

von Auer, Ludwig; Shumskikh, Alena

Article — Published Version

Retrospective Computations of Price Index Numbers: Theory and Application

Review of Income and Wealth

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: von Auer, Ludwig; Shumskikh, Alena (2022) : Retrospective Computations of Price Index Numbers: Theory and Application, Review of Income and Wealth, ISSN 1475-4991, Wiley, Hoboken, NJ, Vol. 70, Iss. 1, pp. 60-79, <https://doi.org/10.1111/roiw.12627>

This Version is available at:

<https://hdl.handle.net/10419/290196>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>

RETROSPECTIVE COMPUTATIONS OF PRICE INDEX NUMBERS: THEORY AND APPLICATION

BY LUDWIG VON AUER*

Universität Trier

AND

ALENA SHUMSKIKH

Universität Trier

Due to outdated weighting information, a Laspeyres-based consumer price index (CPI) is prone to accumulating upward bias. Therefore, the present study introduces and examines simple and transparent revision approaches that retrospectively address the source of the bias. They provide a consistent long-run time series of the CPI and require no additional information. Furthermore, a coherent decomposition of the bias into the contributions of individual product groups is developed. In a case study, the approaches are applied to a Laspeyres-based CPI. The empirical results confirm the theoretical predictions. The proposed revision approaches are adoptable not only to most national CPIs but also to other price-level measures such as the producer price index or the import and export price indices.

JEL Codes: C43, E52

Keywords: CPI, decomposition, Laspeyres, revision, substitution bias

1. INTRODUCTION

Various important fields of economic analysis rely on indicators such as real gross domestic product, real wages, real interest rates, and real public debt. To obtain a long-run time series of such an indicator, the time series of the corresponding nominal indicator is deflated by some appropriate price-level measure. Furthermore, central banks need a price-level measure to anchor their monetary policy. Finally, various forms of contractual payments are indexed to such price-level measures. National statistical offices (NSOs) are entrusted to provide these price-level measures in an accurate and timely manner.

The wide range of uses necessitates a variety of national price-level measures. Probably the most important measure is the consumer price index (CPI). In their

Note: The authors presented their research at the conferences “Messung der Preise, 2022” and “Statistische Woche, 2022.” Helpful comments and suggestions from participants are gratefully acknowledged. The study benefited from the generous advice of two anonymous referees. The supportive comments that we received from Patrick Schwind, Olivia Stähl, Can Tongur, and Sebastian Weinand are greatly appreciated.

*Correspondence to: Ludwig von Auer, Fachbereich IV - VWL, Universität Trier, Universitätsring 15, 54286 Trier, Germany (vonauer@uni-trier.de).

© 2022 The Authors. *Review of Income and Wealth* published by John Wiley & Sons Ltd on behalf of International Association for Research in Income and Wealth.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

CPI compilations, NSOs relate the *comparison period's* (or current period's) price level of final consumption to the price level of some former *price reference period* (or base period). In the applied price index formulas, the weighting of the various consumption items reflects the consumer expenditures on these items. Most NSOs apply an index derived from the Laspeyres index; in other words, a Laspeyres-type index. Usually this is a Lowe index (e.g., Australia, Canada, Switzerland, and the USA) or a Young index (e.g., Denmark, Georgia, and South Africa). Only few NSOs use a proper Laspeyres index (e.g., Germany and Japan). While the weighting of a *proper* Laspeyres index exclusively depends on expenditure information of the price reference period, the weighting of a *Laspeyres-type* index may draw on expenditure information from periods other than (and usually preceding) the price reference period.

Unfortunately, the Laspeyres index and the Laspeyres-type indices are based on item weights determined by the expenditure shares of the past and, therefore, fail to incorporate the substitution behavior reflected in the consumed quantities of the comparison period (e.g., ILO *et al.* 2004, p. 4). The lower the frequency of expenditure weight updates, the more outdated are the expenditure data processed by the index. The problem is aggravated by the lag between the collection of new expenditure information and the implementation of this information into the index. The more outdated the expenditure information—and, consequently, the item weights in the index—the larger the risk of substitution bias and the associated accumulating upward bias in the long-run time series of monthly price levels. Among the more recent studies that provide empirical evidence of this bias are Lebow and Rudd (2003), Hansen (2007), Nimmo *et al.* (2007), de Haan *et al.* (2010), Greenlees and Williams (2010), Bishop and van Kints (2013), Akem and Opryshko (2014), Armknecht and Silver (2014), Huang *et al.* (2015), (2017), Klick (2018), and Herzberg *et al.* (2021).

Upward substitution bias is worrying for institutions and analysts that rely on an unbiased long-run time series of the CPI. Upward biased CPI numbers would result in downward biased real economic indicators such as real wages, real gross domestic product, and real growth rates. Biased economic indicators can lead to ill-designed economic and social policies. Problems arise also for parties involved in contracts that specify indexed payments (e.g., rent contracts) because their willingness to sign such a contract was built on false presumptions. For example, if rents are indexed to an upward biased CPI, the landlord gains at the expense of the tenant.

To reduce the risk of biased CPI numbers, many NSOs attempt to reduce the time interval between the comparison period and the period to which the item weights relate. To this end they may implement a higher frequency of collecting and processing expenditure information and/or they may reduce the lag between the collection of the expenditure information and its implementation into the index. Most OECD countries update their expenditure information every year. The exceptions include Canada (2 years), Czech Republic (2 years), Germany (5 years), and Japan (5 years). Another strategy is to wait until sufficiently up-to-date information on expenditures and, thus, on weighting is available and then to publish revised index numbers. However, in many countries such revisions raise legal issues. Furthermore, NSOs may worry that revisions undermine the credibility of their published

numbers. Therefore, many NSOs are reluctant to retrospectively revise their initial results.

Theoretically, the revised numbers can heal both, the low frequency of expenditure weight updates and the lag between the collection and implementation of new expenditure information. Therefore, it is acknowledged that retrospectively revised numbers tend to be more accurate than original ones. Even though knowledge about past CPI bias comes too late for the realignment of policies and indexed payments determined in the past, it provides additional guidance for future policies and contract negotiations. More NSOs might be willing to consider the provision of revised numbers if the extra work load appears manageable. Therefore, the compilation of the revised numbers should require only information that the NSOs collect and process anyway.

The first contribution of the present paper is to present and elaborate two general approaches of such work-saving retrospective revision methods. The two approaches are denoted as the correction approach and the imputation approach. For the implementation of both approaches, a wide range of options are available. Therefore, they are general enough to be applied by most NSOs. The approaches effectively address the source of the accumulating substitution bias, and they generate a consistent long-run time series of monthly price levels. The revised series also indicates the upward substitution bias inherent in the non-revised series.

In addition, it is shown that the imputation approach can decompose the upward substitution bias into the contributions of individual groups of items. This decomposition method is the paper's second contribution. In the future, this method may help to mitigate the bias already from the outset.

The paper's third contribution is empirical. To quantify the relevance of the problem and to illustrate the efficacy of the proposed revision approach, it is applied to the upper-level aggregation of the official German CPI, an index that is retrospectively revised, but not along the lines suggested in the present paper. It is shown that, in spite of the revision, the official German CPI numbers still suffer from upward substitution bias.

Even though the study relates to the CPI, its insights carry over to other price-level measures such as the producer price index or the import and export price indices. Furthermore, there are some implications for the Harmonized Index of Consumer Prices (HICP), which feeds into the monetary policy of the European Central Bank (ECB).

The rest of the article is organized as follows. Section 2 recaptures the relevant concepts from index theory. The correction approach and the imputation approach are introduced in Section 3. In Section 4, an empirical application of the new revision concept is presented. The decomposition of the upward bias into the contributions of individual groups of items is developed in Section 5 and applied in Section 6. Section 7 relates the article's findings to the HICP. Section 8 concludes.

