Discussion Paper

Index methodology for transaction data in the Dutch CPI

Antonio G. Chessa

February 2026

Abstract

This report documents the development of methods for measuring inflation using transaction data at Statistics Netherlands since the late 1990s. These historical developments provide essential context for understanding the methodological choices underlying the index methods currently applied in the Dutch CPI and HICP.

Central to the current approach is the use of the multilateral Geary–Khamis (GK) method to compute transitive, drift-free index series on rolling windows of typically 25 months. The GK framework yields item-specific reference prices that measure real value and are robust to the diverse price dynamics characteristic of transaction data, including high item churn, clearance prices and seasonal patterns.

To produce non-revisable index series suitable for the publication of official inflation figures, index series computed on successive time windows are linked using the HYCS extension method, and in some cases with the earlier FBEW method. While transitivity is necessarily lost when extending index series to future months as new data become available, empirical studies have shown that the combined GK–HYCS approach effectively suppresses chain drift. This major finding is contextually grounded with analytical arguments.

Overall, the index method currently in use represents a fully data-driven and dynamically refined implementation of the Lowe index method introduced in 2002, reflecting more than two decades of methodological development in response to the challenges posed by high-frequency transaction data.

Keywords: Consumer price index, inflation, multilateral methods, Geary–Khamis, index series extension, HYCS, FBEW, scanner data.

Contents

1 Introduction 4

2 Terminology, taxonomy of methods and scope 6

2.1 Terminology and notation 6

2.2 Taxonomy of index methods and scope. 7

3 Transaction data over the years 9

3.1 How it started 9

3.2 The years 2002-2009 11

3.3 The years 2010-2013 13

3.4 From 2014 until present 15

4 The Geary-Khamis method 19

4.1 Multilateral methods 19

4.2 Why GK?. 20

4.3 Index formulas 22

4.4 How quality fits in 23

4.5 Index calculation 25

5 Index extension methods 26

5.1 Non-revisability and drift 26

5.2 The FBEW method 28

5.3 The method HYCS. 30

5.4 Other valid methods 34

6 Final remarks 35

References 36

Appendices 40

A Convergence of the GK fixed-point algorithm 40

1 Introduction

The Consumer Price Index (CPI) and the Harmonised Index of Consumer Prices (HICP) compiled by EU member states are considered as primary indicators of inflation. Most National Statistical Institutes (NSIs) publish inflation figures on a monthly basis (US, UK, Eurozone), while some NSIs publish on a quarterly basis (Australia recently switched to a monthly CPI, New Zealand will follow early 2027). Publishing inflation figures on such a frequent basis means that consumer price departments of NSIs have to follow a tight schedule that covers different stages: from collecting prices of different consumer goods and services and checking the data to compiling and analysing price indices at different levels of product aggregation before publishing these indices.

Consumer goods and services in the CPI and HICP are classified according to the universally adopted COICOP system. Prices are traditionally collected in shops for samples or 'baskets' of items that are established by NSIs for different product aggregates (COICOPs). These baskets are annually updated in order to capture trends in product evolution and consumer behaviour.

Price collection in physical shops became problematic during the COVID-19 pandemic because of the lockdowns imposed by governments. As a result, many NSIs exploited the knowledge and experience that was being built up during the previous decade with scraping prices and relevant product information from websites (Eurostat 2020). Webscraping has been one of the major recent innovations in price collection within the context of the CPI and HICP.

Another key innovation goes further back in time, which finds its roots in the early barcode designs in the 1940s-50s. The modern, widely adopted system of what now is officially known as Global Trade Item Number (GTIN) has its origins in the early 1970s, with industry leaders agreeing on the UPC barcode in 1973, leading to the first commercial scan in Marsh supermarkets in the US in June 1974. This concept of assigning unique identifiers to products soon expanded internationally, and eventually took shape as the GTIN system, driven by the global harmonisation efforts of company GS1 in the 2000s (GS1 2023).

GTINs are widely adopted by businesses and are assigned to a wide variety of products sold by supermarkets, do-it-yourself (DIY) stores, clothing and footwear shops, pharmacy stores, and so on. Expenditures, numbers of items sold, along with other information, are electronically stored by shops when items are scanned. The automated storage of transaction data provides an interesting opportunity for NSIs to contact retail chains about the possibility of acquiring these data sets as an alternative to the time consuming traditional price collection.

The first contacts between Statistics Netherlands and supermarket chains were established in the 1990s. The Dutch statistical institute became a pioneer in the acquisition of transaction data and the exploration of its possibilities for CPI compilation. The acquisition and use of transaction data gradually expanded in terms of retail chains and product types in the next decades, which in terms of product weights constituted about 35 per cent of the Dutch CPI in 2024 and close to 40 per cent in the Dutch HICP (mainly because owner occupied housing is not included in the HICP).

Transaction data have clear advantages over traditional price collection:

- Manual price collection is replaced by automated data collection, as retail chains periodically transfer their data to NSIs.
- Transaction data sets contain data of all items purchased by consumers, which means a huge upscaling of 'baskets' compared with the small traditional samples.
- Transaction data contain both expenditures and numbers of items sold, while traditional forms of price collection only collect prices.

The third point also holds when comparing transaction data to other forms of price collection like webscraping. The availability of both expenditures and numbers of sold items for every GTIN makes transaction data a unique source of data for CPI purposes. These data are extremely valuable for price index compilation, since the availability of expenditures allows for the possibility to assign weights to products in an index formula based on their expenditure shares. Price changes of products with high expenditure shares will then have a big impact on aggregate price change.

Transaction data also pose new challenges to NSIs:

- A key principle in price index compilation is that products should be defined in a way that like-with-like comparisons can be made between two periods, thus excluding quality changes. With transaction data, NSIs are served such products on a silver platter, with the very tight characterisations at the GTIN level. However, barcodes may change, also when quality remains unaffected (GS1 2023). Old and new barcodes of the same products have to be linked in order to capture possible price changes enforced by retailers with such product relaunches.
- Transaction data reveal the true dynamics in the evolution of product assortments and in prices and sold quantities over time. This is an unusual setting for traditional price index compilation, which relies on fixed baskets with prices being followed from period to period for each item. How can the typical forms of dynamics in transaction data be dealt with?
- Processing large data sets also sets requirements for the IT system that is used for CPI compilation. A main challenge is how to classify thousands of GTINs according to the COICOP system in an efficient way.

The first and third point are beyond the scope of this document, although some words will be spent about how these challenges have been dealt with. The main focus is on the second point. Statistics Netherlands has built up many years of experience with transaction data, as the first introduction of this kind of data in its CPI and HICP goes back to 2002. The reader is guided through the historical developments and choices that were made in various implementations of transaction data in the Dutch CPI and HICP over time (Section 3), after introducing terminology used in this document and a taxonomy of index methods and choices that have to be made when developing a method (Section 2).

The historical developments provide some context in order to better understand the considerations that have led to the adoption of the more technical Geary–Khamis (GK) index method, which is currently used in the Dutch CPI and HICP. The GK method is described in Section 4. This is a so-called *multilateral index method*, which computes indices simultaneously for different periods in a time window.

Adjusting the window when new data become available will normally change index values in past periods. However, it is not allowed to modify published index values since the CPI cannot be revised. Here we enter the specific problem of how to extend CPI series under the non-revisability condition. The extension methods currently in use at Statistics Netherlands are described in Section 5. Section 6 concludes with some final remarks.

2 Terminology, taxonomy of methods and scope

2.1 Terminology and notation

EU member states compile their national CPI and a harmonised version, the HICP. While recognising the differences between the two systems for inflation measurement, reference will only be made to the CPI in the remainder of this document for convenience. The index methods that are described are the same for the CPI and the HICP. So, where reference is made to the CPI in this text, the assertions are equally valid for the HICP.

The CPI is a monthly statistic in most countries. When describing index methods, months will therefore be used as time units. The generic term 'period' is also used in more general situations.

A *price index* measures relative change in prices of a set of goods or services between two periods. An index is a dimensionless ratio number. Values greater (smaller) than 1 mean that prices have increased (decreased) 'on average', while a value equal to 1 means no change in price level. Index numbers are often multiplied by 100 by NSIs, which is also the value in the reference period.

There is no universally agreed notation for price indices. This document uses the notation $P_{0,t}$ for the price index in month t compared with month 0. The superscript position is left free for additional symbols in more complex settings.

According to the proposed notation, $P_{t-12,t}$ denotes the price index in month t compared with 12 months ago. In other words, it is the year-on-year index in month t, which is the commonly used measure for inflation. A synonymous term for inflation is *12-month rate of change*, which in publications is usually expressed as a per cent change that can be calculated as follows: $100(P_{t-12,t} - 1)$.

In index number theory, it is common to use the notation p_i^t and q_i^t for product prices and quantities, with product i in the subscript and time t in the superscript. The author uses the different notation $p_{i,t}$ and $q_{i,t}$, with both product and time placed in a distinctive way in the subscript.

Electronic data sets containing expenditures and quantities of items purchased by consumers are often called *scanner data* in the literature. This document uses the broader term *transaction data*. Scanner data typically refer to data sets in which expenditures and quantities are specified by GTIN or a retailer's own item codes. However, such codes are not available for all types of goods and services.

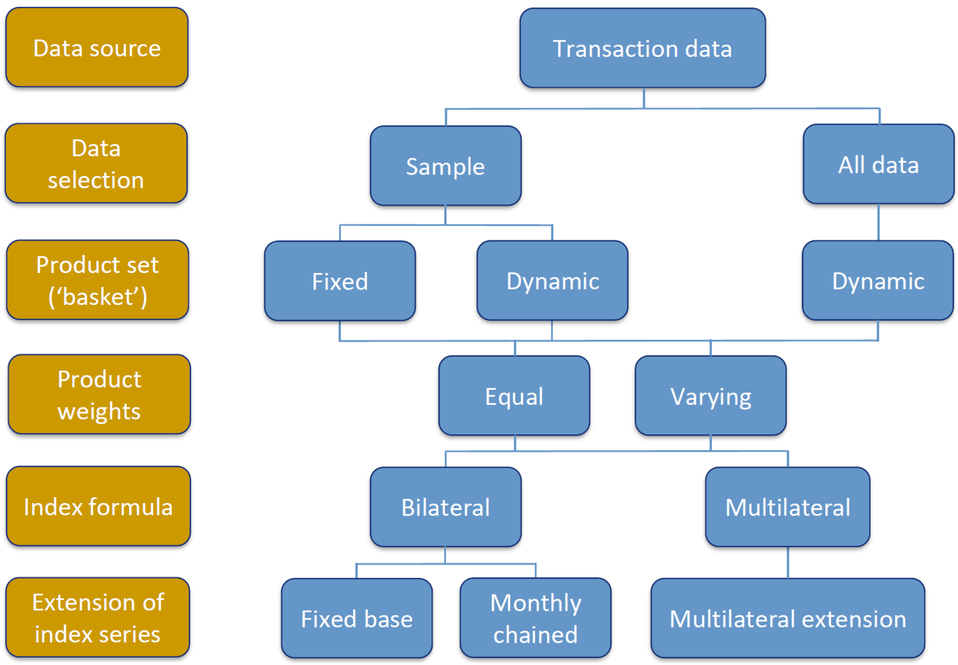
The terms *item* and *product* are both used in this document. The term 'item' is equivalent here with the GTIN level of product definition, and is chosen in accordance with the meaning of GTIN (Global Trade Item Number). The term 'product' is used as a generic concept. It encompasses any type of stratification of items into tight or broad strata, which also includes the item or GTIN level as a special case.

2.2 Taxonomy of index methods and scope

Different choices have to be made when developing an index method for the CPI. These choices do not only apply to the analytical form of an index formula, but also how to define the products and how to treat the data that enter an index formula.

