Multilateral approaches in inflation measurement: Why does the TPD method fail us and can we do something about it?*

Ludwig von Auer† Universität Trier

Sebastian Weinand Eurostat

Draft version: May 2, 2024

Abstract

Some product prices react more elastic to general price level changes than others. This study explains why heterogeneous price level elasticities of product prices invalidate any statistical inference based upon the TPD regression approach and why these heterogeneous elasticities in conjunction with systematic data gaps lead to biased TPD estimates of the price levels. The GEKS approach and the GK approach represent no solution. Therefore, this paper introduces the NLTPD regression, a non-linear generalisation of the TPD regression. Since this novel method estimates the heterogeneous price level elasticities, it avoids the issues of the TPD regression and the bias inherent in the GEKS approach and the GK approach. A simulation compares the performances of the NLTPD regression, the TPD regression, the GEKS method, and the GK method.

Keywords: multilateral price index · TPD method · GEKS method · Geary-Khamis method · measurement bias · simulation

JEL classification: C43 · E31 · E52

---

* The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the EU Commission or the European Statistical System.

† Correspondence to Ludwig von Auer (Universität Trier, Germany, vonauer@uni-trier.de).

1 Introduction

Chaining bilateral price indices of monthly (or even weekly) data often generates considerable measurement bias. This phenomenon is usually denoted as "chain drift bias". As a remedy, Kokoski et al. (1999, p. 141) propose to use a rolling window variant of a multilateral index method. Among the most widely used methods are the time product dummy (TPD) regression, the Gini-Éltető-Köves-Szulc (GEKS) method, and the Geary-Khamis (GK) method.[1] In a comprehensive simulation, Auer (2024a, pp. 12-17) verifies that the TPD regression, the GEKS method as well as the GK method in conjunction with a sensible splicing approach (it connects the rolling windows) effectively curtail the chain drift issue. However, this finding does not preclude that multilateral price indices are prone to other types of bias. The present paper examines this conjecture.

It starts with a critical analysis of the TPD regression approach. Like the other multilateral index methods, the TPD regression aggregates the individual price trends of a group of products into some general price level trend. In contrast to the other methods, the TPD regression allows for statistical inference. However, this advantage rests on an implicit assumption: for each product and each period, the price level elasticity of product prices (that is, the elasticity of a product's price with respect to the general price level) is equal to one. In reality, some products have larger price level elasticities than others. As its first contribution, this study explains why heterogeneous price level elasticities make statistical inference invalid.

The paper's second contribution is related to the issue of data gaps (products with positive turnover that are missing in the available data set). Such gaps can arise in different areas and on different levels of price measurement. It is shown that data gaps exacerbate the problems of the TPD regression. More specifically, heterogeneous price level elasticities in conjunction with missing observations usually lead to biased estimates of the price levels. Only if the data gaps occur completely at random, the estimates of the TPD regression remain unbiased. However, they are still inefficient.

As its third contribution, the present paper examines whether the GEKS method and the GK method are immune to these problems. It is shown that they are not.

Thus, an alternative approach is required. To this end, this paper introduces the *NLTPD regression*, a non-linear generalisation of the TPD regression. The concept of the

[1] Recent surveys of the various methods include Chessa et al. (2017), Chessa (2019), Haan and Krsinich (2014), or Diewert and Fox (2022). The GK method was introduced by Geary (1958) and Khamis (1972) in an interregional price comparison context. This is true also for the TPD regression developed by Summers (1973). Balk (1980) adapts the TPD regression to intertemporal price comparisons of seasonal products. The acronym GEKS honours the publications of Gini (1924) and Gini (1931), Eltetö and Köves (1964), and Szulc (1964). They, too, developed this approach for interregional price comparisons. Balk (1981, pp. 73-74) adapts this approach to an intertemporal context. Ivancic et al. (2011, p. 33) introduce a rolling-window variant of the GEKS approach.

NLTPD regression has been proposed and elaborated in an interregional price comparison context (Auer and Weinand, 2022). The concept can be adapted to the intertemporal context. This is the fourth contribution of this study.

In contrast to the TPD regression, the NLTPD regression provides for each product an estimate of its price level elasticity. As a consequence, the NLTPD regression can be used even when systematic data gaps exist and the products exhibit heterogeneous price level elasticities. In addition, the variance of the estimators can be estimated, providing a basis for valid statistical inference.