2. FOUR TYPES OF SYMMETRIC PRICE INDICES

Let p_i^t and x_i^t denote the price and the quantity of item $i \in A = [1, \dots, N]$ at time period $t \in [0, \dots, T]$. The expenditures on item i are denoted by $v_i^t = p_i^t x_i^t$. The

symbols \sum and \prod are the shorthand notation for the sum and the product over all N items in set A . Only for periods 0 and T , the expenditure shares of the N items are known:

$$s_i^0 = \frac{v_i^0}{\sum v_j^0} \quad \text{and} \quad s_i^T = \frac{v_i^T}{\sum v_j^T} .$$

For each comparison period $t \in [0, \dots, T]$, a Laspeyres index can be compiled that compares the average price level of the comparison period t to the average price level of the price reference period 0. The resulting sequence of Laspeyres indices, $P_L^{0 \rightarrow t}$, is

$$(1) \quad P_L^{0 \rightarrow 0} = \sum s_i^0 \frac{p_i^0}{p_i^0} = 1, \quad P_L^{0 \rightarrow 1} = \sum s_i^0 \frac{p_i^1}{p_i^0}, \quad \dots \quad , \quad P_L^{0 \rightarrow T} = \sum s_i^0 \frac{p_i^T}{p_i^0} .$$

In this study, the period that provides the information for the calculation of the expenditure weights is denoted as the *expenditure reference period*. In the sequence of Laspeyres indices (1), period 0 is both, the expenditure reference period and the price reference period.¹

Once the expenditure shares of the new expenditure reference period T become available, a retrospective price index can be compiled for each comparison period $t \in [0, \dots, T]$. Possible candidates for such retrospective price indices are appropriately modified versions of the so-called *symmetric* standard price index formulas. ILO *et al.* (2004, pp. 5-6) attach the label “symmetric” to price index formulas that give equal importance to the information of the first expenditure reference period 0 and the second expenditure reference period T . However, the phrase “equal importance” turns out to be rather ambiguous because it can come in different forms. Therefore, we propose to distinguish between four types of symmetry.

A. Symmetric Treatment of Quantities

The quantities, x_i , represent some combination of $x_i^0 = v_i^0/p_i^0$ and $x_i^T = v_i^T/p_i^T$. Examples are:

$$(2) \quad \begin{aligned} \text{Marshall-Edgeworth : } P_{ME}^{0 \rightarrow T} &= \frac{\sum p_i^T (x_i^0 + x_i^T)}{\sum p_i^0 (x_i^0 + x_i^T)} \\ \text{Walsh : } P_W^{0 \rightarrow T} &= \frac{\sum p_i^T \sqrt{x_i^0 x_i^T}}{\sum p_i^0 \sqrt{x_i^0 x_i^T}} . \end{aligned}$$

¹The expenditure reference period should not be confused with the *weight reference period* as defined in ILO *et al.* (2004, p. 3). The latter is the period whose *quantities* are actually used in the index. In a Laspeyres index of the form (1), the weight reference period and the expenditure reference period coincide. Another type of reference period is the *index reference period*. This is the period for which the index is set equal to 100.

B. Symmetric Treatment of Expenditures

The expenditures, v_i , represent some combination of v_i^0 and v_i^T . Well-known examples are:

$$\text{Theil : } \ln P_{\text{Th}}^{0 \rightarrow T} = \sum \left[\frac{\sqrt[3]{\frac{1}{2} (v_i^0 + v_i^T) v_i^0 v_i^T}}{\sum \sqrt[3]{\frac{1}{2} (v_j^0 + v_j^T) v_j^0 v_j^T}} \right] \ln \frac{p_i^T}{p_i^0}$$

$$\text{Vartia : } \ln P_{\text{Va}}^{0 \rightarrow T} = \sum \frac{L(v_i^0, v_i^T)}{L(\sum v_j^0, \sum v_j^T)} \ln \frac{p_i^T}{p_i^0},$$

$$\text{where } L(a, b) = \frac{a - b}{\ln a - \ln b} \text{ for } a \neq b \text{ and } L(a, b) = a \text{ for } a = b.$$

C. Symmetric Treatment of Expenditure Shares

The expenditure shares, s_i , represent some combination of s_i^0 and s_i^T . Indices with this type of symmetry include:

$$(3) \quad \text{Törnqvist : } \ln P_{\text{Tö}}^{0 \rightarrow T} = \sum \frac{1}{2} (s_i^0 + s_i^T) \ln \frac{p_i^T}{p_i^0}$$

$$\text{Sato-Vartia : } \ln P_{\text{SV}}^{0 \rightarrow T} = \sum \frac{L(s_i^0, s_i^T)}{\sum L(s_j^0, s_j^T)} \ln \frac{p_i^T}{p_i^0}.$$

D. Symmetric Treatment of Indices

The indices, $P^{0 \rightarrow T}$, represent some combination of an index with expenditure reference period 0 and some index with expenditure reference period T . For this purpose, the Laspeyres index, $P_L^{0 \rightarrow T}$, and the Paasche index,

$$P_P^{0 \rightarrow T} = \frac{\sum p_i^T x_i^T}{\sum p_i^0 x_i^T} = \left[\sum s_i^T \left(\frac{p_i^T}{p_i^0} \right)^{-1} \right]^{-1},$$

are particularly popular. The best known example is the index of

$$(4) \quad \text{Fisher : } P_F^{0 \rightarrow T} = \sqrt{P_L^{0 \rightarrow T} P_P^{0 \rightarrow T}}.$$

Some price index formulas can be assigned to more than one type of symmetry. For example, the Walsh index can also be written in the forms

$$(5) \quad P_W^{0 \rightarrow T} = \sum \frac{\sqrt{v_i^0 v_i^T (p_i^T / p_i^0)^{-1}}}{\sum \sqrt{v_j^0 v_j^T (p_j^T / p_j^0)^{-1}}} \frac{p_i^T}{p_i^0}$$

$$(6) \quad = \sum \frac{\sqrt{s_i^0 s_i^T (p_i^T / p_i^0)^{-1}}}{\sum \sqrt{s_j^0 s_j^T (p_j^T / p_j^0)^{-1}}} \frac{p_i^T}{p_i^0}.$$

This is a weighted arithmetic mean of the price ratios (p_i^T / p_i^0) where the weights depend not only on the expenditures (or expenditure shares) of the price reference period, but also on the expenditures (or expenditure shares) of the comparison period T deflated to the price reference period. Therefore, the Walsh index can be interpreted as a representative of symmetry types A, B, and C.

It is not possible to compute symmetric price indices in real time because the expenditures related to period T become available only with a considerable lag. However, a retrospective computation of the listed price indices (and several other symmetric price indices) is feasible and suitable. The authors of ILO *et al.* (2004, p. 173) advocate the superlative price indices (2), (3), or (4). However, also the other price indices have their specific strengths (e.g., Sato-Vartia and Marshall-Edgeworth) and should be included in the short list.

All the listed symmetric price indices compare the prices of the comparison period T to those of the price reference period 0. However, we need a price index formula that can relate not only period T but also all intermediate comparison periods $t \in [1, \dots, T - 1]$ to the price reference period 0. What can be achieved if neither the quantities, nor the expenditures, nor the expenditure shares of these intermediate comparison periods are known? The objective is a price index formula that provides a more reliable series of retrospective index numbers than the Laspeyres-based series (1).

When the quantity and expenditure information of the intermediate periods ($t = 1, \dots, T - 1$) is completely missing, we must make use of the available quantity or expenditure information of the two expenditure reference periods, $t = 0$ and $t = T$, and we must invoke some form of interpolation for the intermediate periods.