Sections 3–5 give descriptions of different index methods that have been developed at Statistics Netherlands for transaction data. Given the variety of decision variables and the range of possible choices for each variable, it may be useful for the reader to have an overview of the factors that are taken into consideration when developing and implementing an index method. Such an overview helps to characterise and distinguish index methods, and to establish a direction in which index methods evolve over time. Figure 2.1 gives an overview of these factors and the possible choices for each factor.

Figure 2.1 Decision variables (brown boxes) and choices (blue boxes) leading to different classes of index methods for transaction data.



A combination of choices for each factor defines a particular class of index method. Although transaction data sets contain expenditures and sold quantities of all items purchased by consumers, NSIs may want to draw samples of items from these data sets for compiling their CPI. The transition from the relatively small product samples in traditional price collection to thousands of items in transaction data is a big change for NSIs that are building up their first experiences with large electronic data sets.

A set of items that will be followed in a year is typically decided at the end of the previous year in traditional price collection, and is kept fixed throughout a year. This choice can also be made with transaction data samples; an example of an index method using a fixed basket of items selected from transaction data is given in Section 3.2. Samples can also be allowed to vary from month to month in order to include new items in the index calculations; an example of such a method is presented in Section 3.3. The decision to select all transaction data automatically implies that the full dynamic set of items will be processed.

The data for fixed or dynamic sets of items can be used to calculate indices with either bilateral or multilateral index formulas. Within these two broad classes of index formulas it can be decided to use a formula that assigns the same weight to each item or weights that vary across items. Statistics Netherlands has made different choices with regard to index formula and weights in various implementations of transaction data. The considerations behind these choices will be explained in sections 3–5.

Bilateral index formulas either use only prices or both prices and quantities from two periods. An index formula has to be used in some way to build up an index series over multiple periods. This can be done by linking an index calculated in the current period to an index published in a past period. This linking produces an extension of an index series from the previous period to the current period.

The first period in an index formula refers to a past period and is known as the *reference period*. There are two well-known choices for the reference period when computing bilateral indices in the CPI: a fixed reference month—typically December of the previous year in the CPI—and the previous month. These two choices generate extensions of index series that are respectively known as the *direct* or *fixed-base method* and the *monthly chained method*.

The range of methods for extending index series over time is much broader for multilateral index formulas. Multilateral methods compute indices for multiple periods simultaneously. This implies that the number of linking options increases well beyond the options available for bilateral index formulas. Including all linking possibilities for multilateral index formulas in Figure 2.1 would make the diagram too technical at this stage, resulting in a loss of clarity. The linking and extension possibilities are described in Section 5.1.

The index methods classified according to the scheme in Figure 2.1 are applied to the most detailed product aggregates in the CPI. This can be the most detailed COICOP level—5-digit level or higher if there is no differentiation at lower levels—or the most detailed product aggregates defined by NSIs within the COICOP hierarchy. The index series computed for the most detailed product aggregates are subsequently combined in a Laspeyres-type method in order to compile index series for higher aggregates, as established by EU Regulation 2016/792 (Eurostat 2024, pp. 256–257).

3 Transaction data over the years

3.1 How it started

The origins of the use of transaction data in the Dutch CPI can be traced back to the 1990s. High awareness about the potential of transaction data was already well rooted at Statistics Netherlands at the time, both with regard to quality gains (much larger numbers of priced items, real transaction prices, lower sensitivity to errors, detailed and better product weights) and regarding gains in efficiency (less survey work and less administrative burden for retailers).

First contacts with supermarket chains were established in 1995. The contact persons at the main offices were initially hesitant about a future co-operation, as they felt there could be risks with a possible disclosure of the data. Follow-up contacts eventually led to formal, written agreements with two supermarket chains in 2000, establishing confidential treatment and periodic delivery of transaction data in order to ensure timely monthly compilation and publication of the CPI.

The first data sets that were supplied by the two supermarket chains contained weekly expenditures and numbers of scanned items for a total of about 29,000 barcodes sold at a large number of outlets, which is much higher than the 450 items for which supermarket prices were collected at the time. The transaction data sets also contained other relevant information about each item, so the data came in a format that was already very much in line with today's requirements for transaction data.

Once the first data were received, it became clear that one of the main challenges was how to assign such a large number of GTINs to their respective COICOP aggregates in an efficient way. Statistics Netherlands reached an agreement with Centraal Bureau Levensmiddelenhandel (CBL), the Netherlands Food Retail Association, for using their classification of GTINs. The CBL classification provided a convenient intermediate level for mapping GTINs to COICOPs. Although a part of the mapping had to be done manually, the solution found for the classification problem was felt as an important step towards the implementation of transaction data in the CPI.

A second major challenge was how the first transaction data, with prices and sold quantities for each GTIN, could be exploited by an index formula that uses both types of information. The dynamics of the evolution of product assortments was an additional factor that was taken into consideration, as GTINs may disappear from the market, while other GTINs are introduced as new items. Some preliminary studies with transaction data of market research companies were already carried out at Statistics Netherlands before the first transaction data were received.

The first ideas about the choice of index formula gradually took shape. The sights were set on using a monthly chained *Fisher index*, with GTINs as unique homogeneous products. The Fisher index in month t with respect to month $t-1$ is defined as follows:

$$P_{t-1,t} = \left(P^{\mathsf{L}}_{t-1,t}\, P^{\mathsf{P}}_{t-1,t}\right)^{\frac{1}{2}} \tag{1}$$

$$= \left(\frac{\sum_{i \in G_{t-1,t}} p_{i,t} q_{i,t-1}}{\sum_{i \in G_{t-1,t}} p_{i,t-1} q_{i,t-1}} \, \frac{\sum_{i \in G_{t-1,t}} p_{i,t} q_{i,t}}{\sum_{i \in G_{t-1,t}} p_{i,t-1} q_{i,t}} \right)^{\frac{1}{2}} \tag{2}$$

where $G_{t-1,t}$ denotes a set of 'matched' items sold in both months $t-1$ and t. The indices $P^{\mathsf{L}}_{t-1,t}$ and $P^{\mathsf{P}}_{t-1,t}$ respectively denote the month-on-month *Laspeyres* and *Paasche* indices, which are the first and second ratio within brackets in (2). So the Fisher index is an equally-weighted geometric mean of the Laspeyres and Paasche indices.

In terms of the classification scheme presented in Figure 2.1, the plan was to use a dynamic set of items in a monthly chained bilateral index method. The Fisher index uses product weights that vary across items and also over time. Index formula (2) was chosen for a number of reasons:

- The Fisher index satisfies all 20 so-called "tests" or "axioms" on bilateral price indices, and it is the only one to do so (ILO et al. 2004).
- The Fisher index is also among the "best" formulas according to the economic approach to index number theory (ILO et al. 2004, Ch. 17).
- The month-on-month chaining allows to process all items over time, also new and disappearing items, thus fully accounting for the dynamic characteristics of product assortments in the final results.

The scene was set for an ideal scenario. The availability of both price and quantity data created the perfect stage to put the Fisher "ideal index" into practice for the first time. But once the method was implemented and run, the first results showed heavily drifting index series.

In spite of its excellent axiomatic and satisfactory economic-theoretic properties, the Fisher index formula suffers from *chain drift* (Ivancic et al. 2011). This means that direct measurement of aggregate price change between two periods and indirectly through chaining across intermediate periods does not give the same result. In other words, the Fisher index formula is *not transitive*.

There are two main causes of chain drift:

- Changes in the set of sold items over time.
- Dependence of monthly chained indices on prices and quantities in intermediate periods.

The first point causes changes between fixed-base and monthly chained indices because the two index calculations use different sets of items. The second point arises from the analytical form of an index formula, in which prices and quantities of intermediate periods do not cancel out when calculating chained indices.

The differences between the monthly chained Fisher indices and the indices computed with the traditionally collected prices in the two supermarket chains caused a significant impact in the all-items CPI, although the total weight of the two chains was about 4 per cent. A number of driving factors behind chain drift were identified, which can be explained from the aforementioned two causes:

- Temporary unavailability of seasonal items. Items that reappear after months without sales are treated as new items by monthly chained bilateral index methods. Attempts to mitigate drift by imputing prices in months without sales did not have any effect. High-frequency chaining is problematic under highly volatile prices and quantities as in seasonal items.
- The so-called phenomenon of *price bouncing*, which typically refers to the situation where an index does not return to its original level after discounts, which are characterised by high sales.

In addition, the Fisher index evidences problems with rapid price decreases of items that are about to be removed from the market. This is a typical strategy applied by retailers before introducing new items or when replacing disappearing items with items of the same quality under a new barcode ("relaunches"). The Fisher index is very sensitive to downward drift under these pricing strategies. This is caused by the Laspeyres index, since it combines rapid price decreases of disappearing items with quantities in the reference period, which can still be high for such items. So this type of drift is a consequence of the analytical form of the Fisher index formula.

As a result, the chained Fisher index was never used in the compilation of the Dutch CPI. The first experiences with transaction data-based index calculation motivated additional research, which eventually led to the development and implementation of a different index method. More details about the first experiences with transaction data, before the introduction of these data in the Dutch CPI, can be found in Schut (2001).

3.2 The years 2002-2009

The first transaction data-based index method in the Dutch CPI made use of a subset of GTINs, which were selected for each of the two supermarket chains by following a number of steps:

- Expenditures and numbers of items sold of every GTIN were summed over outlets, as is still done in the currently used methods.
- The selection of GTINs was based on the total expenditure of each GTIN in the previous calendar year, which also served as the reference year in the index calculations.
- GTINs had to be available in at least 48 weeks and also in the last 4 weeks of the reference year.
- A cut-off value of 80 per cent was set on the total expenditure share of the selected GTINs.

The selected items were assigned to COICOPs by making use of the CBL classification referred to in Section 3.1. A number of product types were not included in the selection procedure, like seasonal items. This could be related to the negative experiences with the chained Fisher method for these products. Transaction prices of seasonal items were used to replace the manually collected prices in the existing index method.

Beside the implementation of a semi-automated classification method for GTINs, a second major innovation that accompanied the introduction of transaction data in the Dutch CPI was the use of annually updated weights at retailer-COICOP level for the two supermarket chains, based on the expenditures in the reference year. At the time, the Dutch CPI used a basket of goods and services based on the consumption pattern of households from 1995.[1)]

The selection procedure resulted in 6-7,000 GTINs for the index calculations of each supermarket chain. The index formula for the most detailed product aggregates is described below. First, some additional notation is introduced. Let G^r denote a set of items sold in reference year r, which are selected according to the aforementioned steps. For convenience, reference years are relabelled as integers $r \geq 1$. December of year r can thus be denoted as month $12r$.

The transaction price of item $i \in G^r$ in month t is denoted as $p_{i,t}$. The index method compares item prices in month $t \in [12r + 1, 12(r + 1)]$ of the current year $r + 1$ with average item prices p_i^r of reference year r. These average prices are computed as *unit values*, that is, by dividing total item expenditure by the sum of sold item quantities q_i^r in year r.

The formulas presented in the appendix of the report by Schut (2002) can be rewritten into the following form for the price index in month t with respect to reference year r:

$$P_t^r = \frac{\sum_{i \in G^r} p_{i,t} q_i^r}{\sum_{i \in G^r} p_i^r q_i^r}. \tag{3}$$

The CPI typically builds up index series in a calendar year with December of the previous year as the *price reference period*. The annual index series are subsequently linked to the index compiled in December of the previous year in order to obtain an index series covering multiple years. In order to achieve this, expression (3) has to be normalised according to the value in the price reference period, which is month $12r$. This normalisation yields a *Lowe index* $P_{12r,t}^r$ in month t with *weight reference period* r (ILO et al. 2004, p. 3):

$$P_{12r,t}^r = \frac{\sum_{i \in G^r} p_{i,t} q_i^r}{\sum_{i \in G^r} p_{i,12r} q_i^r}. \tag{4}$$

This formula satisfies $P_{12r,12r}^r = 1$, as desired. Price indices were calculated by making use of the first two full weeks of each month.