The final contribution of this paper is a simulation that takes into account typical characteristics of actual purchasing behaviour such as stockpiling (triggered by sales) and consumption smoothing (caused by adjustment costs). The simulation compares the performances of the NLTPD regression, the TPD regression, the GEKS method, and the GK method in the presence of systematic data gaps and heterogeneous price level elasticities.

The rest of the paper is organized as follows. Using a simple illustrative example, Section 2 explains the TPD regression's failure. Section 3 elucidates that the GEKS method as well as the GK method exhibit the same failure. The NLTPD regression avoids these problems. Its rationale is outlined in Section 4. Section 5 presents a more formal treatment of this method. Section 6 develops a comprehensive simulation that compares the various approaches. Section 7 concludes.

2 Problem

Different products usually exhibit different price trends. The task of the price statistician is to derive from the individual price trends the average price development of the aggregated products. Table 1 shows a stylized example. It lists the prices of three products (i = A, B, C) of some elementary aggregate (or product group) in four consecutive periods (t = 1, 2, 3, 4). For simplicity, it is assumed that within each period the expenditure share of each product is the same.

The elementary aggregate's price levels of the four periods can be calculated using the TPD regression approach. Its advocates point out that it is a simple and transparent regression approach that allows for statistical inference. This feature differentiates the

Period	1	2	3	4
Product A:	22	22	21	20
Product B:	30	32	37	40
Product C:	49	50	52	54

Table 1: Prices of three products during four periods.

TPD regression from many other approaches to index number theory.

The spatial analogue of the TPD regression is the country product dummy (CPD) method introduced by Summers (1973) for interregional price comparisons. The statistical issues of the CPD regression are explored in Auer and Weinand (2022). Therefore, the following presentation draws on their discussion.

The TPD regression implicitly assumes that all products possess the same price level elasticities of product prices. The prices in Table 1 violate this assumption. The general price level increases, but the prices of Product A show a small decline, that is, a negative price level elasticity. The prices of Product B and Product C exhibit positive price level elasticities. In the following, it is shown that, with product-specific price level elasticities, the TPD regression produces biased price levels, barring two cases that are rarely satisfied in real-world measurement problems. Even if those two exceptional cases applied, the TPD regression would still be inefficient and inference would remain invalid.

The price of product i in period t is denoted by p_i^t. The TPD regression assumes that each price can be explained by the linear relationship

$$\ln p_i^t = \ln \pi_i + \ln P^t + u_i^t \,, \tag{1}$$

where P^t is the price level of period t, π_i is the general value of product i, and $u_i^t \sim N(0, \sigma^2)$ is an error term. The TPD model (1) can be transformed into a regression equation with a set of dummy variables that represent the periods and the products (e.g., Haan et al., 2021, p. 397). Ordinary least squares yield estimates of the values of $\ln P^t$ and $\ln \pi_i$.

In the example presented in Table 1, the TPD regression yields estimates of the four periods' logarithmic price levels, $\widehat{\ln P^t}$. Taking anti-logs and normalizing the results such that $\widehat{P}^1 \cdot \widehat{P}^2 \cdot \widehat{P}^3 \cdot \widehat{P}^4 = 1$, yields the following price levels:

$$\widehat{P}^1 = 0.95 \,, \quad \widehat{P}^2 = 0.98 \,, \quad \widehat{P}^3 = 1.02 \,, \quad \widehat{P}^4 = 1.05 \,. \tag{2}$$

The upper left panel of Figure 1 illustrates the TPD regression's derivation of these price levels. The vertical axis shows the observed values of the dependent variable, $\ln p_i^t$, while the horizontal axis shows the unknown logarithmic price levels, $\ln P^t$. For each period t, three price observations exist. In the diagram, these three observations are depicted by a red circle (Product A), a green triangle (Product B), and a blue square (Product C). The three observations are positioned along a dashed vertical line. The intersection of such a dashed line with the horizontal axis is the TPD regression's estimated value $\widehat{\ln P^t}$. Thus, the four intersection points indicated in the upper left panel of Figure 1 are the logarithms of the price levels listed in (2). To each product, i, a coloured straight line with slope 1 is depicted. This slope is implied by the TPD regression (1). More specifically, the slope is the products' price level elasticity, that is, the elasticity of the product price with respect to changes in the general price level. The intersection of the coloured line with the vertical axis is the estimated value $\widehat{\ln \pi_i}$.