To this end, Diewert *et al.* (2009, p. 129) propose a retrospective price index formula that relates to symmetry type D (symmetry of indices):

$$(7) \quad P_{\text{DHK}}^{0 \rightarrow t} = \left(\frac{\sum p_i^t x_i^0 \sum p_i^t x_i^T}{\sum p_i^0 x_i^0 \sum p_i^0 x_i^T} \right)^{1/2}.$$

This is the geometric average of a Laspeyres index and a Lowe index that uses the quantities of period T . The subscript “DHK” stands for the authors Diewert, Huwiler, and Kohli. For period $t = T$, the DHK index (7) simplifies to the Fisher index (4) and for period $t = 0$ the index yields $P_{\text{DHK}}^{0 \rightarrow 0} = 1$. Inspired by Hansen (2007, p. 12), Diewert *et al.* (2009, p. 131) also propose the following formula:

$$(8) \quad P_{\text{DHH}}^{0 \rightarrow t} = \left[\sum s_i^0 \left(\frac{p_i^t}{p_i^0} \right) \right]^{1/2} \left[\sum s_i^T \left(\frac{p_i^t}{p_i^0} \right)^{-1} \right]^{-1/2},$$

where the additional “H” in the subscript stands for Hansen. This index is a geometric average of a Laspeyres index and a Paasche-type index which uses the expenditure weights of period T instead of period t . Again, for $t = T$, the DHKH index (8) simplifies to the Fisher index and for period $t = 0$ the index yields $P_{\text{DHKH}}^{0 \rightarrow 0} = 1$.

However, formulas (7) and (8) have a common issue. For example, consider the price comparison of the neighboring periods 0 and 1. In the sequence of Laspeyres indices (1), the index $P_L^{0 \rightarrow 1}$ uses quantity information of period 0 only. Since period 1 is the neighboring period, it is unlikely that a significant change in the quantities occurs between periods 0 and 1. Therefore, the substitution bias of the Laspeyres index, $P_L^{0 \rightarrow 1}$, is likely to be very small.

By contrast, the indices $P_{\text{DHK}}^{0 \rightarrow 1}$ and $P_{\text{DHKH}}^{0 \rightarrow 1}$ incorporate not only the quantities of period 0 but also the quantities of the last period, T . If periods 1 and T are far apart, the quantities of period 0 usually represent a much better approximation to the quantities of period 1 than do the quantities of period T . In other words, $P_L^{0 \rightarrow 1}$ is more reliable than $P_{\text{DHK}}^{0 \rightarrow 1}$ and $P_{\text{DHKH}}^{0 \rightarrow 1}$. The retrospective correction causes more harm than good.

Of course, as time progresses, the reliability of the Laspeyres index, $P_L^{0 \rightarrow t}$, erodes, while the reliability of a retrospective index improves until in period T the retrospective index reaches its full advantage over the Laspeyres index. Therefore, the basic concept of constructing a retrospective price index should be preserved, but it must be refined such that only in the final period T the full extent of retrospective correction is reached, while in the initial periods hardly any retrospective correction arises.

The following section introduces a typology of such refined approaches. The typology reveals that a wide range of options is available to price statisticians and that only very few of them have been applied in the price index literature.

3. RETROSPECTIVELY COMPUTED PRICE INDICES

The construction of a retrospective price index formula can follow at least two alternative principles. The index can be defined as

- the product of the original Laspeyres index (or Laspeyres-type index) and some *correction* factor that accounts for the gradually increasing deviation between the Laspeyres index and a symmetric price index as t progresses from 0 to T , or as
- a symmetric price index formula that *imputes* the missing information on x_i^t , y_i^t , or s_i^t ($t = 1, \dots, T - 1$) where the imputed values are weighted averages of the values of periods 0 and T , and where the weight of the value of period 0 gradually decreases from 1 to 0 as t progresses from 0 to T .

We denote the two principles as *correction approach* and *imputation approach*. The two approaches can be implemented in various ways. In both approaches, the variable $\lambda_t = t/T$ plays a crucial role. This variable reflects the relevance of the second expenditure reference period ($t = T$) relative to the relevance of the first expenditure period ($t = 0$). As time progresses, λ_t increases from $\lambda_0 = 0$ to $\lambda_T = 1$, reflecting the gradually diminishing relevance of period $t = 0$ and the increasing relevance of period $t = T$.

Correction Approach: Using the Walsh index (2), the series of Laspeyres indices (1) can be revised by the following retrospective price index formula:

$$(9) \quad P^{0 \rightarrow t} = (P_L^{0 \rightarrow t}) \left(\frac{P_W^{0 \rightarrow T}}{P_L^{0 \rightarrow T}} \right)^{\lambda_t}, \quad t = 0, \dots, T.$$

For $t = 0$, we have $\lambda_0 = 0$. Therefore, the correction factor $(P_W^{0 \rightarrow T} / P_L^{0 \rightarrow T})^{\lambda_t}$ is equal to 1 (i.e., no correction), and the retrospective price index formula gives $P^{0 \rightarrow 0} = P_L^{0 \rightarrow 0} = 1$. As time progresses toward period $t = T$, λ_t increases from 0 to 1, the correction factor reaches its maximum deviation from 1, and the retrospective price index gradually turns into the Walsh index: $P^{0 \rightarrow T} = P_W^{0 \rightarrow T}$.

Instead of the Walsh index, other symmetric price index formulas could be used to define the correction factor. For example, Auer and Shumskikh (2022) apply the Törnqvist index (3). Using the Fisher index (4) gives:

$$(10) \quad P^{0 \rightarrow t} = (P_L^{0 \rightarrow t}) \left(\frac{P_P^{0 \rightarrow T}}{P_L^{0 \rightarrow T}} \right)^{\lambda_t/2}, \quad t = 0, \dots, T,$$

and, therefore, $P^{0 \rightarrow 0} = 1$ and $P^{0 \rightarrow T} = P_F^{0 \rightarrow T}$. The same two values are generated by a retrospective price index formula that is proposed in de Haan *et al.* (2010, p. 30). They replace the correction factor of formula (10) by $(P_{Lo}^{0 \rightarrow t} / P_L^{0 \rightarrow t})^{\lambda_t/2}$, where the Lowe index $P_{Lo}^{0 \rightarrow t}$ is defined as $P_{Lo}^{0 \rightarrow t} = (\sum p_i^t x_i^T) / (\sum p_i^0 x_i^T)$.

Instead of the multiplicative correction of the Laspeyres index, $P_L^{0 \rightarrow t}$, applied in (9), one could use an additive one:

$$(11) \quad P^{0 \rightarrow t} = (P_L^{0 \rightarrow t}) + \lambda_t (P_W^{0 \rightarrow T} - P_L^{0 \rightarrow T}), \quad t = 0, \dots, T.$$

Again, one obtains $P^{0 \rightarrow 0} = 1$ and $P^{0 \rightarrow T} = P_W^{0 \rightarrow T}$.