According to the scheme in Figure 2.1, the index method can be characterised as a method that uses a sample from the transaction data, with the set of corresponding items kept fixed during a year, which are processed with fixed-base bilateral index formula (4). The product weights are fixed during a year, but vary across items. The decision to keep certain factors fixed evidences the prudent approach followed by Statistics Netherlands

1) The base year was changed to 2000 in January 2003. Annually updated weights according to the expenditure pattern of the previous year were introduced in January 2007.

during the development of the first index method for transaction data in the Dutch CPI after the experiences with the Fisher index.

It is important to note that index formula (4) yields transitive index series in calendar years, so that a solution was found to the previously encountered chain drift in the Fisher index series. However, the transitivity property holds when the set of items G^{τ} remains fixed over time. Product assortments may rapidly change, also in supermarkets. Ongoing research showed high rates of churn in the supermarket data, as about 20 per cent of the items sold in 2000 were no longer available in 2001 (Schut 2002).

Although a substantial part of the disappearing items was found to have low expenditure shares, there was a clear need to keep the basket up to date. A number of measures were defined to achieve this:

- Item replacements were found for GTINs with an expenditure share of at least 2.5 per cent within CBL groups. A number of possible actions were considered when comparing the disappearing and replacement item:
  - whether or not to adjust the price of the replacement item for quality change;
  - to adjust the price of the replacement item if it had a different net content.
- The prices of GTINs with expenditure shares below 2.5 per cent were imputed according to the month-on-month index of the GTINs for which prices are available in the previous and current month within the same CBL group.
- New GTINs that were not used as replacement items were considered for inclusion in next year's basket according to the rules set out at the beginning of this subsection.

Also Lowe index formula (4) is sensitive to downward drift in months with clearance prices, for the same reasons noted for the Laspeyres index at the end of Section 3.1. However, the interventions that were implemented for disappearing items rendered the drift issue negligible.

The method described above was introduced in the Dutch CPI for two supermarket chains in June 2002. Efforts to acquire transaction data sets from other retail chains continued. Several years later, agreements were reached with five supermarket chains for the delivery of transaction data.

While the first index method that was developed for processing transaction data in the Dutch CPI feels carefully crafted, it took considerable time to maintain the method, as the item replacements were handled manually. Extending this work to other supermarket chains was found to be very demanding with the available resources and strict deadlines. These considerations eventually led to the development of a new index method.

3.3 The years 2010-2013

During the development of the new index method, high priority was given to two points:

- It was desirable to develop a method that keeps maintenance work in production limited.
- New items should be included in the index calculations in the month of introduction to the market.

When the second point was addressed, thoughts went back to the monthly chained Fisher index that was originally targeted as the index method to be used in the Dutch CPI. Because of the chain drift evidenced by this method, the index formula eventually chosen was the Jevons index. This formula was used to compute monthly chained index series according to the following expression for the month-on-month index:

$$P_{t-1,t} = \left(\prod_{i \in G_{t-1,t}} \frac{p_{i,t}}{p_{i,t-1}} \right)^{\frac{1}{N_{t-1,t}}}, \tag{5}$$

where $G_{t-1,t}$ denotes a set of matched items in months $t-1$ and t and $N_{t-1,t}$ is the number of items in this set. GTINs were chosen as unique products also in this method.

Jevons index series are transitive when a set of items remains the same over time. This cannot be guaranteed in practice, since the composition of product assortments often changes over time. This implies that chain drift cannot be avoided in monthly chained Jevons index series either when new items are included in the index calculations from the first month in which these items are sold.

An additional drawback of the Jevons formula is that all items are assigned the same weight, irrespective of their expenditure shares. Items with very low expenditure shares, or with rapidly decreasing expenditures and prices when items are about to be removed from the market, may significantly distort aggregate price change, while their influence should be limited.

As a consequence, a number of measures were implemented in order to limit these adverse effects on aggregate price change, which were balanced with the amount of maintenance work in production:

- Prices of items with average expenditure shares over two consecutive months below a specific threshold were excluded.
- Thresholds were also set in order to detect rapid decreases in expenditures and prices of items that were about to be cleared. The detected prices were subsequently removed.
- Items that were about to be cleared were not linked to possible replacement items, which means that relaunches were ignored.

Prices that were excluded after applying these thresholds were imputed according to the month-on-month index of all other items within the same aggregate. Prices were also imputed for temporarily unavailable items, such as seasonal goods. An imputation period of 14 months was established as a general rule in order to allow temporarily unavailable items to re-enter the index calculations when they re-appeared in an assortment, and in order to better account for annual shifts in the availability of certain seasonal items, such as Easter eggs.

A comparative study between the Jevons-based method and the Geary–Khamis method for supermarket data showed that almost 70 per cent of the GTIN prices were excluded after applying the expenditure and clearance price thresholds in the Jevons method, which correspond to an expenditure share of almost 20 per cent (Chessa et al. 2018,

p. 6). This shows that the results of the Jevons method heavily depended on imputed prices. While a monthly chained index method allows to process all items, the limitations of the Jevons index formula for transaction data led to an implementation with a drastically reduced sample from the data.

More details about the Jevons-based method can be found in Grient and Haan (2010) and in the guide for processing transaction data developed by Eurostat, where this method is referred to as the "dynamic approach" (Eurostat 2017). Details about the threshold values can be found in Haan and Grient (2011, p. 45).

The Jevons method was introduced in the Dutch CPI in January 2010. By then, the number of supermarket chains for which transaction data were used in the CPI had increased to six. Meanwhile, the CPI department of Statistics Netherlands was trying to acquire transaction data from additional supermarket chains and also from other retail chains.

In 2013, the number of retail chains with transaction data in the Dutch CPI had increased to 15. Beside supermarkets, transaction data were also used for chains of department stores and DIY stores. The landscape of index methods used to process transaction data was changing as more transaction data sets were used in the CPI compilation process: Jevons was used for supermarkets, a Laspeyres-type method was used for the department stores, and yet another, sample-based method was used for DIY stores.

The equally-weighted Jevons cannot be applied to arbitrary types of product. For instance, iPhones have a big market share and a different pricing strategy compared with smartphones of other manufacturers. As a consequence, the Jevons may give a distorted picture of the aggregate price change of mobile phones. This example explains why the search was extended to index methods beyond the Jevons as the number and variety of retail chains and products expanded.

Also different choices regarding product definition were made under the growing diversity of product types. The department stores sell a broad variety of products, among which clothing. The GTIN level is not suited to these products because of the high rates of churn and the frequent use of high initial prices for new product lines (Chessa 2016). The price increases implied by such strategies would be missed by the Jevons method for supermarkets, since this method ignores relaunches. Broader item strata were defined for clothing in order to capture price changes associated with relaunches.

Given the recent methodological developments, a major goal was set by the end of 2013: to prevent the landscape of index methods in the Dutch CPI from further fragmentation as new transaction data sets were acquired. The question was: *Could an index method be found that can be applied to transaction data of different retail chains selling a broad variety of product types?*

3.4 From 2014 until present

The question at the end of the previous subsection marked the beginning of a new phase in the development of index methods for transaction data. A comparative study set up at Statistics Netherlands included bilateral and multilateral index methods, which were applied to transaction data of a department store chain. These data were chosen because of the broad variety of products sold by the department stores.

Multilateral methods are more technical than bilateral methods, but have clear advantages over the latter class of methods:

- Multilateral methods generate transitive index series on time windows containing multiple periods.
- All items can be processed, including new and disappearing items, without compromising transitivity.
- Items can be weighted according to their period-specific expenditure shares, which does not affect transitivity either.

From this list of benefits it will be clear that the problems and shortcomings indicated in the previous sections with bilateral methods can be dealt with by using multilateral methods. However, the problem of chain drift still needs attention for various reasons, which will be further explained in Section 5.

The research phase initiated at Statistics Netherlands led to the Geary–Khamis (GK) method as the recommended multilateral method for transaction data in the Dutch CPI (Section 4.3). After successful implementation and a series of test calculations in production, the GK method was introduced in the Dutch CPI for transaction data of mobile phones in January 2016. Index series were extended from month to month with the FBEW method, which is described in Section 5.2. A detailed description of the compound GK–FBEW method can be found in Chessa (2016).

Statistics Netherlands became the first country in Europe using a multilateral method in its CPI. A sequence of events followed after presentations of the method at the 2015 Scanner data workshop and the 2016 Meeting of the CPI Group of Experts. A timeline of events is shown in Figure 3.1. The timeline contains different lines of research initiated at Statistics Netherlands since 2014, the multilateral methods introduced in its CPI, initiatives that highlight the strategy of Statistics Netherlands towards modernising its data collection, and the international dimension.

A first meeting organised and hosted by Eurostat in December 2016 on the GK method and other index methods eventually led to the Task Force on Multilateral Methods (TF MLM) launched by Eurostat several years later, which published the first edition of the Guide on Multilateral Methods (Eurostat 2022). The task force constitutes a platform where NSIs and other international organisations can exchange research results and experiences, especially with multilateral methods for CPI purposes. The formation of the TF MLM was a logical response to the extensive research that has been carried out worldwide on index methods for transaction data in a relatively short period of time (ABS 2017; Auer 2017; Białek and Bobel 2019; Białek 2025; Chessa et al. 2017; Dalén 2017; Diewert and Fox 2017; Krsinich 2014; Lamboray 2017; Loon and Roels 2018; Tardos 2023; Zhang et al. 2019).

Other important research developments were initiated on the topics of product definition or stratification and on the problem of extending multilateral index series over time.

Figure 3.1 Timeline of events during the development of new methods for transaction data.

2014-01	Research on index methods at CBS
2015-01	Implementation and tests of chosen method
2016-01	GK–FBEW in Dutch CPI for mobile phones
2016-01	Relations management group formed
2016-12	Meeting at Eurostat on CBS research
2017-01	Research on product stratification at CBS
2018-01	GK–FBEW in CPI for supermarkets, DIY and dept. stores
2018-08	Stratification method MARS finalised
2018-09	Research on extension methods at CBS
2019-11	MARS applied to mobile phones in Dutch CPI
2020-01	End of price collection in physical shops
2020-09	TF MLM launched by Eurostat
2022-04	Guide on Multilateral Methods published
2024-01	GK–HYCS in Dutch CPI for clothing, footwear and furniture
2025-01	GK–HYCS applied to almost all transaction data in Dutch CPI

Various studies on these topics indicate that the range of choices on product stratification and index extension have a bigger impact on aggregate price change than choosing among multilateral methods with expenditure-based product weights (e.g. see Chessa (2016) and Chessa (2021b) and Chessa et al. (2017)).

Research on product stratification carried out at Statistics Netherlands resulted in the method MARS (Chessa 2021a). This method aims at finding and combining product information such that the tightness of the resulting item strata (*product homogeneity*) is balanced in an optimal way with the degree by which item strata can be matched in two periods (*product match*). This allows to capture relaunches and associated price changes and, importantly, to automate the definition of products in the CPI. An illustration of the effectiveness of product information to define homogeneous item strata that are capable of detecting relaunches and changes in package content can be found in a recent study on shrinkflation (Chessa and Barb 2025).

The emphasis in the studies on methods for transaction data was mostly on the multilateral index formulas. How to extend multilateral index series when data of the next month enter the index calculations was initially given less importance. This problem is concerned with linking index series computed on successive time windows. There are many ways of linking, which may result in substantially different behaviour with regard to chain drift (Section 5). Different years of research carried out at Statistics Netherlands on this topic eventually led to the extension method HYCS (Section 5.3), which has now replaced the originally developed FBEW method for almost all transaction data in the Dutch CPI.