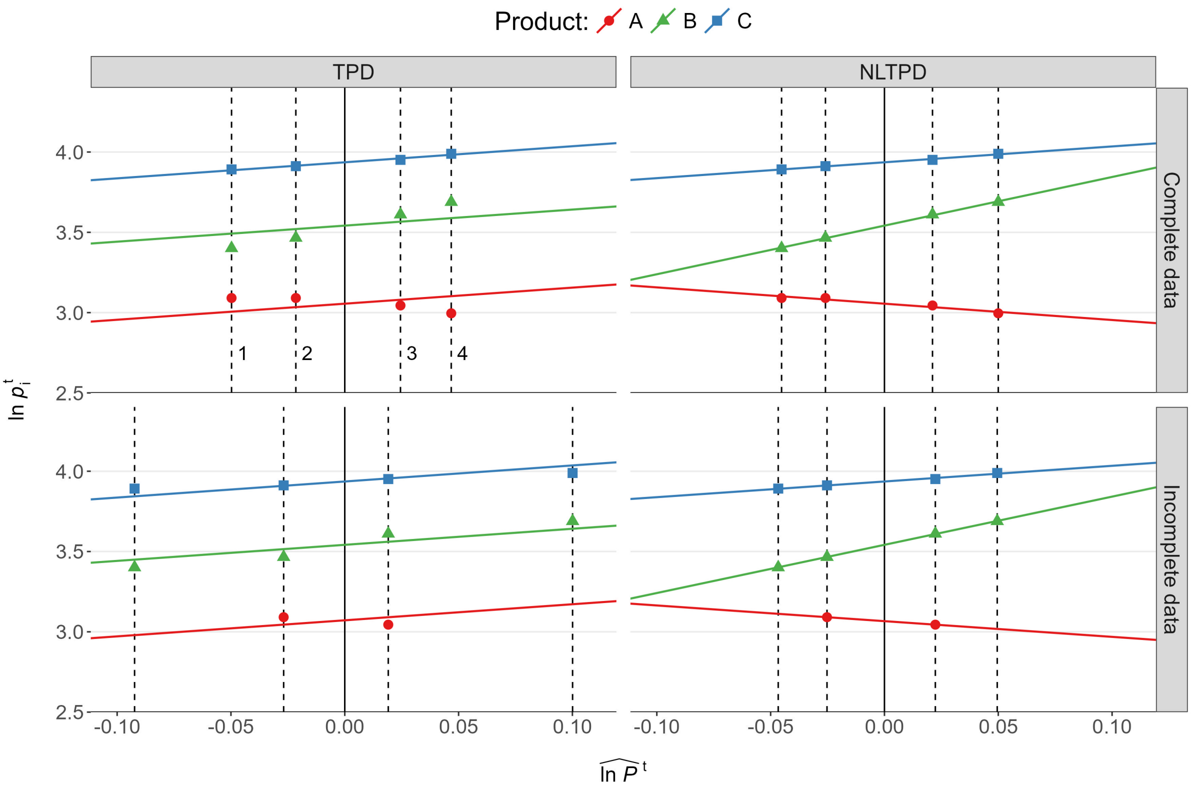

Figure 1: TPD and NLTPD regressions for the complete price data of Table 1 (top panels) and for the same data but with two missing prices of Product A (bottom panels).

A change in the estimated value $\widehat{\ln \pi_i}$ causes a parallel vertical shift of the coloured solid line relating to product i. A change in the estimated value $\widehat{\ln P^t}$ causes a horizontal shift of the dashed vertical line of period t and, therefore, of the three observations relating to that period. Both types of shifts would alter the vertical distance between the observations and their respective solid line. This vertical distance is the residual, $\widehat{u}_i^t$. Graphically speaking, the TPD regression simultaneously shifts the solid lines and the dashed vertical lines (together with their three observations) such that the sum of the (squared) vertical distances between the observations and their respective solid lines is minimized. The upper left panel of Figure 1 depicts the solution to this minimization problem.

Irrespective of the set of missing price observations, the TPD regression assumes that the conditional expected value of the error term is zero: $\mathrm{E}(u_i^t|\mathbf{x}_i^t) = 0$, where $\mathbf{x}_i^t$ represents the values of the two sets of dummy variables underlying the ordinary least squares regression. However, a systematic pattern of missing observations leads to $\mathrm{E}(u_i^t|\mathbf{x}_i^t) \neq 0$, violating the TPD regression's basic assumption. This violation is illustrated in the left panels of Figure 1.

