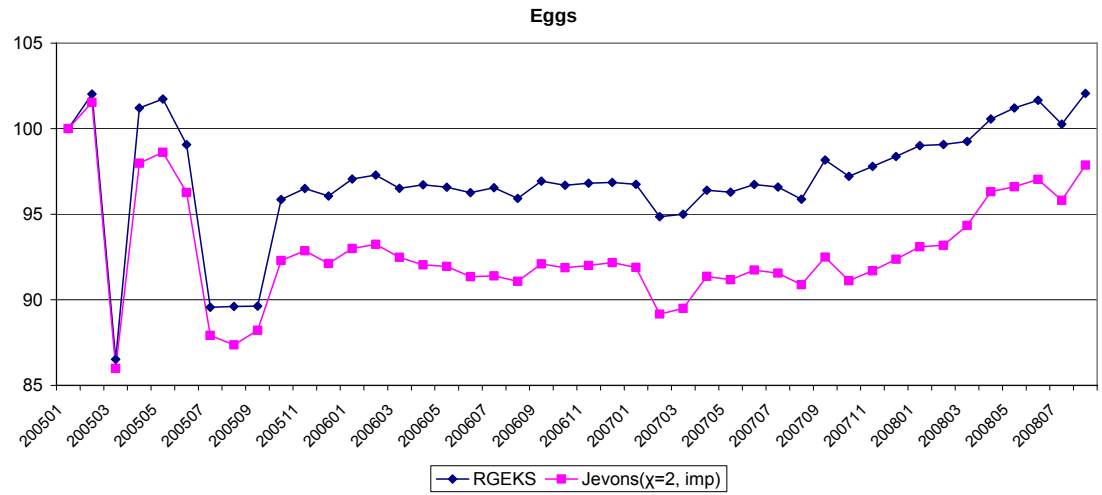


Rolling year GEKS-Törnqvist indexes and chained Jevons indexes (imputations and cut-off sampling)





Rolling year GEKS-Törnqvist indexes and chained Jevons indexes (imputations and cut-off sampling)



Conclusion and future work

IFD method works as expected; no sign of (chain) drift

Future work: computing rolling year GEKS indexes

- for strongly seasonal items
- at higher aggregation levels
- for weekly and monthly data

Plus: paper on second IFD method - ‘Time-Product-Dummy Approach’ (kind of pseudo hedonic imputation approach)