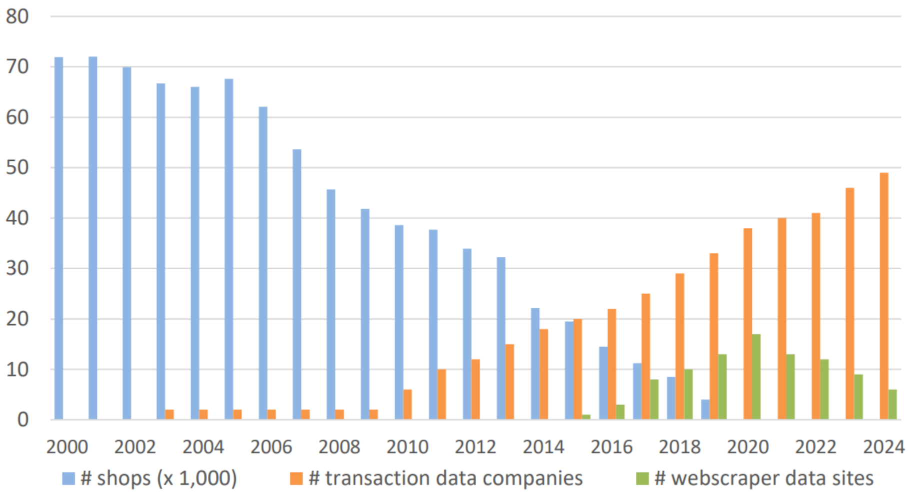
Figure 3.2 gives a summary overview of the index methods that have been implemented for supermarket transaction data in the Dutch CPI since 2002. The methods are characterised according to the scheme in Figure 2.1. The different phases clearly show a transition towards a fully data-driven and dynamic index method, which combines the multilateral GK method with the extension method HYCS that makes use of a rolling time window.

Figure 3.2 Characteristics of the various index methods developed for supermarket transaction data in the Dutch CPI.

Choice	Method 1 (2002)	Method 2 (2010)	Method 3 (2018)	Method 4 (2025)
Data selection	Sample	Sample	All data	All data
Basket	Fixed	Dynamic	Dynamic	Dynamic
Product weights	Varying	Equal	Varying	Varying
Index formula	Bilateral (Lowe)	Bilateral (Jevons)	Multilateral (GK)	Multilateral (GK)
Index extension	Fixed base	Monthly chained	FBEW (fixed base)	HYCS (rolling)

Beside the many research activities, other important initiatives were created that had, and still have, a big impact on the Dutch CPI and deserve to be mentioned. In 2016, a separate unit within the CPI department of Statistics Netherlands was specifically formed to take care of the relations with retail chains. The acquisition of transaction data to replace other forms of price collection is set as the main priority. The number of retail chains delivering transaction data sets that are used in the Dutch CPI kept increasing since 2016, which in 2024 went close to an impressive 50 retail chains (Figure 3.3).

Figure 3.3 Number of retail chains delivering transaction data, scraped websites and visited physical shops in the Dutch CPI during 2000–2024.



In the first years of the 2010s, Statistics Netherlands also started to investigate the possibilities of using web scraped data as an alternative source to replace manually collected prices. This led to the use of web scraped data in the Dutch CPI in January 2015. However, after adding more scrapers to the CPI it became clear that the use of scrapers resulted in high maintenance costs. Scraping code has to be adjusted after website modifications, which is particularly costly for the front-end scraping technique adopted at Statistics Netherlands. Several years ago it was therefore decided to stop building new scrapers and to replace existing scrapers with transaction data when possible.

The increase in the number of retail chains that provide transaction data, and initially also in the number of scrapers, significantly contributed to reducing the number of physical shops visited by price collectors. A major goal was reached in January 2020: price collection in physical shops came to an end, less than three months before the COVID-19 pandemic was declared (Figure 3.3).

4 The Geary-Khamis method

4.1 Multilateral methods

The experiences described in sections 3.1–3.3 evidence a dilemma when applying bilateral index methods to transaction data:

- Chained index methods that make use of both item prices and quantities are able to exploit various forms of dynamics in transaction data sets, but are sensitive to chain drift.
- Chain drift can be better controlled, or even eliminated, by avoiding certain choices in an index formula, such as the use of monthly expenditure shares as product weights or the use of a dynamic basket. Concessions are then made in terms of the dynamics of product assortments and consumer behaviour.

These points summarise the limitations when applying bilateral methods to transaction data. At the beginning of Section 3.4 it was argued that these limitations can be dealt with by using multilateral methods.

Multilateral methods were originally developed for international price comparisons. An important feature of these methods is that the resulting purchasing power parities are independent of the choice of reference country, and consequently generate transitive indices. Well-known examples of multilateral methods are the GEKS method (Eltetö and Köves 1964; Gini 1931; Szulc 1964), the Country Product Dummy method (Summers 1973), and the Geary–Khamis method (Geary 1958; Khamis 1972). The reader may also consult Balk (2008) for details about these and other multilateral methods.

Although multilateral methods have a long history in the geographical domain, studies concerning their potential for price comparisons over time are of a much later date, inspired by the increased interest of NSIs in transaction data (Haan and Grient 2011; Ivancic et al. 2011). Sets of countries or regions are replaced by months when applying multilateral methods in the CPI. A set of consecutive months is usually referred to as a *time window*. Multilateral methods generate index numbers simultaneously for all months in a time window. This is an important difference with bilateral methods, which are restricted to price comparisons between two periods.

The independence of reference country obviously also holds when applying multilateral methods to the time dimension. The independence of reference period can be achieved in different ways with multilateral methods:

1. By calculating an average of bilateral price comparisons over a time window. This approach is used by the GEKS family of index methods (Haan and Grient 2011; Ivancic et al. 2011).
2. By comparing item prices in different periods with average or *reference prices* computed over a full time window. The Geary–Khamis (GK) method (Chessa 2016) and the Time Product Dummy (TPD) method (Krsinich 2014) are well-known examples of methods that follow this approach.

A comparison of these multilateral methods when applied to transaction data will be made in the next subsection, followed by a motivation of the chosen method in the Dutch CPI.

4.2 Why GK?

The three types of multilateral method mentioned in the previous subsection share important features:

- All methods satisfy transitivity, the identity test in the 'narrow sense' (ILO et al. 2004, p. 293, fn. 21), the proportionality in prices tests—which are important for the constant-tax CPI—and commensurability.
- As was already stated, multilateral methods are capable of processing data of dynamic product assortments and changes in expenditure patterns over time, without affecting transitivity.

The first experiences at Statistics Netherlands with applications of the Fisher index method on transaction data make clear that satisfactory behaviour on axioms and on economic-theoretic grounds does not guarantee adequate responses to typical phenomena occurring in transaction data, which is at least as important. Although different multilateral methods are capable of dealing with various types of dynamics observed in transaction data, the two approaches distinguished in Section 4.1 react differently in some situations.

The GEKS family of methods uses bilateral indices in order to compute multilateral index series. This class of methods therefore inherits the *matched-model* principle from bilateral index methods. This principle has an important implication: products that are sold in most or in all periods of a time window enhance their contribution to an index over products that are sold in a few periods, since more bilateral comparisons can be made with the former products. This property may distort aggregate price change in various cases:

- Products that are removed from the market. Clearance prices will have a disproportionate downward impact on price change when the corresponding products are still sold in many periods of a time window. This effect is particularly felt when new products just enter the market.
- Seasonal items. Smoothing effects in aggregate price change may occur, in particular when certain items are sold in short periods of a year.

The negative impact of clearance prices will be amplified when using the Törnqvist or Fisher index formula to compute GEKS indices (Chessa et al. 2017, pp. 19–22). This is strongly related with the discussion on the Fisher index at the end of Section 3.1. The Walsh index is more robust to clearance prices. Unfortunately, undesirable effects on aggregate price change may still occur. A recent study on shrinkflation evidenced serious problems with GEKS–Walsh to measure the impact of changes in package content on inflation. Also these problems can be ascribed to the use of the matched-model approach (Chessa and Barb 2025).

Clearance prices can be removed by making use of the thresholds used in the Jevons-based method (Section 3.3). However, this is a less straightforward operation as might seem. An example included in the cited shrinkflation study shows that products that are about to be cleared are not always easy to identify.

Applications of GEKS methods to transaction data also make use of techniques to impute prices of removed items and of (temporarily) unavailable products. This can be seen as a response to the unequal treatment of products with different numbers of bilateral comparisons in a time window. While this may mitigate the aforementioned problems, imputing prices in order to fill the whole time–product panel does not resolve these problems because of the lower weights of items in periods without sales.

The problems signalled with the GEKS class of methods do not occur with GK and the expenditure-based weighted version of TPD. Price comparisons with reference prices are independent of the number of periods in which prices are available. In addition, clearance prices and out-of-season prices do not distort GK and TPD indices because of the very low expenditure shares associated with these prices.

The GK method has a number of advantages over GEKS and TPD:

- Compared with GEKS:
  - appropriate handling of all types of dynamics in transaction data, including clearance prices and seasonalities;
  - no need of thresholds—only for outliers, due to data errors—and imputations;
  - as a consequence, GK does not require the data engineering and fabrication of artificial microdata commonly applied with GEKS. The method treats transaction data in a natural way, as it is collected and delivered by retail chains.
- Compared with GEKS and TPD:
  - GK assigns *equal importance* to every single transaction, in contrast with GEKS (matched model), but also with TPD. The latter uses expenditure shares as weights for the time periods in the reference prices. These weights make sense in the product dimension, but not in the time dimension;
  - GK price indices simplify to aggregate expenditure change when baskets and quantities between two periods are fixed, thus satisfying the *fixed-basket test*. This situation occurs in the CPI (e.g. with rents);
  - GK price indices move towards unit value indices as quality differences between products become smaller. Unit value indices should be calculated for products of the same quality (ILO et al. 2004, p. xxii).

These observations motivate the choice in favour of GK for multilateral index calculations with transaction data in the Dutch CPI.

4.3 Index formulas

Like other multilateral methods, the GK method computes price indices in each period of a time window simultaneously. Interestingly, the index expressions of this method can be directly derived as dynamic refinements of the first-used Lowe index method in the Dutch CPI (Section 3.2). In order to show this, some additional notation is introduced.

Let $[0, T]$ denote a time window containing $T + 1$ successive months $0, 1, \dots, T$. The number of months $T + 1$ is the length of the time window. Let G_t denote a set of items sold in month $t \in [0, T]$ with strictly positive average transaction prices (unit values) $p_{i,t} > 0$ and sold quantities $q_{i,t} > 0$ of item $i \in G_t$.

Consider expression (3) for the (unnormalised) Lowe index, with the following changes applied to this formula:

- The reference year r is replaced by the time window $[0, T]$.
- Aggregate price levels $P_t^{[0,T]}$ are computed in each month of the time window.
- The set of items G^r is replaced by the set G_t in month $t \in [0, T]$.
- The item quantities q_i^r of reference year r are replaced by the quantities $q_{i,t}$ sold in month t.
- The item-specific reference prices p_i^r are replaced by reference prices $\nu_i^{[0,T]}$ defined on the time window $[0, T]$. Also the $\nu_i^{[0,T]}$ are computed as unit values, but in this case the prices $p_{i,t}$ are deflated by $P_t^{[0,T]}$ in every $t \in [0, T]$.

One of the most essential differences with the Lowe index is that the months in which indices are calculated fall within the time window. The transaction data in the window are used to determine a reference price for all items, including new items. This choice makes it possible to process all items in a window. The modifications listed above give rise to the following system of equations:

$$P_t^{[0,T]} = \frac{\sum_{i \in G_t} p_{i,t} q_{i,t}}{\sum_{i \in G_t} \nu_i^{[0,T]} q_{i,t}}, \tag{6}$$

$$\nu_i^{[0,T]} = \frac{\sum_{z=0}^{T} p_{i,z} q_{i,z} / P_z^{[0,T]}}{\sum_{z=0}^{T} q_{i,z}}, \tag{7}$$

for all $i \in G_t, t \in [0, T]$. These equations are equivalent with the Geary–Khamis system in the time domain.