The upper left panel shows the scenario with complete data. The two outer vertical dashed lines indicate the estimated logarithmic price levels of periods 1 and 4, respectively. In the lower left panel of Figure 1, Product A is observed in periods 2 and 3, but missing in periods 1 and 4. Thus, the large positive disturbance relating to Product A in period 1 (red circle in the upper left panel of Figure 1) vanishes in the lower left panel. There, the error terms of the remaining products available during period 1 yield $\mathrm{E}(u_i^1|\mathbf{x}_i^1) < 0$. To reduce the sum of squared residuals of the two remaining price observations of period 1, the TPD regression moves the vertical dashed line of period 1 to the left (see lower left panel of Figure 1). More generally, when in times of inflation a product with a relatively low price level elasticity is missing in one of the early periods, the TPD method's estimated price level of that period always decreases below the level with complete data – in other words, downward bias arises. Similarly, the missing observation in one of the last periods leads to $\mathrm{E}(u_i^4|\mathbf{x}_i^4) > 0$ and the dashed vertical line of that period moves to the right causing upward bias (see lower left panel of Figure 1).

Normalizing the resulting price level estimates in the same way as the price level estimates in (2), yields

$$\widehat{P}^1 = 0.91\,,\quad \widehat{P}^2 = 0.97\,,\quad \widehat{P}^3 = 1.02\,,\quad \widehat{P}^4 = 1.11\,.$$

With incomplete data, the price increase between periods 1 and 4 is 21.2%, while it was 10.1% when the data were complete. This indicates an upward bias. The bias would have the opposite sign if in periods 1 and 4 the price observations of Product B (the product prices with the largest price level elasticity) were missing instead of the price observations of Product A (the product prices with a negative price level elasticity).

There are two exceptional cases in which the TPD regression remains unbiased: a complete data set (upper left panel of Figure 1) or price observations missing completely at random. However, even in those two exceptional cases, the TPD regression would be inefficient and inference would be invalid because the residuals would be both, autocorrelated and heteroskedastic. The upper left panel of Figure 1 illustrates that the general price level increases over time, while the prices of Product A show a weak downward trend. As a consequence, a strong negative correlation between the residuals $\widehat{u}^t_{\text{Product A}}$ and the estimated values of the general price levels, $\ln P^t$, arises. The upward trend in the prices of Product B is more pronounced than the upward trend of the general price level. This leads to a strong positive correlation between the residuals $\widehat{u}^t_{\text{Product B}}$ and the estimated values of $\ln P^t$. Only the price trend of Product C is similar to that of the general price levels. As a consequence, the TPD regression's residuals related to Product C vary less than those related to Product A and Product B. Thus, heteroskedasticity arises.

The residuals' autocorrelation and heteroskedasticity imply that the TPD regression is inefficient and that the estimation of the disturbances' standard deviation is biased. Therefore, inference is invalid, even with complete data. These conclusions are formally proven in Appendix A.4 in Auer and Weinand (2022). Theoretically, the issue of invalid inference could be remedied. Crompton (2000, p. 368) proposes to use White's heteroskedasticity-robust specification of the variance matrix. However, this remedy requires unbiased estimates of the price levels, $\ln P^t$. This, in turn, necessitates either complete price data (as in the upper left panel of Figure 1) or data gaps that arise completely at random. Real world data sets rarely do us this favour.

3 No Solution

In the price statistical literature, the TPD regression approach is not the only multilateral index method. Many writers have advocated the GEKS approach. When weights are available, the GEKS approach computes the price change between two periods by the geometric average of all possible pairs of chained bilateral Törnqvist indices (or Fisher indices) linking the two periods by some third period. In Table 1, no weights are available. In such cases, the GEKS approach routinely uses the Jevons index (geometric mean of the products' intertemporal price ratios) instead of the Törnqvist index (or Fisher index). What happens when the GEKS-Jevons approach is applied to Table 1?

Let $P^{s,t}$ denote the Jevons index relating some price comparison period t to some price reference period s. The price change between periods 1 and 4 as measured by the GEKS approach is the geometric mean of the four chain indices $P^{1,1}P^{1,4}$, $P^{1,2}P^{2,4}$, $P^{1,3}P^{3,4}$, and $P^{1,4}P^{4,4}$. With complete data, the stable prices of Product A cushion the price increases as measured by the Jevons indices $P^{1,2}$, $P^{1,3}$, $P^{1,4}$, $P^{2,4}$, and $P^{3,4}$. When in periods 1 and 4 the observations of Product A are missing, the five listed Jevons indices exclusively use the prices of Products B and C. The resulting index values are larger than with complete data and, therefore, the geometric mean of the four listed chain indices (that is, the GEKS-Jevons index) is larger than with complete data. In other words, upward bias arises. If, instead of Product A, Product B were missing, a downward bias would arise.