Imputation Approach: When the Walsh index (2) is applied, the unknown quantities, x_i^t ($t = 1, \dots, T - 1$), must be substituted by imputed values, \hat{x}_i^t . If we assume that the rate of change of x_i^t is constant between periods 0 and T , the imputed values are given by the weighted geometric average

$$(12) \quad \hat{x}_i^t = (x_i^0)^{1-\lambda_t} (x_i^T)^{\lambda_t},$$

and we obtain $P^{0 \rightarrow 0} = 1$ and $P^{0 \rightarrow T} = P_W^{0 \rightarrow T}$. Alternatively, other symmetric price indices can be used. Furthermore, the imputed values, \hat{x}_i^t , can be obtained from a weighted arithmetic average:

$$(13) \quad \hat{x}_i^t = (1 - \lambda_t)x_i^0 + \lambda_t x_i^T.$$

The same options are available, if the missing values of the expenditures, v_i^t , or the expenditure shares, s_i^t , have to be imputed.²

²A related interpolation problem arises in the construction of the international purchasing-power-parity (PPP) exchange rates compiled by the International Comparison Program (ICP) and published in the Penn World Tables. These PPPs are constructed for each basic heading

An early application of the imputation approach can be found in de Haan *et al.* (2010, p. 29). They propose to use the Törnqvist index (3) or the Fisher index (4) and to impute the unknown expenditure shares, s_i^t ($t = 1, \dots, T - 1$), by $\hat{s}_i^t = (1 - \lambda_t)s_i^0 + \lambda_t s_i^T$.

In this section, we have presented two different construction principles for a retrospective price index: the correction approach and the imputation approach. For both approaches, a large set of symmetric price indices is available. In the correction approach, a multiplicative or an additive form of correction can be applied. In the imputation approach, the imputed values can be obtained from weighted arithmetic or geometric averages. Since $\lambda_T = 1$, the index number relating to the comparison period $t = T$ (the second expenditure reference period) is independent from the choice between the correction and the imputation approach.

Both, the correction approach and the imputation approach, have intuitive appeal. The correction approach appears somewhat less intrusive than the imputation approach. Therefore, in the following section, we show how the correction approach can be implemented in the German CPI. By contrast, the imputation approach allows for an additive decomposition of the overall price change. This property is useful in various ways. For example, it helps to identify those groups of items that cause the upward substitution bias of the Laspeyres index. This identification is a prerequisite for reducing the bias already from the outset. Therefore, in Section 6 the imputation approach is applied to identify the critical items in the German CPI.

4. RETROSPECTIVE CORRECTION OF THE GERMAN CPI

Destatis (shortened form of “Statistisches Bundesamt”) compiles proper Laspeyres indices. The most recent price and expenditure reference period of Destatis is the year 2015. The previous ones were the years 2010, 2005, and so on. For example, the Laspeyres index of the comparison period January 2000 was published in February 2000 and used the year 1995 as price and expenditure reference period. The same price and expenditure reference period was used for the subsequent months.

When in 2003 the expenditures of the year 2000 became available, Destatis recalculated the index numbers of the comparison periods starting in January 2000. The revised Laspeyres indices use the year 2000 as the price and expenditure reference period. These index calculations were continued until in 2008 the expenditure

for each benchmark year, that is, years for which expenditure surveys were conducted. The latest benchmark years are 1996, 2005, 2011, and 2017. Interpolations are provided for the years between the benchmark years. To this end, Feenstra *et al.* (2013, pp. 17-20) take a weighted arithmetic average of the PPPs of the two benchmark years where the weight is constructed in the same way as λ_t . That is, as time progresses the impact shifts from the older benchmark year toward the newer benchmark year. Rao *et al.* (2018, pp. 6-10) recommend a modification of that approach and discuss various alternative options that have been published in the literature. These options include the econometric approach of Rao *et al.* (2010, pp. S68-S69) and the approach of Diewert and Fox (2015, pp. 26-31).

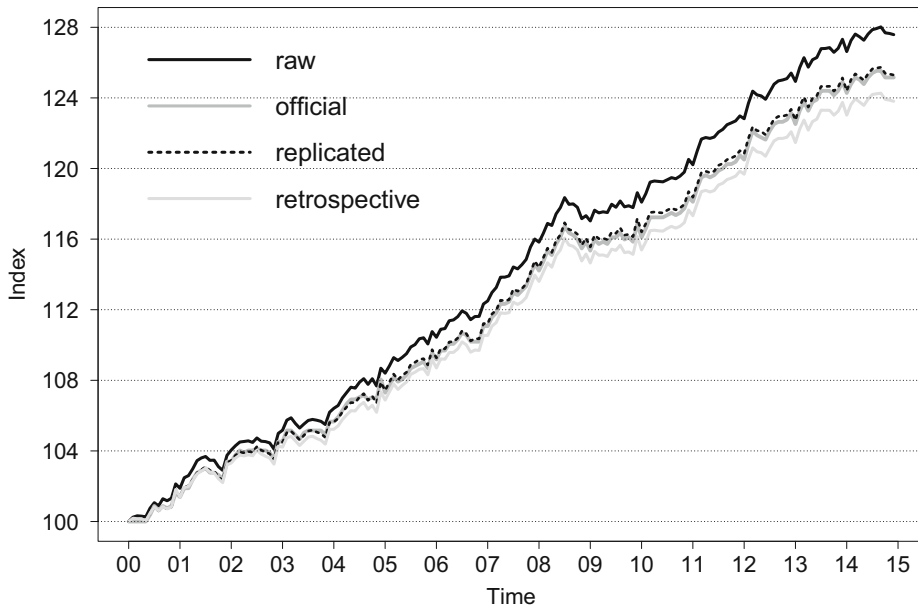


Figure 1. German Consumer Price Index (January 2000 = 100) for January 2000 to December 2014.
Source: Own calculations based on data of Destatis.

weights of the year 2005 became available. This new information prompted a retrospective revision of the Laspeyres index numbers starting in January 2005. Analogous retrospective revisions occurred in 2013 when the expenditures of the year 2010 became available.

In total, these compilations generate a sequence of three consecutive time series of monthly Laspeyres indices, each covering a 5-year period. These three series are chained, such that a consistent time series starting in January 2000 is obtained. The details of the official CPI compilation procedure are documented in Appendix A. The official source is Statistisches Bundesamt (2018, pp. 6-7, 15-19).³ The official time series of monthly price levels is depicted in Figure 1 (dark-gray line labeled as “official”).

We cannot use the official time series as a reference because we do not have access to the same data set. The data set available to us is described in Appendix A. Using that data set, we replicate the compilation procedure of Destatis. The resulting time series will serve as a reference for the time series derived from the retrospective index formula. The reference time series is depicted in Figure 1 by the black dotted line labeled as “replicated.” The differences between the officially published Laspeyres CPI index and our replicated Laspeyres index are hardly visible. In the final comparison period (December 2014), the deviation between the official and the replicated index number is merely 0.14 percentage points. This is a

³Additional descriptions are provided by Egner (2003), Elbel and Egner (2008), Egner (2013), and Egner (2019).

remarkable result because the information accessible for our research purposes is not as granular as the information processed by Destatis (see the data description in Appendix A).

Now, if Destatis had ruled out retrospective revisions, in 2003 it would have continued with the old price and expenditure reference period (1995) until January 2005, the month scheduled for the introduction of regular methodological modifications and other updates of the index. Thus, January 2005 would have been the time to switch from the old (1995) to the new (2000) price and expenditure reference period. Subsequently, an analogous changeover would have been conducted in January 2010.

Based on the data set available to us, we simulate this non-revisionary compilation procedure. The details are documented in Appendix A. In Figure 1, the black line labeled as “raw” depicts the resulting time series of unrevised monthly price levels. It covers the time interval January 2000 to December 2014 (the index numbers of earlier months would require the expenditure information of the year 1990 which we do not have). This time series is obtained from Laspeyres index numbers based on weights that are outdated by 5–10 years. As a result, it is prone to substantial upward substitution bias.

The comparison of the two time series “raw” and “replicated” reveals that the retrospective revisions of Destatis successfully curbed the long-run upward substitution bias. Did these revisions even eliminate the bias? To answer this question, a time series of index numbers is required that can be considered as free of long-run substitution bias.