Note that $P_t^{[0,T]}$ and $\nu_i^{[0,T]}$ appear in both expressions, which consequently cannot be computed directly. How indices and reference prices satisfying equations (6) and (7) can be computed will be treated in Section 4.5. For a unique solution (up to a scale factor) to exist for these equations, it is sufficient to assume that there is no partition of the set of months $\{0, 1, \dots, T\} = A \cup B$ such that every item is observed exclusively in either A or B. This condition ensures connectedness of all months in the time window, either directly or indirectly. This means there is a single coherent universe of item prices and quantities over time.

The deflation of item prices in expression (7) leads to a decomposition of the time–product panel of item prices and quantities into item-specific reference prices and aggregate price levels. The deflation of item prices projects these prices onto an arbitrary common reference month across items, which assigns a precise meaning to expression (6): an aggregate price level is measured by relating nominal to real expenditure, which is the ratio on the right-hand side of (6).

Like Lowe expression (3), GK index formula (6) does not return the value 1 for an arbitrarily chosen reference month. Also in this case a normalisation is needed. Let the first month 0 of the time window be the reference month. The price index $P_{0,t}^{[0,T]}$ in month t with respect to the reference month can then be obtained as follows:

$$P_{0,t}^{[0,T]} = \frac{P_t^{[0,T]}}{P_0^{[0,T]}} = \frac{\sum_{i \in G_t} p_{i,t} q_{i,t} / \sum_{i \in G_0} p_{i,0} q_{i,0}}{\sum_{i \in G_t} \nu_i^{[0,T]} q_{i,t} / \sum_{i \in G_0} \nu_i^{[0,T]} q_{i,0}}. \tag{8}$$

Note that the deflators in (7) can be replaced by $P_{0,z}^{[0,T]}$, which leaves index expression (8) invariant:

$$\nu_i^{[0,T]} = \frac{\sum_{z=0}^{T} p_{i,z} q_{i,z} / P_{0,z}^{[0,T]}}{\sum_{z=0}^{T} q_{i,z}}. \tag{9}$$

For convenience, the same notation as in (7) is used for the reference prices.

It is easily verified that formula (8) generates transitive index series on $[0, T]$, as the reference prices $\nu_i^{[0,T]}$ are constant over the full time window. GK price index (8) results as a ratio of an expenditure index to a weighted quantity index. The quantity index is completely determined by the reference prices and the quantities. This shows again the importance of the deflators in (7) and (9), since aggregate price change is excluded from the quantity index.

4.4 How quality fits in

The GK method offers a formal framework in which the change in total expenditure is decomposed into aggregate price and quantity change. The deflation of item prices in the reference prices plays an essential role in separating price from quantity change.

The topic of product quality has not explicitly been addressed so far, so it is important to clarify how quality is incorporated in the formulas presented in the previous subsection. Quantities of items with different quality characteristics cannot be simply summed. Quality differences are taken care of in the denominators of expressions (6) and (8), with the item-specific reference prices $\nu_i^{[0,T]}$ absorbing these differences across items in the multilateral window.

If the reference prices have the same value for all items, it is easy to see that expression (6) simplifies to a unit value and expression (8) to a unit value index. This corresponds with the special case where all items have the same quality, which leads to the desired index formula (ILO et al. 2004, p. xxii). For this reason, expression (8) was named *Quality-adjusted Unit value index* (QU index) in Chessa (2016). Reference prices can be defined in different ways. So GK represents a specific member of the QU family of index methods, and is often referred to as QU–GK. More details about other multilateral QU-type methods, including comparisons with GK, can be found in Chessa 2016.

Expressions (6)–(9) are specified at the item/GTIN level. Applying these formulas at this level of product definition does not give correct results when product relaunches occur with 'hidden' price changes. The reference prices would then differ for items with different GTINs for which quality has remained the same. The reference prices of such items should have the same value.

This example shows that interventions may be needed before running GK in order to tailor the reference prices to quality differences and changes across items and over time. The type of intervention depends on the type of relaunch:

- Relaunches without changes in the consumable part of an item. Changes in the design or the packaging of an item may occur, as long as these changes do not affect quality and net content.
- Relaunches where an item's net content or quality characteristics are modified.

The first case can be handled by defining broader strata of items with common product characteristics. The strata should ideally be tightly defined, but not too tight in order for disappearing and reintroduced items to be captured within the same stratum, which allows to detect possible price changes associated with relaunches. The method MARS was specifically designed for this purpose (Chessa 2021a). This approach results in quality adjustments that are known as *direct comparison* within strata, while new items forming new strata will not affect a price index in the period of introduction (Eurostat 2024, Sec. 6.5).

Relaunches that are accompanied by changes in net content require explicit adjustments of item prices and quantities in order to obtain the effective price change per unit of measurement (e.g. per kg or litre). Such cases can also be handled by finding broader product definitions, so that items with both the original and modified package content can be combined within the same stratum. An application of MARS and GK can be found in the previously cited shrinkflation study (Chessa and Barb 2025).

Relaunches with quality changes represent the most difficult case. The question is whether a modified item should be considered as a totally new product—with price change fully reflecting quality change—or as an existing item with limited changes in quality. A question that arises in the second case is whether the quality changes can be safely ignored or should be used for price adjustments. These questions typically emerge for products characterised by frequent technological changes, notably consumer electronics. This is an area of the CPI that attracts continuous debates, attention and research.

The above problems not only illustrate the complexity of the product definition problem and the related topic of quality, and how to account for these problems in an index formula, but also highlight the importance of the availability of product information in transaction data.

Returning to GK, it is important to emphasise that the formulas should be applied in the same form at any stratification level, with items replaced by strata of items. Expenditures and quantities are summed over all items within the same stratum. It is easy and also interesting to see that formulas (6) and (8) can be expressed in terms of expenditures and quantities at the item level also when using broader product definitions. The item-specific reference prices will then have the same value for all items within the same stratum. This means that items within the same stratum are correctly treated as items of the same quality by formulas (6) and (8).

4.5 Index calculation

Equations (6)–(7) and (8)–(9) form coupled systems in which the aggregate price levels and reference prices depend on each other. As a consequence, it is not possible to calculate price indices directly from index formula (8). However, this can be achieved by making use of an iterative algorithm that successively updates the price indices and the reference prices until these values converge to a stable solution.

Equations (6)–(7) and (8)–(9) can be written in the form $\mathbf{P} = \Phi(\mathbf{P})$. In the case of (8)–(9), $\mathbf{P}$ is a $(T+1)$-dimensional vector of price indices, and the mapping $\Phi : \mathbb{R}^{T+1}_{++} \to \mathbb{R}^{T+1}_{++}$ returns $P^{[0,T]}_{0,t}$ for month t. A vector $\mathbf{P}$ satisfying $\mathbf{P} = \Phi(\mathbf{P})$ is known as a *fixed point* in mathematics. From an economic point of view, the GK fixed point is an equilibrium in which the item-specific reference prices and price indices are consistent with all observed transactions and expenditures. There is only one equilibrium when the connectedness condition stated in Section 4.3 is satisfied.

GK price indices can be obtained by using the relation $\mathbf{P} = \Phi(\mathbf{P})$ to update the price indices in an iterative algorithm, by substituting the result $\mathbf{P}^k$ of the k-th iteration in expression (9) and then use the resulting $\nu_i^{[0,T]}$ to derive an updated vector $\mathbf{P}^{k+1}$ according to (8) in the next iteration. In summary, the adopted scheme is: $\mathbf{P}^{k+1} := \Phi(\mathbf{P}^k)$ after an initialisation of either the item-specific reference prices or the price indices in each month of the window.

An iterative algorithm for obtaining price indices according to expressions (8) and (9) can be set up as follows:

1. A convenient initialisation can be obtained by ignoring the price indices to deflate the prices in (9), which yields an initial set of values for the item-specific reference prices. Proceed to step 2.
2. The reference prices are substituted in expression (8), which yields an initial set of price indices in every month of the time window. Proceed to step 3.
3. The price indices are substituted in expression (9) to obtain updated reference prices. Proceed to step 4.
4. The updated reference prices are substituted in expression (8), which yields updated price indices. Proceed to step 5.
5. If the absolute differences between the index values computed in the last two iterations drop below a predefined threshold, then the algorithm terminates and the last set of index values is accepted as the final result. Otherwise, take the result of step 4 and continue with step 3.

The initial set of price index values generated by steps 1 and 2 are named *augmented Lehr indices* in Auer (2017). It should be clear that the use of the price indices as deflators in expression (9) must be restored in step 3 after the two initialisation steps. The algorithm converges to a unique solution, which is independent of the values chosen in the initialisation step.

A concise and generic proof of convergence is given in Appendix A. The proof also applies to other QU-type multilateral index formulas (Chessa 2016) and generalises to TPD, which is also a fixed-point method.

5 Index extension methods

5.1 Non-revisability and drift

The GK method described in the previous section generates index series on a given time window $[0, T]$. Calculating indices with new data after month T while keeping the window fixed would bring us nearly back to the Lowe index method described in Section 3.2. The fixity of the window will still not allow us to include new items that cannot be linked to existing or removed items.

The aim is to process new items as well, and in a timely way. This can be accomplished by adapting the initial window $[0, T]$ so that all data from the next month $T + 1$ can be included in the window. A new index series can then be calculated on the adapted window, which yields an index in $T + 1$. The inclusion of new data will normally result in different index values on the new window until month T compared with the values obtained on the initial window $[0, T]$. Index values that change in past months as a window is adapted to accommodate new data create a problem: price indices cannot be revised in the CPI once published.

This *non-revisability* condition has important implications for the construction of CPI series:

- Item-specific reference prices and price indices computed on each successive window have to be treated as fixed values.
- A linking mechanism has to be developed that combines the results of consecutive windows into a non-revisable index series for CPI publication.
- Because of the fixity of window estimates, non-revisable index series are not transitive.

The first two points tell us that CPI series cannot be generated by exclusively using a multilateral index formula. The purpose of a multilateral index formula in a CPI context is to deliver window estimates, which are subsequently used by a linking technique—usually referred to as *index extension method* or *splicing method*—to evaluate price change with respect to a chosen reference month and to use this result to extend CPI series from one month to the next. In the case of the Dutch CPI this means that the series that are eventually published are not classical GK indices, but the result of applying a larger framework.

Chain drift is reintroduced when generating index series under the non-revisability condition. The key question therefore is whether chain drift can be controlled in extended index series. Fortunately, multilateral methods have a clear advantage over bilateral methods also with regard to this specific problem: the larger number of variables in multilateral methods offers more possibilities to link indices of consecutive windows and to control chain drift.

Multilateral methods allow to make different choices on various variables, as described in Chessa (2023) and Chessa (2025):

- The *length* of the time window.
- The type of *window adaptation* to incorporate data from the next month in a window, which can be achieved by: (1) expanding the current window with one month, (2) shifting a fixed-length rolling window by one month.
- The *linking interval*, which is the number of months that separates the current month from the month in which an index is linked (the *linking month*) to extend an index series.
- The *linking index* in the linking month, which can be: (1) the first calculated index, or (2) a recalculated index.
- The listed variables could be kept fixed over time or varied (*static* vs *adaptive splicing*).

An illustration of how these variables can be applied to extend index series is presented in Figure 5.1. Two index extensions are shown, which consequently produce two published indices. The two indices are obtained by first deriving indices from series 2 and 3 for the time lag specified by the linking interval, with the linking month acting as reference month.