This intuition is confirmed by the numerical results derived from Table 1. When Product A is missing in periods 1 and 4, the price levels computed by the GEKS approach yield $\widehat{P}^1 = 0.91\,,\ \widehat{P}^2 = 0.96\,,\ \widehat{P}^3 = 1.03\,,\ \widehat{P}^4 = 1.11$. This is roughly the same price increase as the upward biased price increase indicated by the TPD regression.

There are similar reservations about the GK method because it usually approximates the numerical results of a TPD regression. The simulation presented in Section 6 confirms this conjecture.

The prices listed in Table 1 were used to illustrate the upward bias inherent in multilateral price index approaches when data gaps are correlated with the price level elasticities of product prices. The missing observations of Product A in periods 1 and 4 lead to an upward biased estimate of the price change between periods 1 and 4.

The source of the problem can also be derived from the perspective of a strictly bilateral price comparison of periods 1 and 4 . In Table 1, the price change between periods 1 and 4 could be measured by a Jevons index that relates the prices of period 4 to those of period 1. When the prices of a product are missing, it is unknown whether the associated product prices decelerate or accelerate the change in the general price level. For example, when Product A of Table 1 is missing in periods 1 and/or 4, the Jevons index for these two periods includes only Products B and C. The resulting index number is upward biased. Conversely, the absence of Product B results in a downward biased Jevons index. In other words, data gaps that are correlated with the products' price level elasticities cause biased bilateral index numbers.

The problem could be mitigated or even avoided if one knew whether the missing products tend to be products with low or large price level elasticities. Then, the direction of the bias could be anticipated and, to some degree, corrected. Thus, it is important to learn about the price level elasticity of the products that are missing in periods 1 and 4. This requires to look at other periods than periods 1 and 4. The perspective changes from bilateral to multilateral. For example, when Product A of Table 1 is missing in periods 1 and 4, Product A's prices in periods 2 and 3 indicate that Product A has a low price level elasticity and that a price comparison of periods 1 and 4 without Product A would be upward biased.

Each of the multilateral index approaches discussed (TPD, GEKS, and GK) process the prices of Product A in periods 2 and 3. However, they do it in a way that does not address the source of the bias because they do not identify the price level elasticities of the individual products. Despite using the prices of Product A in periods 2 and 3, the multilateral index methods are biased when the prices of Product A in periods 1 and 4 are missing. Therefore, the present paper advocates an alternative multilateral index method.

4 Solution

For scenarios with complete price data, the TPD regression remains unbiased even in the presence of heterogeneous price level elasticities of product prices. However, the TPD regression is inefficient and inference is invalid. The slope of the three coloured solid lines in the upper left panel of Figure 1 is the price level elasticity of product prices. By assumption, this slope is 1. If each solid line had its own slope, the residuals could be substantially reduced. Thus, the regression model

$$\ln p_i^t = \ln \pi_i + \delta_i \ln P^t + u_i^t \tag{3}$$

generalizes the TPD model (1) by the unknown slope parameters δ_i. These parameters must be estimated, too. Regression model (3) is denoted as NLTPD regression where "NL" stands for "non-linear".

The additional parameters, δ_i, represent the products' price level elasticities. A one per cent increase in the price level P^t changes the period t price of product i by δ_i per cent. Products with elasticities larger than one exhibit a stronger upward trend than the price levels, P^t. In Table 1, this certainly applies to the prices of Product B: $\delta_{\mathrm{B}} > 1$. Since the prices of Product A are decreasing over time, the slope parameter δ_{A} appears to be negative.

Usually, the aggregated products have their individual price level elasticities δ_i. Therefore, most of these elasticities differ from their average, that is, from the trend of the general price level, P^t. The price data listed in Table 1 show the same heterogeneity. In such cases, the TPD model (1) is misspecified. In contrast, the NLTPD model (3) accounts for product-specific price level elasticities, δ_i. In Section 5, it will be shown that these elasticities lead to the restriction $\sum_{i=1}^{3} \delta_i/3 = 1$.

The estimation of the NLTPD model (3) uses the same set of dummy variables as the estimation of the TPD model (1). No additional information is required. The TPD model (1) implicitly assumes that all price level elasticities, δ_i, are equal to one. Thus, it automatically satisfies the restriction $\sum_{i=1}^{3} \delta_i/3 = 1$. For the NLTPD model (3), this restriction must be appropriately implemented in the estimation procedure.