Therefore, we compile a time series of retrospective Walsh indices and compare it with the time series of replicated Laspeyres indices. The Walsh-based series applies the correction approach. The comparison between the Laspeyres- and Walsh-based series gives us clues about the existence and the extent of the remaining distortion in the long-run time series of the official CPI. The retrospective Walsh index numbers of the time interval 2015–2020 would require the expenditure weights of the year 2020. These will become available not before 2023. Therefore, we decided to restrict the analysis to the interval January 2000 to December 2014. The formal details of the compilation procedure of the retrospective Walsh index are explained in Appendix B.

The result of this procedure is the light gray line in Figure 1 labeled as “retrospective.” The graph confirms the theoretical predictions. The retrospective Walsh index runs below the replicated Laspeyres index, indicating that the official index still suffers from upward substitution bias. The deviation accumulates over time and in December 2014 it reaches almost 1.5 percentage points. This deviation represents only the distortion that can be attributed to the upper-level aggregation, that is, the aggregation of the 102 classes into the overall CPI. Since the subclasses at the lower-level aggregation exhibit a larger substitutability than the classes at the upper-level aggregation, it is quite likely that the actual upward bias is larger than the 1.5 percentage points observed here.

On a per-year basis, the estimated upper-level substitution bias of the replicated Laspeyres index is 0.087 percentage points each year. How does this number compare to the upper-level substitution bias estimated for the CPI-U, the headline CPI compiled by the US Bureau of Labor Statistics? Their expenditure information is

updated every 2 years, and the expenditure shares are drawn from two subsequent calendar year expenditures. For example, beginning in January 2022, the CPI-U index numbers are based on expenditure information from the years 2019 and 2020. This expenditure information is price-updated to the price reference period. Therefore, the CPI-U is a Lowe index instead of a Laspeyres index.

Greenlees and Williams (2010, p. 749) found a yearly deviation of 0.316 percentage points between the CPI-U and a retrospectively computed yearly chained Törnqvist index for the time interval 1999 to 2007. This estimate is quite in line with the estimate of Lebow and Rudd (2003, p. 162) for the time interval 1987–2000. Similarly, Klick (2018, p. 4) found a yearly deviation of 0.297 percentage points between the CPI-U and the retrospectively computed C-CPI-U (a monthly chained Törnqvist index) for the time interval 2000–2012, while Armknecht and Silver (2014, p. 240) obtained a yearly deviation of 0.159 for the time interval 2002–2010.

5. DECOMPOSITION OF THE BIAS

It would be valuable to know which groups of items are responsible for the upward substitution bias of the Laspeyres index. Usually the CPI is formed by major expenditure categories where each category comprises several items. More formally, the set of items, A , can be partitioned into the subsets A_k where $k \in K$ and K is the set of (major expenditure) categories.

We know from formulas (1) and (6) that both the Laspeyres index and the Walsh index can be expressed in the additive form

$$(14) \quad P = \sum_{i \in A} z_i \frac{p_i^T}{p_i^0},$$

where p_i^T/p_i^0 is the price ratio of item i and, therefore, the *primary attribute* of the index and z_i is the *secondary attribute*. It can be interpreted as the weight of the primary attribute. In this section, we drop the superscript “0 → T” at the index P . Since $\sum_{i \in A} = \sum_{k \in K} \sum_{i \in A_k}$, index formula (14) can also be written in the form

$$(15) \quad P = \sum_{k \in K} Z_k P_k,$$

where $P_k = \sum_{i \in A_k} (z_i/Z_k)(p_i^T/p_i^0)$ is the price index computed for category k and $Z_k = \sum_{i \in A_k} z_i$ is the weight of category k , with

$$(16) \quad z_i = z_{i,L} = s_i^0 \quad (\text{Laspeyres index})$$

$$(17) \quad z_i = z_{i,W} = \frac{\sqrt{s_i^0 s_i^T (p_i^T/p_i^0)^{-1}}}{\sum_j \sqrt{s_j^0 s_j^T (p_j^T/p_j^0)^{-1}}} \quad (\text{Walsh index}).$$

Equations (14) and (15) imply that the same index number is obtained, regardless of whether the index is computed in a single stage over all items in set A or in two stages, where, on the first stage, for each subset A_k the price index P_k and the aggregated weight Z_k are computed and, on the second stage, these values are used to compute the overall result P . In that second stage, the same index formula is used as in the first stage and the single stage computation. Z_k and P_k simply replace z_i and p_i^t/p_i^0 , respectively. Therefore, Auer and Wengenroth (2021) denote the Laspeyres and the Walsh index as *consistent in aggregation with respect to the secondary attributes* (16) and (17), respectively. The authors also show that all other indices listed in Section 2 of the present study are consistent in aggregation with respect to some appropriately defined secondary attribute (e.g., Törnqvist index, Fisher index).⁴

Usually, the property of consistency in aggregation would be used to decompose the overall price change into the contributions of the individual categories. However, this property also allows for the decomposition of the deviation between two different index formulas into the contributions of the individual categories. Since we interpret the deviation between the Laspeyres index and the Walsh index as the upward substitution bias of the Laspeyres index, we obtain a decomposition of that bias.

The categories' Z_k -values can be considered as the weight of the category in the overall index computation. Definitions (16) and (17) imply that the notation must distinguish between the Z_k -values of the Laspeyres index, $Z_{k,L} = \sum_{i \in A_k} z_{i,L}$, and the Z_k -values of the Walsh index, $Z_{k,W} = \sum_{i \in A_k} z_{i,W}$. Noting that $\sum_{k \in K} Z_{k,L} = \sum_{k \in K} Z_{k,W} = 1$ and that the Laspeyres index and Walsh index are consistent in aggregation with respect to the secondary attributes (16) and (17), respectively, we propose the following decomposition of the overall bias:

$$(18) \quad P_L - P_W = \sum_{k \in K} Z_{k,L} P_{k,L} - \sum_{k \in K} Z_{k,W} P_{k,W} = \sum_{k \in K} [Z_{k,L} (P_{k,L} - 1) - Z_{k,W} (P_{k,W} - 1)],$$

where

$$P_{k,L} = \sum_{i \in A_k} z_{i,L} \frac{p_i^T}{p_i^0} \quad \text{and} \quad P_{k,W} = \sum_{i \in A_k} z_{i,W} \frac{p_i^T}{p_i^0}.$$

This decomposition would work for any pair of price indices that are consistent in aggregation with respect to a single secondary attribute.

6. DECOMPOSITION OF THE BIAS IN THE GERMAN CPI

The German CPI is formed by 12 major expenditure categories, $K = [1, \dots, 12]$. The names of these 12 categories are listed in the first column of

⁴The definition of “consistency in aggregation with respect to some secondary attribute” (Auer and Wengenroth 2021, p. 602) relies on the authors’ notion of consistent aggregation rules (p. 593). In Proposition 2 (p. 594), they show that quasi-additive aggregation rules and, thus, also additive aggregation rules are always consistent aggregation rules. The Fisher index and the Törnqvist index can be expressed as quasi-additive aggregation rules.