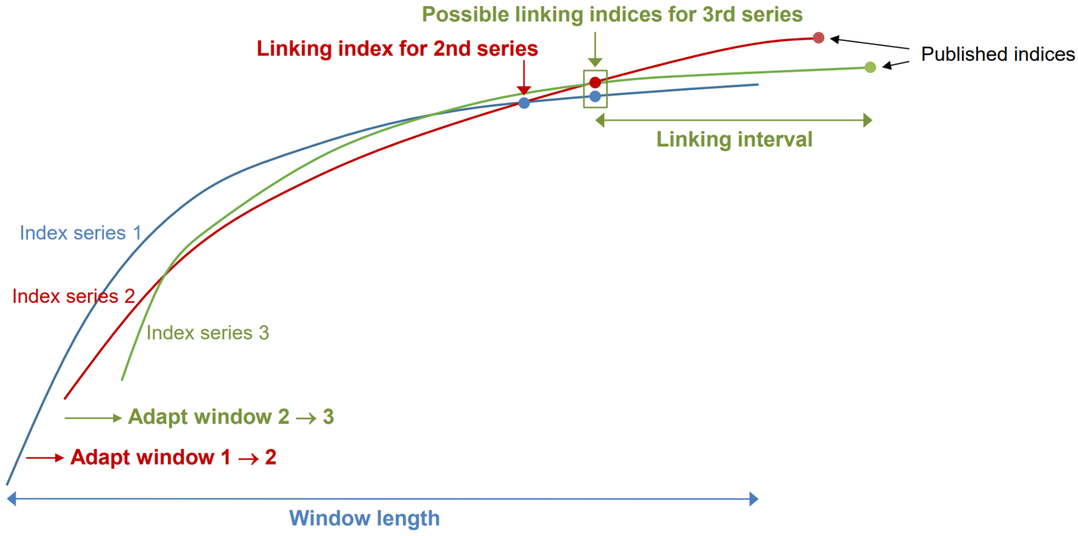
In the next step these indices are linked on an index in the linking month. There is only one index to link on after computing index series 2, which is the index of the first series in the linking month. There are two candidate indices to link on after computing the third index series: the index of the first series in the linking month and the index of the second series that is obtained after linking this series to the first index series. The index in the linking month of the linked second series is a recalculated index. In the example of Figure 5.1 the third index series is linked on the recalculated index.

Indices in past months are recalculated as windows are adapted over time. As Figure 5.1 illustrates, using a recalculated index to produce extended index series does not lead to a revised index. The method HYCS described in Section 5.3 is an example of a method that links on recalculated indices.

The fifth variable in the above list is not explicitly shown in Figure 5.1. The choice made for this variable is ‘static splicing’, since the same window length and linking interval are used in the two index extensions, while the choice to link on (the last) recalculated index would also be made in subsequent index extensions.

A broad range of index extension methods can be derived by combining the different choices for the above variables (Chessa 2023; Chessa 2025). The next two subsections describe the extension methods that have been used on a large scale in the Dutch CPI.

Figure 5.1 Extension variables (bold face) and how these can be used to extend index series.



5.2 The FBEW method

The first extension method used in the Dutch CPI is the Fixed Base Expanding Window (FBEW) method. This method was adopted together with the GK method in January 2016. Details of the FBEW method can be found in Chessa (2016), where it was given the name "real time method". The description of the method in this section makes use of the previously introduced notation.

As the name suggests, the FBEW method uses an expanding window. December of the previous year is the first month of each time window, which thus contains two months in January of the current year, three months in February, and is extended until it contains the maximum number of 13 months in December of the present year. The GK index series computed on each window are used to derive a fixed-base index with respect to December of the previous year. The fixed-base index is linked on the index published in December.

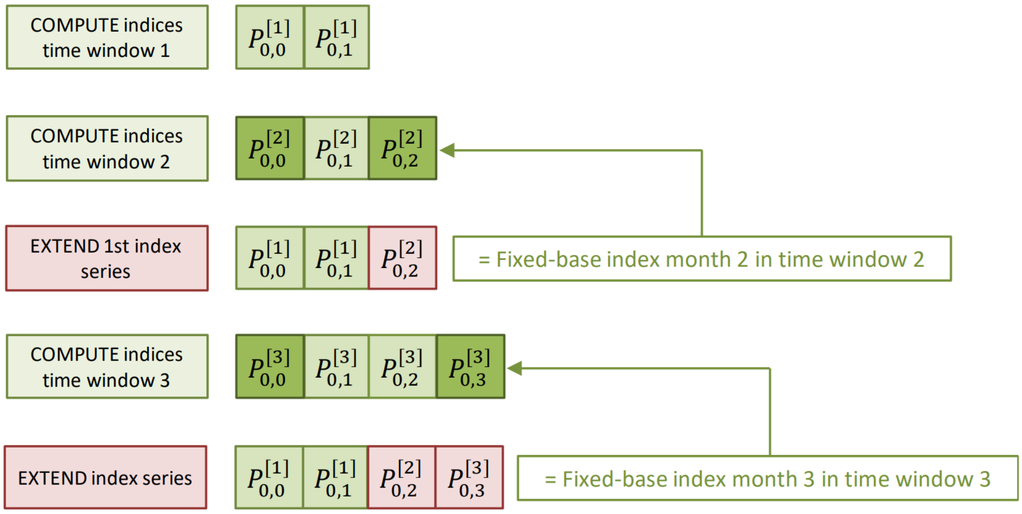
This procedure can be formalised as follows. An index in month t of an extended non-revisable index series starting in month 0 is denoted as $P_{0,t}$. The symbol r for reference year, which was used in Section 3.2, is also used here. The expanding window in month t of year $r+1$ used by the FBEW method is $[12r, t]$, where month $12r$ represents December of year r. Following the notation introduced in Section 4.3, the fixed-base index in month t derived from the GK index series on window $[12r, t]$ is $P_{12r,t}^{[12r,t]}$. The extended index in month t, which will be eventually published according to the FBEW method, can be obtained as follows:

$$P_{0,t} = P_{0,12r} P_{12r,t}^{[12r,t]}, \tag{10}$$

with $t \in [12r + 1, 12(r + 1)]$. An illustration of the FBEW method is given in Figure 5.2.

The FBEW method was introduced in the Dutch CPI together with GK for mobile phones in January 2016. The application of the GK–FBEW method was gradually expanded to

Figure 5.2 Illustration of the FBEW method for two index extensions.



transaction data of other retail chains in successive years (Figure 3.1). The replacement of the Jevons-based method by GK–FBEW for all supermarket chains in January 2018 marked a significant moment, as Statistics Netherlands became the first statistical office worldwide to apply a multilateral method on such a large scale. An operation of this size requires careful preparations, deep analyses and reporting the impact of the new method on inflation to Eurostat (Chessa et al. 2018).

The FBEW method is still used in the Dutch CPI for mobile phones. It is also used for energy prices, which represents a case where the FBEW method is tailored to the data. Part of the variables in the energy data are defined on a yearly basis, so that December plays a double role in the construction of index series: to compute an index with the data of the current year and to initialise a new series that will use the new yearly values of the next year.

A number of properties of the FBEW method are worth emphasising:

- Index series are free of drift at the end of each year, since the FBEW indices in December will coincide with the December indices of multilateral series computed on 13-month windows.
- The method does not require data from past years.
- The FBEW method perfectly fits with typical end-of-year CPI practices, such as preparing next year's CPI compilation for changes in classification and product stratification. The index series in each year only make use of the classification and stratification established for each respective year.

The FBEW method provides an illustration of how drift can be controlled in extended non-revisable index series. Chain drift cannot be excluded in each month, but the linking on the December indices ensures that the index values in December of each year will equal transitive indices computed on 13-month windows spanning periods between two consecutive December months. This prevents an FBEW index series to deviate from its transitive counterpart at the end of each year.

The drift-free property of the FBEW method is valid conditional upon the length of the time window; in other words, when 13-month windows can be considered ‘sufficiently’ long. Additional details and discussions on the behaviour of the FBEW method and expanding window variants can be found in Diewert and Shimizu (2024).

5.3 The method HYCS

Beside the nice properties mentioned at the end of the previous subsection, the FBEW method also has some limitations:

- Windows of 13 months may be too short in situations with high product churn and for seasonal items.
- The use of a fixed linking month (December of the previous year) also fixes the linking interval in each month. This removes a degree of freedom from the range of choices listed in Section 5.1 to control drift.
- The number of months in the expanding window is small in the first months of a year.

A comparative study on extension methods applied to transaction data of seasonal items has shown that 13-month windows are generally too short for these types of product (Chessa 2021b). Large deviations of extended index series from transitive ‘benchmark’ series computed on a longer window were found in particular for garden furniture. This type of product contains many items that are only sold in spring and summer. Such items will therefore fill a relatively small number of months in a 13-month window.

A logical response to limited data coverage in 13-month windows is to consider longer windows. The cited study shows that much better results are obtained with 25-month windows. Apart from better data coverage, the use of an additional year is advantageous for seasonal items also because price comparisons between successive years can be made in every season.

The second limitation holds for all extension methods that apply fixed-base index linking. When also considering the third limitation, the use of a fixed-length rolling window—with the linking interval to be determined—is a better alternative with more flexibility when making choices about extension variables. Beside choosing an appropriate window length and linking interval, the question is on which index to link in the linking month.

Since the 12-month rate of price change is the commonly used inflation measure in policy, an interesting choice would be to link the year-on-year indices computed on each time window on the indices published 12 months ago. This way of extending index series leads to the *Half Splice on Published indices* (HASP) method when using 25-month windows. This method has the big advantage of preserving the 12-month rates of change computed on successive time windows in extended index series, which will therefore also appear in published figures.

However, recent studies have also shown that results for other time lags, like monthly rates of change, can be very inaccurate (Chessa (2023, pp. 40–41);Chessa (2025, pp. 45–46)). The month-on-month indices calculated with HASP depend on the indices published in past periods. Inaccuracies in past indices may accumulate and propagate to current values.

Accurate monthly rates of change can be obtained by linking on recalculated indices, in particular on the *last* recalculated index. While this type of linking might sound complicated, index series can be extended in a simple and elegant way with this approach.

Linking on the last recalculated index of L months ago produces a month-on-month chained non-revisable index series. The lag L is the linking interval for which a suitable value has to be determined. In order to explain and formalise this type of extension method, let $[0, T]$ denote the window on which the first index series is generated. The next fixed-length rolling window is $[1, T + 1]$, where $T + 1$ is the month in which the first index extension takes place. The index in month $t > T$ is obtained by linking a month-on-month index $P_{t-1,t}$ on the index published in month $t - 1$:

$$P_{0,t} = P_{0,t-1} P_{t-1,t}. \tag{11}$$

The month-on-month index $P_{t-1,t}$ can be derived as follows. Let $P^*_{0,t-L}$ denote the last recalculated index obtained in month $t - L$ with respect to month 0 after extending an index series until month $t-1$. The index series is extended to month t by linking the index $P^{[t-T,t]}_{t-L,t}$ in t with respect to the linking month $t - L$, which is computed on the window $[t - T, t]$, on the last recalculated index in month $t - L$. This procedure is illustrated in Figure 5.1 with the linking of the third index series. The index $P_{0,t}$ of the extended series in month t is thus obtained as follows:

$$P_{0,t} = P^*_{0,t-L} P^{[t-T,t]}_{t-L,t}. \tag{12}$$

Multiplying $P^*_{0,t-L}$ with the index $P^{[t-1-T,t-1]}_{t-L,t-1}$ in month $t - 1$ with respect to linking month $t - L$, which is computed on the previous window $[t - 1 - T, t - 1]$, gives the index $P_{0,t-1}$ of the extended non-revisable index series in month $t - 1$ (see also Figure 5.1). Expression (12) can now be written as follows:

$$P_{0,t} = P_{0,t-1} \frac{P^{[t-T,t]}_{t-L,t}}{P^{[t-1-T,t-1]}_{t-L,t-1}}. \tag{13}$$

When linking on the last recalculated index of L months ago, the month-on-month index $P_{t-1,t}$ results as a ratio of indices computed on the last two time windows, which is the ratio on the right-hand side of (13). Note that the two indices in this ratio are computed with respect to the same reference month, which is linking month $t - L$.