The right panels of Figure 1 illustrate the NLTPD regression. The upper panel shows how the NLTPD regression lines are fitted to the price data listed in Table 1. The estimates of the slopes of the coloured regression lines are $\widehat{\delta}_{\mathrm{A}} = -1.02$, $\widehat{\delta}_{\mathrm{B}} = 3.03$, and $\widehat{\delta}_{\mathrm{C}} = 0.99$. Applying the same type of normalization as before, the NLTPD regression yields the following price levels:

$$\widehat{P}^1 = 0.96\,, \quad \widehat{P}^2 = 0.97\,, \quad \widehat{P}^3 = 1.02\,, \quad \widehat{P}^4 = 1.05\,. \tag{4}$$

Since no prices are missing, these numbers are very similar to those obtained from the TPD regression.

The lower right panel of Figure 1 shows the same price scenario as the lower left panel. The prices of Product A are missing in periods 1 and 4. In contrast to the TPD regression, these data gaps cause hardly any change in the estimated price levels $\widehat{P}^1$ to $\widehat{P}^4$ in Eq. (4). In other words, incomplete data no longer leads to estimation bias. The key to this success is the estimation of the products' price level elasticities δ_i.

Furthermore, the NLTPD method allows for inference because it provides meaningful estimates of the standard errors of all estimated parameters (formally shown in Appendix A.4 of Auer and Weinand, 2022). Thus, one can examine the statistical significance of the coefficients $\widehat{\ln P^r}$, $\widehat{\ln \pi}_i$, and $\widehat{\delta}_i$. If at least one coefficient $\widehat{\delta}_i$ significantly deviates from one, the TPD model is misspecified (unless the data are complete or missing completely at random).

Finally, the model fit of the NLTPD regression is much better. In the case of complete data (upper panels of Figure 1), the sum of squared residuals divided by the degrees of freedom falls from 0.055 (TPD regression) to 0.004 (NLTPD regression).

5 Method

The NLTPD model (3) is a generalization of the linear TPD model (1). The model function is non-linear in its parameters. Consequently, parameter estimates must be derived by non-linear regression.

5.1 Weights and Elasticities

Let $\mathcal{N} = \{1, \ldots, N\}$ denote a set of products offered during one or more periods of the consecutive periods $\mathcal{T} = \{1, \ldots, T\}$. Product i's share of total expenditures during period t is denoted by w_i^t with $\sum_{i\in\mathcal{N}} w_i^t = 1$. Averaging the weights w_i^t over all periods, yields the weights $w_i = (1/T)\sum_{t\in\mathcal{T}} w_i^t$. Thus, $\sum_{i\in\mathcal{N}} w_i = 1$.

Assuming $u_i^s = u_i^t = 0$, the NLTPD regression (3) yields

$$\frac{p_i^t}{p_i^s} = \left(\frac{P^t}{P^s}\right)^{\delta_i} , \tag{5}$$

for all i, s, and t. Consequently, the price ratios p_i^t/p_i^s depend not only on the price level ratio P^t/P^s but also on δ_i and, therefore, are product-specific. Taking logarithms on both sides of Eq. (5), multiplying by w_i, adding over all N products, and exponentiating these sums yields

$$\exp\left(\sum_{i\in\mathcal{N}} w_i \ln \frac{p_i^t}{p_i^s}\right) = \exp\left(\ln \frac{P^t}{P^s} \sum_{i\in\mathcal{N}} w_i\delta_i\right) .$$

The left-hand side is the weighted geometric mean of the N products' price ratios, p_i^t/p_i^s. In a meaningful index approach, this weighted mean should coincide with the price level ratio P^t/P^s. Thus, on the right-hand side, one obtains the restriction

$$\sum_{i\in\mathcal{N}} w_i\delta_i = 1 . \tag{6}$$

In the bilateral case ($T = 2$), this restriction causes the NLTPD price level estimates to coincide with a Törnqvist index (Auer and Weinand, 2022, p. 11).

5.2 Estimators

The residuals $\hat{u}_i^t$ of the NLTPD regression model (3) are defined by $\hat{u}_i^t = \ln p_i^t - \widehat{\delta}_i \widehat{\ln P^t} - \widehat{\ln \pi_i}$. Accordingly, the weighted sum of squared residuals, $S_{\hat{u}_i^t \hat{u}_i^t}$, can be written as

$$S_{\hat{u}_i^t \hat{u}_i^t} = \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{N}_t} w_i^t \left(\ln p_i^t - \widehat{\delta}_i \widehat{\ln P^t} - \widehat{\ln \pi_i} \right)^2 = \sum_{i \in \mathcal{N}} \sum_{t \in \mathcal{T}_i} w_i^t \left(\ln p_i^t - \widehat{\delta}_i \widehat{\ln P^t} - \widehat{\ln \pi_i} \right)^2 \ , \quad (7)$$

where $\mathcal{N}_t$ denotes the set of products for which a price is available in period t. Analogously, $\mathcal{T}_i$ defines the set of periods in which product i is priced. This set's number of products is denoted by T_i.