TABLE 1

LASPEYRES INDEX NUMBERS AND WALSH INDEX NUMBERS FOR DECEMBER 2014 ($P_{k,L}$ AND $P_{k,W}$; JANUARY 2010 = 100), WEIGHTS ($Z_{k,L}$ AND $Z_{k,W}$; IN PERCENT), AND CONTRIBUTIONS TO THE BIAS OF THE 12 MAJOR EXPENDITURE CATEGORIES OF THE GERMAN CPI

Expenditure Category	Laspeyres		Walsh		Contrib. to Bias
	$P_{k,L}$	$Z_{k,L}$	$P_{k,W}$	$Z_{k,W}$	
Food, non-alcoholic beverages	111.87	10.45	111.74	10.09	0.056
Alcoholic beverages, tobacco	111.78	3.81	111.70	3.78	0.007
Clothing, footwear	109.89	4.49	109.84	4.51	0.000
Housing, water, electricity, gas, other fuels	109.04	32.28	108.88	31.92	0.085
Furniture, other household equipment	102.83	5.06	102.80	5.24	-0.003
Health	102.58	4.52	101.61	4.80	0.039
Transport	105.91	13.71	105.78	13.65	0.021
Communication	90.67	3.06	90.04	3.19	0.032
Recreation, entertainment, culture	110.88	11.69	109.36	11.61	0.186
Education	70.29	0.40	70.29	0.37	-0.008
Restaurant, accommodation services	109.62	4.54	109.63	4.66	-0.012
Miscellaneous goods and services	105.42	5.97	105.13	6.17	0.007
Total (January 2010 to December 2014)	107.75	100.00	107.34	100.00	0.410

Table 1. The decomposition (18) allows us to examine whether the bias in the Laspeyres index is concentrated in only a few of these categories. We start with the comparison period December 2014 ($t = 12/14$) and the price reference period January 2010 ($t = 1/10$). In the notation, we drop the superscript “1/10 → 12/14” at the price index. For the retrospective computation of the Walsh index, the imputation approach is applied. The computational details are described in Appendix C. The same type of analysis is repeated for the time intervals January 2000 to December 2004 as well as January 2005 to December 2009.

The bottom line of Table 1 shows the all-items index numbers for the price reference period January 2010 and the comparison period December 2014. The all-items Laspeyres index (1) gives the number $P_L = 107.75$, whereas the retrospectively computed Walsh index yields $P_W = 107.34$. Thus, the upward bias of the all-items Laspeyres index is 0.410. The index numbers of each category are also listed in Table 1. The last column decomposes the all-items bias (0.410) into the contributions of the 12 individual categories. To this end, we use the decomposition approach presented in formula (18), where

$$P_{k,L} = \sum_{i \in A} s_i^{10} \frac{p_i^{12/14}}{p_i^{1/10}} \quad \text{and} \quad P_{k,W} = \sum_{i \in A} \frac{\sqrt{s_i^{10} s_i^{15} (p_i^{15}/p_i^{10})^{-1}}}{\sum_{j \in A} \sqrt{s_j^{10} s_j^{15} (p_j^{15}/p_j^{10})^{-1}}} p_i^{12/14} p_i^{1/10}$$

$$Z_{k,L} = \sum_{i \in A_k} s_i^{10} \quad \text{and} \quad Z_{k,W} = \sum_{i \in A_k} \frac{\sqrt{s_i^{10} s_i^{15} (p_i^{15}/p_i^{10})^{-1}}}{\sum_{j \in A_k} \sqrt{s_j^{10} s_j^{15} (p_j^{15}/p_j^{10})^{-1}}}.$$

For example, the contribution of the tenth category, “Education”, is negative (-0.008), even though the price indices of that category are identical

($P_{10,L} = P_{10,W} = 70.29$). The negative contribution says that the category “Education” *reduces* the upward bias inherent in the overall Laspeyres index. The cause of that reduction is the difference in the weights. In the Laspeyres index, the price decline receives the weight $Z_{10,L} = 0.40$, while in the Walsh index the weight is only $Z_{10,W-im} = 0.37$. The category “Education” reveals that a comparison of the Laspeyres index and Walsh index of some given category is not sufficient to evaluate that category’s contribution to the overall bias. Also the weights must be considered. Besides “Education”, two other categories mitigate the bias. The strongest positive contribution to the bias comes from the category “Recreation, entertainment, and culture”.

The decomposition of the bias can be conducted also for the intervals January 2000 to December 2004 as well as January 2005 to December 2009. The results are documented in Appendix D. The overall bias for January 2000 to December 2004 and for January 2005 to December 2009 are 0.683 and 0.573, respectively. Looking over all three time intervals, one can say that roughly half of the bias can be attributed to the category “Recreation, entertainment, and culture”. For the compilers of the German CPI, this is a valuable insight. It may prompt a closer inspection of the computational procedures applied in this category. Even though the category “Housing, water, electricity, gas, other fuels” has by far the highest weight, its contribution to the bias is small. The contributions of the categories “Clothing, footwear,” “Education,” “Restaurant, accommodation services,” and “Miscellaneous goods and services” are negligible.

7. IMPLICATIONS FOR THE HICP

The national CPIs are the relevant benchmark indices for wage bargaining as well as the indexation of rents, taxes, social benefits, alimony, and other contractual payments. Furthermore, the national CPIs provide the basic input for the HICP which is the relevant yardstick for the monetary policy of the Governing Council of the ECB. Measurement problems in the CPIs are likely to feed through into the HICP. Therefore, not only the national CPIs but also the HICP should be scrutinized with respect to possible sources of measurement bias.

In contrast to the German CPI, which is a revisable fixed-base index, the HICP is an annually chain-linked index that is generally not revised (an exception was made in 2019).⁵ For example, suppose that the price level for month 6/22 (short for June 2022) is to be published and that the index reference period (period where the index is equal to 100) is month 12/15. The HICP methodology would solve this task by the following chain index:

$$\tilde{P}_{\text{HICP}}^{12/15 \rightarrow 6/22} = 100 \cdot P_{\text{HICP}}^{12/15 \rightarrow 12/16} \cdot P_{\text{HICP}}^{12/16 \rightarrow 12/17} \cdot \dots \cdot P_{\text{HICP}}^{12/20 \rightarrow 12/21} \cdot P_{\text{HICP}}^{12/21 \rightarrow 6/22}$$

⁵Another difference is the scope of the items included. For example, the HICP excludes owner-occupied housing, a category that is integrated in the German CPI by the rental equivalence approach.

If the individual links ($P_{\text{HICP}}^{0 \rightarrow t}$) of this chain index used up-to-date expenditure shares, the issue of substitution bias would be much less relevant than in the German CPI with its expenditure shares that are outdated by several years. However, such up-to-date information is not available. Therefore, the up-to-date expenditure shares must be estimated. This is accomplished by some form of updating outdated expenditure information.

Before 2012, the links of the HICP were conceptualized as proper Lowe indices:

$$P_{\text{Lo}}^{0 \rightarrow t} = \frac{\sum p_i^t x_i^b}{\sum p_i^0 x_i^b},$$

where the benchmark year b (year for which a full-scale household expenditure survey was conducted) precedes the price reference period 0. Writing this index in share form yields

$$(19) \quad P_{\text{Lo}}^{0 \rightarrow t} = \sum s_i^{b \uparrow 0} \frac{p_i^t}{p_i^0},$$

where the arrow \uparrow indicates that the shares, $s_i^{b \uparrow 0}$, are obtained by price-updating the observed expenditures of the benchmark year b to the price reference period 0:

$$(20) \quad s_i^{b \uparrow 0} = \frac{v_i^b (p_i^0 / p_i^b)}{\sum v_j^b (p_j^0 / p_j^b)} = \frac{s_i^b (p_i^0 / p_i^b)}{\sum s_j^b (p_j^0 / p_j^b)}.$$

The comparison period t is the month preceding the month in which the index number is published. The price reference period 0 is December of the previous year. In the German HICP, the benchmark year b is a year that precedes the year of the price reference month 0 by 5–7 years. Formula (20) price-updates the expenditures v_i^b from the benchmark year b to the price reference month 0. To this end, the expenditures are multiplied by the ratio (p_i^0 / p_i^b) .