The question is how to find suitable choices for the window length $T + 1$ and the linking interval L. The month-on-month indices follow from index series on time windows that are shifted by one month. Successive windows therefore have a high degree of overlap, which increases with longer windows. The number of common months in two consecutive 25-month windows is 24 months, which results in a very high overlap percentage. Almost the same data are consequently used in two consecutive windows, which explains the high accuracy in the monthly rates of change for methods that link on the last recalculated index when using 25-month windows (Chessa 2025, p. 44).

With the high accuracy achieved in monthly rates of change when linking on the last recalculated index, the attention can be shifted towards controlling the behaviour of extended series over longer time lags. A time lag of particular interest when measuring inflation is 12 months. The amount of chain drift in month t can be measured by comparing the year-on-year index $P_{t-12,t}$ of an extended index series with the direct year-on-year index $P_{t-12,t}^{[t-T,t]}$ computed on the window $[t-T,t]$.

Let $\Delta_{t-12,t}$ denote the *drift factor* for the year-on-year index $P_{t-12,t}$ of an extended index series, which is defined as follows:

$$\Delta_{t-12,t} = \frac{P_{t-12,t}}{P_{t-12,t}^{[t-T,t]}}. \tag{14}$$

The index $P_{t-12,t}$ of extended index series (13) can be written as a chained index of 12 month-on-month indices:

$$P_{t-12,t} = \prod_{z=t-11}^{t} \frac{P_{z-L,z}^{[z-T,z]}}{P_{z-L,z-1}^{[z-1-T,z-1]}}. \tag{15}$$

The drift factor for this index can be written in the following form when computing GK index series on each window:

$$\Delta_{t-12,t} = \prod_{z=t-11}^{t} \frac{\sum_{i\in G_{z-L}} w_{i,z-L}^{z-1} \frac{\nu_i^{[z-T,z]}}{\nu_i^{[z-1-T,z-1]}}}{\sum_{i\in G_{t-12}} w_{i,t-12}^{z-1} \frac{\nu_i^{[z-T,z]}}{\nu_i^{[z-1-T,z-1]}}} \tag{16}$$

$$w_{i,\tau}^{z-1} = \frac{\nu_i^{[z-1-T,z-1]} q_{i,\tau}}{\sum_{j\in G_\tau} \nu_j^{[z-1-T,z-1]} q_{j,\tau}}. \tag{17}$$

Each of the 12 product terms in expression (16) is a ratio of weighted sums of reference price ratios computed on two successive time windows. The weights apply to items sold in months $z-L$ (numerator) and $t-12$ (denominator) and are given by expression (17). The weights are defined as an item's real expenditure share in month τ (i.e. $z-L$ or $t-12$), with sold units valued according to reference prices computed on the time window in month $z-1$.

Expression (16) is a function of window length $T+1$ and linking interval L. The drift factor can thus be influenced by varying these two factors. There will be no drift in the extended year-on-year indices when $\Delta_{t-12,t}=1$ for all t. The drift factor can be driven towards this ideal value in a number of ways:

- By increasing the window length $T+1$.
- By positioning the 12 linking months $z-L$ in the weights $w_{i,z-L}^{z-1}$, with $z=t-11,\dots,t$, as close as possible to month $t-12$ in which the weights $w_{i,t-12}^{z-1}$ are evaluated.

Increasing the window length $T+1$ will move the ratios of reference prices computed on successive windows to the same value, up to a scale factor which drops out of each ratio in expression (16). Choosing very long windows is beneficial for reducing drift, but is also impractical for different reasons (data over a long period are needed, and data over longer periods are more sensitive to changes applied by retailers). The comparative studies in Chessa (2023) and Chessa (2025) have shown that accurate 12-month rates of change can be obtained with 25-month windows.

Drift in the year-on-year indices can be eliminated by setting the months $z - L$ equal to $t - 12$. The weights $w_{i,z-L}^{z-1}$ will then equal $w_{i,t-12}^{z-1}$, so that $\Delta_{t-12,t} = 1$. By substituting $z - L = t - 12$ in expression (15) it follows that the resulting index extension is equivalent with the method HASP, which links year-on-year indices on the indices published 12 months ago. However, as was already mentioned, this method may produce inaccurate month-on-month indices.

It seems very hard, if possible at all, to improve the accuracy in the monthly rates of change generated by HASP while preserving the direct year-on-year indices in the published inflation figures. When linking on the last recalculated index it is possible to control both rates of change: the accuracy in monthly rates of change is ensured by the high overlap of the two consecutive time windows involved in the calculation of month-on-month indices, while the accuracy in 12-month rates of change can be controlled with suitable choices for the linking interval L.

Optimal values of L can be found by minimising the sum of the distances between the 12 linking months $t - 11 - L, \ldots, t - L$ in the chained index (15) and the reference month $t - 12$ in the direct year-on-year index. The optimal values are $L = 6$ and 7 months. These two values also maximise window overlap as defined in Chessa (2025). The choice of $L = 6$ is equivalent with the extension method referred to as *Half Year reference Chained Splice* (HYCS).

How the two optimal values of L control drift in 12-month rates of change can be understood as follows. The linking intervals satisfy the property $t - 12 \in I_{t,L}$, where $I_{t,L}$ is the set of linking months $\{t - 11 - L, \ldots, t - L\}$. This set can be partitioned as $I_{t,L} = I_{t,L}^{(1)} \cup I_{t,L}^{(2)} \cup \{t - 12\}$, with non-empty sets $I_{t,L}^{(1)} = \{t - 11 - L, \ldots, t - 13\}$ and $I_{t,L}^{(2)} = \{t - 11, \ldots, t - L\}$.

Recall that $\Delta_{t-12,t} = 1$ when $z - L = t - 12$. Trends in expenditure shares are a common phenomenon as the composition of item purchases changes over time. Expression (16) then typically decomposes into a component smaller than 1 on $I_{t,L}^{(1)}$ and a component larger than 1 on $I_{t,L}^{(2)}$ or vice versa. These two components capture deviations of the shares $w_{i,z-L}^{z-1}$ from the shares $w_{i,t-12}^{z-1}$ in the reference month, which cancel each other over time. The comparative studies in Chessa (2023) and Chessa (2025) show no drift in 12-month rates of change at all-items level and at 2-digit COICOP level when using 25-month rolling windows. Also the results at the most detailed retailer-COICOP level are accurate.

The version of HYCS with a 25-month window was introduced in the Dutch CPI in January 2024 and is used on a large scale for transaction data since January 2025. An illustration of the method is presented in Figure 5.3. The visualised index extensions follow the steps in the monthly chained index calculation according to expression (13).

Beside its satisfactory drift behaviour, HYCS also behaves well from an axiomatic perspective. This method satisfies, among others, the proportionality, monotonicity and commensurability axioms. Although transitivity is lost in extended index series, HYCS manages to control chain drift very well, and consequently also limits deviations from

Figure 5.3 Illustration of the method HYCS for two index extensions.

the strict identity and fixed-basket properties.

Another very attractive property of HYCS—which is shared by all methods that link on the last recalculated index—is that index series on arbitrary time intervals are independent of extended indices computed in previous periods. Index series produced by HYCS can thus be said to be 'autonomous index series'. This is a particularly elegant and powerful property in the context of the CPI: for instance, index series computed in calendar years are uniquely determined. This is not the case for a big subclass of methods that link on published indices, like HASP, since the resulting index series depend on the choice of the initial month of the extended series.

It is also important to emphasise that the use of a rolling window in the CPI requires additional effort compared with the FBEW method, since such windows cover more than one calendar year. Rolling windows may therefore contain changes to a new classification or product stratification. These and other problems have been dealt with in separate studies at Statistics Netherlands, for which satisfactory solutions have been found.

5.4 Other valid methods

The broad comparative studies on extension methods described in Chessa (2023) and Chessa (2025) resulted in several other methods that rank among the best regarding their capability of controlling chain drift. Two methods are briefly discussed and compared with HYCS in this subsection.

The use of a single linking interval L as in HYCS could produce results that are sensitive to the specific choice of L in certain months. Choosing a set of linking intervals around the optimal values of 6 and 7 months could mitigate such sensitivities. The trimmed mean splice method TRIMS was proposed for this reason, which computes an average of the extended indices for six values of the linking interval (Chessa 2025). However, the cited study shows that HYCS and TRIMS produce almost the same results, which was subsequently confirmed in Chessa and Barb (2025).

A different approach is to allow the window length and linking interval to vary from month to month. This extension method is called *adaptive splicing* (Chessa 2023). This method requires more data than static splicing methods like HYCS and TRIMS. Adaptive splicing also requires more computational effort since it optimises a drift factor each month. In spite of using additional data and an optimisation procedure, the cited study shows better results for HYCS at all-items level, while improvements of adaptive splicing over HYCS are small for lower aggregates.

The analyses and findings summarised in the previous and present subsection provide a justification for choosing the simpler method HYCS with a 25-month window as the current extension method in the Dutch CPI. This choice also has the advantage of simplifying the derivation of other methods designed for specific purposes in CPI production, such as methods for computing contributions of individual items to aggregate price change.

6 Final remarks

This document presented the various phases in the development of index methods for transaction data in the Dutch CPI. The foundations for the current methodology were laid more than 25 years ago, when the route was planned towards the introduction of transaction data in the Dutch CPI in 2002. Several other big changes were introduced at the time: beside changing from traditional price collection to transaction data for two supermarket chains, a new index method was implemented, a switch was made to annually updated weights at retailer-COICOP level, while efforts were also made to automate the classification of GTINs. All this was achieved in a different time, when the topic of transaction data was totally new, on an international scene where today's frequent exchanges of experiences were lacking. This can truly be considered a big achievement.

Section 4.3 of the present document has shown that the current multilateral methodology can be obtained as a natural refinement of the first method from 2002, which is now fully tailored to the different forms of dynamics that are observed in transaction data sets. A versatile methodology has been found that meets the goal set at the end of 2013, when it became clear that the methodological landscape that was being created was no longer sustainable. The current methodology rests on three pillars:

- The method MARS for defining products and detecting relaunches, which can be handled in an automated way.
- The GK method produces item-specific economically meaningful reference prices on subsequent time windows that are robust to typical phenomena observed in transaction data, like clearance prices.
- The extension method HYCS with a 25-month window then uses these reference prices to build a monthly chained index. The method is explicitly designed to minimise drift in 12-month rates of change, without affecting the high accuracy in short-term changes.

The multilateral method used for almost all transaction data in the Dutch CPI is consequently GK–HYCS. As was stated in Section 5.1, this method is different from classical

GK. The non-revisability condition imposed by CPI series requires a more comprehensive approach: GK is concerned with window estimation, while HYCS uses the estimated item-specific reference prices to evaluate price change and extend CPI series accordingly, thus taking care of the time dimension.

The compound GK–HYCS method has satisfactory axiomatic and other properties, and is able to control chain drift very well. In spite of these merits, the results of any index method depend on the quality of its input. An essential part of the input is information about products in transaction data sets, beside item expenditures and sold quantities. The amount and structure of product information may significantly differ among data providers. Agreeing upon delivery of 'sufficient' product information in a well-structured way is a standard agenda item during negotiations of the relations management unit of the Dutch CPI department with retail chains.

Identifying the relevant pieces of product information for the purpose of product definition is a complex problem. Depending on the quality of the data received, this may involve different steps like data cleaning and feature engineering in order to convert the data into a format suitable for methods like MARS. An illustration of the importance and effectiveness of product information for product definition, index calculation and impact analysis can be found in the recent shrinkflation study of Chessa and Barb (2025) and was also emphasised with other examples in Section 4.4. As efforts to acquire transaction data continue, processing data, assessing and improving data quality should be highly prioritised by statistical institutes.

Acknowledgments

The views expressed in this paper are those of the author(s) and do not necessarily reflect the policies of Statistics Netherlands.Parts of this document are based on research funded by various grants assigned by Eurostat, which is gratefully acknowledged. The author also wants to thank various colleagues at Statistics Netherlands for their useful comments.