The NLTPD-estimators can be derived by minimizing $S_{\hat{u}_i^t \hat{u}_i^t}$ with respect to $\widehat{\ln \pi_i}$, $\widehat{\delta}_i$, and $\widehat{\ln P^t}$. However, since $\widehat{\delta}_j \widehat{\ln P^s} = \left(\widehat{\delta}_j \lambda \right) \widehat{\ln P^s} / \lambda$, the estimated price levels, $\widehat{\ln P^s}$, could be arbitrarily scaled up or down by the parameter λ. Therefore, one $\widehat{\delta}_i$-value must be residually derived by restriction (6) instead of being estimated. Note that the TPD model (1) satisfies restriction (6) by assumption ($\delta_i = 1$ for all $i \in \mathcal{N}$). By contrast, the NLTPD estimates of the price level elasticities, δ_i, have to satisfy restriction (6). In the following, the value of $\widehat{\delta}_1$ is derived from this restriction:

$$\widehat{\delta}_1 = \frac{1 - \sum_{i \in \mathcal{N} \setminus \{1\}} w_i \widehat{\delta}_i}{w_1} \ . \quad (8)$$

In Sections 2 and 3, the normalization $\sum_{t \in \mathcal{T}} \widehat{\ln P^t} = 0$ was used to avoid perfect multicollinearity in the NLTPD model (3). In the following, the more common normalization

$$\widehat{\ln P^1} = 0 \quad (9)$$

is applied.

Using restriction (8) and normalization (9) in the minimization of $S_{\hat{u}_i^t \hat{u}_i^t}$, the NLTPD estimator of the logarithmic general value of product i, $\ln \pi_i$, is

$$\widehat{\ln \pi_i} = \sum_{t \in \mathcal{T}_i} \frac{w_i^t}{\sum_{s \in \mathcal{T}_i} w_i^s} \left(\ln p_i^t - \widehat{\delta}_i \widehat{\ln P^t} \right) \quad \text{for } i = 1, \ldots, N \ . \quad (10)$$

This is the expenditure share weighted average of the deflated prices observed of product i where the deflator is the product-specific price level, $\widehat{P^t}^{\widehat{\delta}_i}$, instead of the general price level, $\widehat{P^t}$.

The NLTPD estimator of the logarithmic price level of period t, $\ln P^t$, is

$$\widehat{\ln P^t} = \frac{\sum\limits_{i \in \mathcal{N}_t} w_i^t \widehat{\delta}_i \left(\ln p_i^t - \widehat{\ln \pi_i} \right)}{\sum\limits_{i \in \mathcal{N}_t} w_i^t \left(\widehat{\delta}_i \right)^2} \quad \text{for } t = 2, \ldots, T \; . \tag{11}$$

The numerator is the covariation (across products) of $\left(\ln p_i^t - \widehat{\ln \pi_i} \right)$ and the price level elasticity $\widehat{\delta}_i$. The denominator is the variation (across products) of $\widehat{\delta}_i$. The same formula would be applied in a simple weighted least squares regression where the dependent variable $\left(\ln p_i^t - \widehat{\ln \pi_i} \right)$ is a linear function of the independent variable $\widehat{\delta}_i$. A negative value, $\widehat{\ln P^t}$, indicates a relatively cheap period. It arises when the numerator is negative, that is, when in period t prices, $\ln p_i^t$, below the general value, $\widehat{\ln \pi_i}$, dominate in the sense that they are either more frequent and/or more often arise for products with a large price level elasticity, $\widehat{\delta}_i$. In expensive periods $\left(\widehat{\ln P^t} > 0 \right)$, prices above the general level dominate.