Economic considerations suggest that such price-updating tends to generate index numbers that are even more upward biased than Laspeyres index numbers (e.g., ILO *et al.* (2004, pp. 272-273). Greenlees and Williams (2010, p. 749), as well as Armknecht and Silver (2014, p. 236) empirically confirm this conjecture for data provided by the US Bureau of Labor Statistics). Huang *et al.* (2015, p. 16) report such evidence for data of Statistics Canada.

Herzberg *et al.* (2021) examine German HICP data. They compute an experimental HICP that uses the pre-2012 HICP methodology ($P_{\text{HICP}}^{\text{pre}2012}$). To evaluate the performance of this index, they compute its deviation from an experimental benchmark index that they consider as the measure of “true” inflation. This index is a retrospectively computed annually chain-linked Törnqvist index ($P_{\text{T}}^{\text{revised}}$) with the following links:

$$(21) \quad P_{\text{T}}^{0 \rightarrow t} = \sum \frac{1}{2} (\hat{s}_i^0 + \hat{s}_i^t) \ln \frac{p_i^t}{p_i^0}.$$

The expenditure shares \hat{s}_i^0 and \hat{s}_i^t must be estimated. This process draws on price data, on the expenditure shares from the benchmark years, and on expenditure information from the *final revision* of the national accounts (for details see Appendix E). Herzberg *et al.* (2021) show that the yearly averages of the monthly ratios $P_{\text{HICP}}^{\text{pre2012}} / P_{\text{T}}^{\text{revised}}$ are larger than unity, except in the year 2000. In other words, the index $P_{\text{HICP}}^{\text{pre2012}}$ is biased upward.

Concerns about upward bias contributed to a methodological overhaul of the HICP. According to European Union (2016) regulation, from January 2012 onwards, formulas (19) and (20) are no longer appropriate. Instead, the links of the HICP must be computed by the following Lowe-type index:

$$(22) \quad P_{\text{HICP}}^{0 \rightarrow t} = \sum \hat{s}_i \frac{P_i^t}{P_i^0}.$$

The estimation of the expenditure shares \hat{s}_i relies on price data, on the expenditure shares from the benchmark years, and on *preliminary* expenditure information from the national accounts (for details see Appendix E).

Whether the new HICP methodology produces more reliable index numbers is an empirical question that is also addressed in Herzberg *et al.* (2021). They compute two additional experimental indices. One of them is the annually chain-linked index $P_{\text{HICP}}^{\text{prelim}}$ prescribed by the new HICP methodology with the Lowe-type links (22). The superscript “prelim” indicates that the index draws on preliminary information from the national accounts. It turns out that the mean deviation between $P_{\text{HICP}}^{\text{prelim}}$ and $P_{\text{T}}^{\text{revised}}$ is smaller than that between $P_{\text{HICP}}^{\text{pre2012}}$ and $P_{\text{T}}^{\text{revised}}$. However, the reformed index $P_{\text{HICP}}^{\text{prelim}}$ exhibits stronger fluctuations around $P_{\text{T}}^{\text{revised}}$ than the original index $P_{\text{HICP}}^{\text{pre2012}}$. The remaining bias of the German HICP is derived from the ratio $P_{\text{HICP}}^{\text{prelim}} / P_{\text{T}}^{\text{revised}}$. It yields an upward bias of roughly 0.1 percentage points per year. Interestingly, this number is not far away from the 0.087 percentage points which the present paper obtained for the German CPI (see Section 4). The results suggest that, in spite of annual chain-linking, the German HICP is not more reliable than the German CPI.

To decompose the remaining bias of the German HICP into the substitution bias and the bias caused by processing preliminary expenditure information, Herzberg *et al.* (2021) compute an experimental Törnqvist index ($P_{\text{T}}^{\text{prelim}}$) that is equal to the benchmark index $P_{\text{T}}^{\text{revised}}$, but uses preliminary instead of revised expenditure information from the national accounts. The comparison of $P_{\text{T}}^{\text{prelim}}$ and $P_{\text{HICP}}^{\text{prelim}}$ is used to quantify the substitution bias. As expected, the substitution bias is clearly positive ($P_{\text{HICP}}^{\text{prelim}} > P_{\text{T}}^{\text{prelim}}$). The effect of using revised instead of preliminary expenditure information from the national accounts is derived from a comparison of $P_{\text{T}}^{\text{prelim}}$ and $P_{\text{T}}^{\text{revised}}$. The effect is also positive ($P_{\text{T}}^{\text{prelim}} > P_{\text{T}}^{\text{revised}}$) and slightly larger than the substitution bias.

In the study of Herzberg *et al.* (2021), the Törnqvist index $P_{\text{T}}^{\text{revised}}$ is considered as the measure of “true” inflation. This annually chained index can be computed only retrospectively. Since for the comparison period (t) and/or the price reference period (0) no full-scale expenditure survey is conducted, the Törnqvist index $P_{\text{T}}^{\text{revised}}$

must be blended by price-updated expenditure information that mixes data from expenditure surveys, national accounts, and possibly some other sources. Thus, it cannot be ruled out that the benchmark index deviates from the “true” inflation. The methodology outlined in Section 3 avoids the use of price-updated information, though at a cost: it is not possible to construct an annually chained-linked index. If Destatis conducted its full-scale expenditure surveys each year (instead of every fifth year), it would be possible to retrospectively compute a chained Törnqvist index without the mix of data sources. Such an index would be more reliable than P_T^{revised} and it would become possible to identify the effect of chaining.

8. CONCLUDING REMARKS

The expenditure information for the weighting of a national CPI is usually outdated by 13–60 months. The longer this lag, the larger the risk of upward substitution bias in the CPI. In the context of the COVID-19 pandemic, the problem of stale weights received fresh attention. For example, the expenditure shares of package holidays and visits to restaurants significantly dropped (e.g., Eurostat 2020, p. 3). In a Laspeyres index or in Laspeyres-type indices, this drop is not reflected in the weights assigned to the price changes of these two products. Therefore, the role of these price changes would be overstated and the index would be biased *downward*.

The measurement problems caused by stale weights have been well known long before the outbreak of the pandemic. The NSOs address the issue in different ways. To keep the lag short, some NSOs collect the expenditure information every year and attempt to speed up the processing of this information. This is certainly a useful strategy. When this strategy reaches its limits and the lag cannot be reduced further, retrospective revisions can help to tackle any remaining substitution bias. In fact, some NSOs regularly revise their headline CPI or publish some retrospectively computed supplementary CPI. However, such revisions may not completely solve the problem.

The German CPI is a case in point. Without any retrospective revision, the long-run upward substitution bias during the time interval January 2000 and December 2014 would have amounted to almost 0.22 percentage points per year. The official retrospective revisions of the German CPI did not eliminate the yearly bias, but reduced it to 0.087 percentage points. This is a conservative estimate, since it does not include the upward bias attributable to the lower level of aggregation.

This study has introduced a procedure that can effectively address the source of the bias without requiring any additional information. This procedure replaces the Laspeyres or Laspeyres-type index (typically a Lowe or a Young index) by a retrospective price index. Such a retrospective price index can be computed in different ways. Here we discussed the correction approach and the imputation approach and presented various options for implementing these approaches. Both approaches rely on some symmetric price index formula. Suitable candidates include the Walsh, Törnqvist, Fisher, Sato-Vartia, or Marshall-Edgeworth index.

The imputation approach allows for a decomposition of the bias into the contributions of individual groups of items. It was shown that half of the bias in the

German CPI can be attributed to a single group of items even though the expenditure share of that group is below 12 percent.