References

ABS (2017). "Making greater use of transactions data to compile the Consumer Price Index." In: *15th Meeting of the Ottawa Group on Price Indices*. https://stats.unece.org/ottawagroup/download/f456.pdf. Eltville am Rhein, Germany.

Auer, L. Von (2017). "Processing scanner data by an augmented GUV index." In: *Eurostat Review on National Accounts and Macroeconomic Indicators* 2017.1. https://ec.europa.eu/eurostat/web/products-statistical-books/-/ks-gp-17-001, pp. 73–91.

Balk, B. M. (2008). *Price and quantity index numbers: Models for measuring aggregate change and difference*. Cambridge University Press.

Białek, J. (2025). "General Classes of GEKS-Type Price Indices With Application to Scanner Data." In: *The Review of Income and Wealth* 71.1. https://doi.org/10.1111/roiw.12726, e12726.

Białek, J. and A. Bobel (2019). “Comparison of price index methods for the CPI measurement using scanner data.” In: *16th Meeting of the Ottawa Group on Price Indices*. https://stats.unece.org/ottawagroup/download/f533.pdf. Rio de Janeiro, Brazil.

Chessa, A. G. (2016). “A new methodology for processing scanner data in the Dutch CPI.” In: *Eurostat Review on National Accounts and Macroeconomic Indicators* 2016.1. https://ec.europa.eu/eurostat/web/products-statistical-books/-/ks-gp-16-001, pp. 49–69.

– (2021a). “A product match adjusted R squared method for defining products with transaction data.” In: *Journal of Official Statistics* 37.2. http://dx.doi.org/10.2478/JOS-2021-0018, pp. 411–432.

– (2021b). “Extension of multilateral index series over time: Analysis and comparison of methods.” In: *Meeting of the Group of Experts on Consumer Price Indices*. https://unece.org/sites/default/files/2021-05/Session_1_Netherlands_Paper.pdf. online meeting.

– (2023). “Controlling drift when extending index series of consumer price change over time.” In: *64th ISI World Statistics Congress*. https://www.isi-next.org/media/abstracts/ottawa-2023_772473ea5ae1bbe86bed356779241963.pdf. Ottawa, Canada.

– (2025). “Extension of price index series over time for inflation measurement.” In: *10th Annual Conference of The Society for Economic Measurement*. https://virtual.oxfordabstracts.com/event/74548/submission/66. Athens, Greece.

Chessa, A. G. and F. Barb (2025). *Impact of shrinkflation in the Dutch CPI*. Tech. rep. To be uploaded on CIRCABC. The Hague: Statistics Netherlands.

Chessa, A. G., K. Link, A. M. van Huizen, and L. J. Peeters (2018). “A comparative study between the Geary–Khamis and chained Jevons method for Dutch supermarkets.” In: *Meeting of the Group of Experts on Consumer Price Indices*. https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/Netherlands_poster_ppt.pdf. Geneva, Switzerland.

Chessa, A. G., J. Verburg, and L. Willenborg (2017). “A comparison of price index methods for scanner data.” In: *15th Meeting of the Ottawa Group on Price Indices*. https://stats.unece.org/ottawagroup/download/f460.pdf. Eltville am Rhein, Germany.

Dalén, J. (2017). “Unit values and aggregation in scanner data – Towards a best practice.” In: *15th Meeting of the Ottawa Group on Price Indices*. https://stats.unece.org/ottawagroup/download/f464.pdf. Eltville am Rhein, Germany.

Diewert, W. E. and K. J. Fox (2017). “Substitution bias in multilateral methods for CPI construction using scanner data.” In: *15th Meeting of the Ottawa Group on Price Indices*. https://stats.unece.org/ottawagroup/download/f461.pdf. Eltville am Rhein, Germany.

Diewert, W. E. and C. Shimizu (2024). “Product Churn and Quality Adjustment: Using Scanner data of Laptop in Japan.” In: *18th Meeting of the Ottawa Group on Price Indices*. https://stats.unece.org/ottawagroup/download/Paper_session_3_

paper_1_cshimizu_scanner_data_product_churn_and_quality_adjustment.pdf. Ottawa, Canada.
Eltetö, O. and P. Köves (1964). "On an Index Computation Problem in International Comparisons." In: *Statiztikai Szemle* 42. In Hungarian, pp. 507–518.
Eurostat (2017). *Practical Guide for Processing Supermarket Scanner Data*. https://stats.unece.org/ottawagroup/download/f477.pdf. Luxembourg: Eurostat.
– (2020). *Practical guidelines on web scraping for the HICP*. https://ec.europa.eu/eurostat/documents/272892/12032198/Guidelines-web-scraping-HICP-11-2020.pdf. Luxembourg.
– (2022). *Guide on Multilateral Methods in the Harmonised Index of Consumer Prices*. https://ec.europa.eu/eurostat/documents/3859598/14503841/KS-GQ-21-020-EN-N.pdf. Publications Office of the European Union, Luxembourg. First edition (second edition expected end of 2026).
– (2024). *Harmonised Index of Consumer Prices (HICP): Methodological Manual*. https://ec.europa.eu/eurostat/documents/3859598/18594110/KS-GQ-24-003-EN-N.pdf. Publications Office of the European Union, Luxembourg.
Geary, R. C. (1958). "A Note on the Comparison of Exchange Rates and Purchasing Power between Countries." In: *Journal of the Royal Statistical Society A* 121. https://doi.org/10.2307/2342991, pp. 97–99.
Gini, C. (1931). "On the circular test of index numbers." In: *Metron* 2, pp. 3–24.
Grient, H. van der and J. de Haan (2010). "The use of supermarket scanner data in the Dutch CPI." In: *Joint UNECE/ILO meeting on Consumer Price Indices*. https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2010/zip.6.e.pdf. Geneva, Switzerland.
GS1 (2023). *GTIN Management Standard*. Tech. rep. https://ref.gs1.org/standards/gtin-management/. GS1.
Haan, J. de and H. A. van der Grient (2011). "Eliminating chain drift in price indexes based on scanner data." In: *Journal of Econometrics* 161. https://doi.org/10.1016/j.jeconom.2010.09.004, pp. 36–46.
ILO, IMF, OECD, UNECE, Eurostat, and World Bank (2004). *Consumer Price Index Manual: Theory and Practice*. https://www.elibrary.imf.org/display/book/9789221136996/9789221136996.xml. Geneva: ILO Publications.
Ivancic, L., W. E. Diewert, and K. J. Fox (2011). "Scanner data, time aggregation and the construction of price indexes." In: *Journal of Econometrics* 161. https://doi.org/10.1016/j.jeconom.2010.09.003, pp. 24–35.
Khamis, S. H. (1972). "A New System of Index Numbers for National and International Purposes." In: *Journal of the Royal Statistical Society A* 135. https://doi.org/10.2307/2345041, pp. 96–121.
Krsinich, F. (2014). "The FEWS index: Fixed effects with a window splice – Non-revisable quality-adjusted price indices with no characteristic information." In: *Meeting of the Group of Experts on Consumer Price Indices*. https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2014/New_Zealand_-_FEWS.pdf. Geneva, Switzerland.

Lamboray, C. (2017). “The Geary Khamis index and the Lehr index: How much do they differ?” In: *15th Meeting of the Ottawa Group on Price Indices*. https://stats.unece.org/ottawagroup/download/f462.pdf. Eltville am Rhein, Germany.

Loon, K. Van and D. Roels (2018). “Integrating big data in the Belgian CPI.” In: *Meeting of the Group of Experts on Consumer Price Indices*. https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/Belgium.pdf. Geneva, Switzerland.

Schut, C. (2001). “Using scanner data to compile price indices: Practical problems.” In: *Joint ECE/ILO Meeting on Consumer Price Indices*. https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ac.49/2001/5.e.pdf. Geneva, Switzerland.

– (2002). *Gebruik van scannerdata van supermarkten in de consumentenprijsindex*. Tech. rep. https://www.cbs.nl/nl-nl/achtergrond/2002/27/gebruik-van-scannerdata-van-supermarkten-in-de-consumentenprijsindex. Statistics Netherlands. In Dutch.

Summers, R. (1973). “International Price Comparisons Based Upon Incomplete Data.” In: *Review of Income and Wealth* 19.1. https://doi.org/10.1111/j.1475-4991.1973.tb00870.x, pp. 1–16.

Szulc, B. (1964). “Index numbers of multilateral regional comparisons.” In: *Przeglad Statystyczny* 3. In Polish, pp. 239–254.

Tardos, A. (2023). “Introduction of scanner data into Austrian CPI and HICP: Practical implementation experience, with a focus on window length options.” In: *Meeting of the Group of Experts on Consumer Price Indices*. https://unece.org/sites/default/files/2023-06/1a.2%20Austrian_CPI_and_HICP_practical_implementation_Experience_Adam_Tardos.pdf. Geneva, Switzerland.

Zhang, L.-C., I. Johansen, and R. Nygaard (2019). “Evaluating unit-value price indices in a dynamic item universe.” In: *16th Meeting of the Ottawa Group on Price Indices*. https://stats.unece.org/ottawagroup/download/f547.pdf. Rio de Janeiro, Brazil.

Appendices

A Convergence of the GK fixed-point algorithm

This appendix gives a proof of convergence of the iterative algorithm described in Section 4.5. Consider expressions (6) and (7) and omit the window $[0, T]$ from aggregate price levels and reference prices for notational convenience. Use the compact notation $\varphi_{i,z}$ for the quantity share $q_{i,z}/\sum_{s=0}^{T} q_{i,s}$ of item $i \in G_t$ in month $z \in [0, T]$.

Consider the ratio $r_t^{k+1} = P_t^{k+1}/P_t^k$ of aggregate price levels in month t obtained in iterations k and $k+1$. The two aggregate price levels are computed by first substituting the results of iterations $k-1$ and k in expression (7) for the reference prices, which are subsequently substituted in expression (6). The aforementioned ratio can thus be written as follows:

$$r_t^{k+1} = \frac{\sum_{i \in G_t} \sum_{z=0}^{T} \varphi_{i,z} \frac{p_{i,z}}{P_z^{k-1}} q_{i,t}}{\sum_{i \in G_t} \sum_{z=0}^{T} \varphi_{i,z} \frac{p_{i,z}}{P_z^{k}} q_{i,t}}. \tag{A.1}$$

By changing the order of summation in the numerator and the denominator, expression (A.1) can be rewritten in the following convenient way:

$$r_t^{k+1} = \sum_{z=0}^{T} w_{t,z}^k r_z^k, \tag{A.2}$$

$$w_{t,z}^k = \frac{\sum_{i \in G_t} \varphi_{i,z} \frac{p_{i,z}}{P_z^k} q_{i,t}}{\sum_{s=0}^{T} \sum_{i \in G_t} \varphi_{i,s} \frac{p_{i,s}}{P_s^k} q_{i,t}}, \tag{A.3}$$

where the weights $w_{t,z}^k \geq 0$ satisfy $\sum_{z=0}^{T} w_{t,z}^k = 1$. Hence each r_t^{k+1} is a convex combination of the previous ratios in $z = 0, 1, \dots, T$.

Let $M_k = \max_t r_t^k$ and $m_k = \min_t r_t^k$. By convexity, $M_{k+1} \leq M_k, m_{k+1} \geq m_k$. As a consequence, $S_k = M_k/m_k$ is non-increasing and bounded below by 1. The lower bound is the infimum of the sequence $\{S_k\}$ when normalising the aggregate price levels in each iteration according to P_t^k/P_0^k. From the monotone convergence theorem it follows that $S_k \downarrow 1$. The iterative fixed-point algorithm thus converges. The solution is unique because of the connectedness assumption stated in Section 4.3.

Colophon

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague, Netherlands
www.cbs.nl

Prepress
Statistics Netherlands Grafimedia

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70
Via contact form: www.cbs.nl/information