Retrospective revisions of price indices generate knowledge about past measurement errors. This knowledge, in turn, may improve future measurement methods and, as a consequence, future policies. Even indexed contracts may benefit from retrospective indices. For example, the monthly payments related to indexed social benefits could be specified as the sum of a preliminary part and an adjustment part. The preliminary part is the payment determined by the new and, therefore, preliminary CPI numbers. The adjustment part would be a payment that compensates for the errors caused by the deviation between past preliminary CPI numbers and their recently revised values. Thus, the adjustment payment can be positive or negative. Such composite payments are conceivable also for other contractual arrangements such as rents.

ACKNOWLEDGMENT

Open Access funding enabled and organized by Projekt DEAL.

REFERENCES

- Auer, L. von and A. Shumskikh, "Substitution bias in the measurement of import and export price indices: causes and correction," *Journal of Official Statistics*, 38(1), 107–26, 2022.
- Auer, L. von and J. Wengenroth, "Consistent aggregation with superlative and other price indices," *Journal of the Royal Statistical Society: Series A*, 184(2), 589–615, 2021.
- Akem, A. and A. Opryshko, *Why 'Lowe' When 'Young' and 'Laspeyres' Are Available? Australian Bureau of Statistics. Reference paper 2 for the 'Higher-level Indices Workshop' United Nations Economic Commission for Europe*, 2014. Available at: <https://unece.org/statistics/events/group-experts-consumer-price-indices> [Accessed 30th June 2022].
- Armknecht, P. and M. Silver, "Post-Laspeyres: the case for a new formula for compiling consumer price indexes," *Review of Income and Wealth*, 60(2), 225–44, 2014.
- Bishop, G. and M. van Kints, *Geometric indexes and substitution bias in the CPI*. Australian Bureau of Statistics. Ottawa Group Meeting 2013, 2013. Available at: <https://www.ottawagroup.org/ottawa/ottawagroup.nsf/4a256353001af3ed4b2562bb00121564/89606143e000a3bcca257bb00004da25?OpenDocument> [Accessed 30th June 2022].
- de Haan, J., B. Balk and C. Hansen, "Retrospective approximations of superlative price indexes for years where expenditure data is unavailable," in L. Biggeri, and G. Ferrari (eds), *Price indexes in time and space. Contributions to statistics*, Physica-Verlag, Heidelberg, 25–42, 2010. https://doi.org/10.1007/978-3-7908-2140-6_2 [Accessed 29th September 2021].
- Diewert, E. W., M. Huwiler and U. Kohli, "Retrospective price indices and substitution bias," *Swiss Journal of Economics and Statistics*, 145, 127–35, 2009.
- Diewert, W. E. and K. J. Fox, *Output growth and inflation across space and time*. School of Economics, The University of New South Wales. Discussion Papers 2015-04, 2015. Available at: <https://ec.europa.eu/eurostat/cros/system/files/eurostat1-2017-art1.pdf> [Accessed 30th June 2022].
- Egner, U., "Umstellung des Verbraucherpreisindex auf Basis 2000," *Wirtschaft und Statistik*, 5, 423–32, 2003.
- Egner, U., "Verbraucherpreisstatistik auf neuer Basis 2010," *Wirtschaft und Statistik*, 5, 329–44, 2013.
- Egner, U., "Verbraucherpreisstatistik auf neuer Basis 2015," *Wirtschaft und Statistik*, 5, 86–106, 2019.
- Elbel, G. and U. Egner, "Verbraucherpreisstatistik auf neuer Basis 2005," *Wirtschaft und Statistik*, 4, 339–50, 2008.
- European Union, *Regulation (EU) 2016/792 of the European Parliament and the Council of 11 May 2016*. Official Journal of the European Union, 2016. Available at: <https://eur-lex.europa.eu/eli/reg/2016/792/oj> [Accessed 30th June 2022].
- Eurostat, *Guidance on the compilation of HICP weights in case of large changes in consumer expenditures*. Directorate C, Macro-economic Statistics, 2020. Available at: <https://ec.europa.eu/eurostat/> [Accessed 30th June 2022].

- Feenstra, R. C., R. Inklaar and M. Timmer, *The next generation of the Penn World Table*. National Bureau of Economic Research. Working Paper 19255, 2013. Available at: <https://doi.org/10.3386/w19255> [Accessed 29th June 2022].
- Greenlees, J. S. and E. Williams, “Reconsideration of weighting and updating procedures in the US CPI,” *Jahrbücher für Nationalökonomie und Statistik*, 230(6), 741–58, 2010.
- Hansen, C. B., Recalculations of the Danish CPI 1996 - 2006. In: *Paper presented at the 10th Meeting of the International Working Group on Price Indices (Ottawa Group)*, 9–12 October 2007, Ottawa, 2007.
- Herzberg, J., T. A. Knetsch, P. Schwind and S. Weinand, *Quantifying bias and inaccuracy of upper-level aggregation in HICPs for Germany and the Euro Area*. Deutsche Bundesbank. Discussion Paper 06/2021, 2021. Available at: <https://www.bundesbank.de/resource/blob/861158/4cc8ed30bf2fce8ff50e08c7119b9cf3/mL/2021-03-16-dkp-06-data.pdf> [Accessed 29th March 2021].
- Huang, N., W. Wimalaratne and B. Pollard, *Choice of index number formula and the upper-level substitution bias in the Canadian CPI*. Consumer Prices Division, Statistics Canada. 14th Ottawa Group Meeting, 2015. Available at: https://www.stat.gov.jp/english/info/meetings/og2015/pdf/t1s4p13_paper.pdf [Accessed 30th June 2022].
- Huang, N., W. Wimalaratne and B. Pollard, “The effects of the frequency and implementation lag of basket updates on the Canadian CPI,” *Journal of Official Statistics*, 33(4), 979–1004, 2017.
- ILO, IMF, OECD, UNECE, Eurostat, The World Bank, *Consumer price index manual: theory and practice*. ILO/IMF/OECD/UNECE/Eurostat/The World Bank, Geneva, International Labour Office. Technical report, 2004. Available at: https://www.ilo.org/public/english/bureau/stat/download/cpi/cpi_manual_en.pdf [Accessed 29th March 2021].
- Klick, J., “Improving initial estimates of the chained consumer price index,” *Monthly Labor Review*, 2018, 1-15. <https://doi.org/10.21916/mlr.2018.6> [Accessed 29th March 2021].
- Lebow, D. E. and J. B. Rudd, “Measurement error in the consumer price index: where do we stand?” *Journal of Economic Literature*, 41(1), 159–201, 2003.
- Nimmo, B., S. Hayes and C. Pike, *Choice of index number formula and the upper-level substitution bias in the Canadian CPI*. Prices, Statistics New Zealand. Ottawa Group Conference, 2007. Available at: <https://www.ottawagroup.org/ottawa/ottawagroup.nsf/4a256353001af3ed4b2562bb00121564/3eab247794e923e0ca25757700800ac8?OpenDocument> [Accessed 30th June 2022].
- Rao, D. S. P., R. Inklaar and A. Rambaldi, *Options for producing smoothed PPP time-series for the years between reference year comparisons*. International Comparison Program. Second Meeting of the Technical Advisory Group, 2018. Available at: <https://thedocs.worldbank.org/en/doc/836211539796300849-0050022018/original/ICPTFPCC03DOCS03OptionsforproducingsmoothenedPPPTSforInterimPR.pdf> [Accessed 30th June 2022].
- Rao, D. S. P., A. Rambaldi and H. Doran, “Extrapolation of purchasing power parities using multiple benchmarks and auxiliary information: a new approach,” *Review of Income and Wealth*, 56, S59–98, 2010.
- Statistisches Bundesamt, *Preise: Verbraucherpreisindex für Deutschland*, Statistisches Bundesamt, Wiesbaden, 2018.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher’s web site:

Appendix S1 Supporting Information

Appendix S2 Supporting Information