The parameter δ_i represents the elasticity of product i's price with respect to the general price level, $\ln P^t$. The parameter's estimator is

$$\widehat{\delta}_i = \frac{\sum_{t \in \mathcal{T}_i} w_i^t \widehat{\ln P^t} \left(\ln p_i^t - \widehat{\ln \pi_i} \right)}{\sum_{t \in \mathcal{T}_i} w_i^t \left(\widehat{\ln P^t} \right)^2} \quad \text{for } i = 2, \ldots, N \; . \tag{12}$$

The numerator is the covariation (across periods) of the logarithmic price levels, $\widehat{\ln P^t}$, and $\left(\ln p_i^t - \widehat{\ln \pi_i} \right)$. The denominator is the variation (across periods) of the logarithmic price levels. Therefore, the estimator (12) can be viewed as the simple weighted least square estimator of the slope parameter of a simple linear model where $\left(\ln p_i^t - \widehat{\ln \pi_i} \right)$ is regressed on $\widehat{\ln P^t}$. The covariation represented by the numerator is usually positive. The larger this covariation, the stronger the price level elasticity of the product prices, p_i^t. If some product i has a uniform price, then $\widehat{\ln \pi_i} = \ln p_i^t$ and, therefore, the fraction becomes 0.

5.3 Solution Algorithm

The non-linear least squares formulas (10), (11), and (12) do not provide explicit solutions for the coefficients $\widehat{\ln \pi_i}$, $\widehat{\ln P^t}$, and $\widehat{\delta}_i$. Instead, an iterative optimization routine with appropriate start values is necessary. For $\widehat{\ln \pi_i}$ and $\widehat{\ln P^t}$, such start values can be obtained from the TPD regression. Their insertion into formula (12) yields start values for $\widehat{\delta}_i$. The values of $\widehat{\delta}_i$ and $\widehat{\ln P^t}$ can then be included in the calculation of new $\widehat{\ln \pi_i}$-values. The values of $\widehat{\delta}_i$ and the new $\widehat{\ln \pi_i}$-values are used in the computation of new $\widehat{\ln P^t}$-values. The new $\widehat{\ln \pi_i}$- and $\widehat{\ln P^t}$-values are used in the calculation of new $\widehat{\delta}_i$-values, and so on. This cyclical iteration is repeated until the coefficients $\widehat{\ln \pi_i}$, $\widehat{\ln P^t}$, and $\widehat{\delta}_i$ have converged to stable values. This cyclical iterative process is computationally much faster than more sophisticated algorithms.[2]

The TPD method's weighted least squares estimators, $\widehat{\ln \pi_i'}$ and $\widehat{\ln P^{t\prime}}$, are special cases of the NLTPD method. They can be derived from Eqs. (10) and (11) by setting $\widehat{\delta}_i = 1$ ($i = 1, \ldots, N$):

$$\widehat{\ln \pi_i'} = \sum_{t \in \mathcal{T}_i} \frac{w_i^t}{\sum_{s \in \mathcal{T}_i} w_i^s} \left(\ln p_i^t - \widehat{\ln P^t} \right) \quad \text{for } i = 1, \ldots, N \tag{13a}$$

$$\widehat{\ln P^{t\prime}} = \sum_{i \in \mathcal{N}_t} \frac{w_i^t}{\sum_{i \in \mathcal{N}_t} w_i^t} \left(\ln p_i^t - \widehat{\ln \pi_i} \right) \quad \text{for } t = 2, \ldots, T \tag{13b}$$

Taking the exponential of Eqs. (13a) and (13b), respectively, yields the same definitions as in the Rao (1990) system. Rao (2005) has shown that this system gives identical estimates as an expenditure share weighted TPD method. Consequently, the TPD parameter estimates can be obtained by some iterative alternating algorithm exploiting Eqs. (13a) and (13b) (e.g., Maddison and Rao, 1996, pp. 14-17) or by linear regression. Note the similarity of such an iterative alternating algorithm to the cyclical iteration process previously described.

5.4 Limitations

The performance of the NLTPD regression hinges upon its ability to produce reliable estimates of the price level elasticities, δ_i. Problems may arise when there are only a few observations per product, when the elasticities are similar, when sales are concentrated either at the end or at the beginning of the time window covered by the regression, or when periods exist in which only products are offered that exhibit constant prices over time. In such cases, the TPD regression is likely to produce more reliable index numbers than the NLTPD regression. As an alternative, one may consider a restricted version of the NLTPD regression in which some δ_i-values are predetermined (in TPD regressions, all of them are predetermined) and the remaining ones are estimated.

6 Simulation

The illustrative example introduced in Section 2 revealed that the TPD regression is inefficient and that inference based on TPD estimates is invalid. Even worse, in the presence of data gaps, the TPD regression is usually biased. When the δ_i-values of the products are positively correlated with their number of gaps, the bias in the estimated price levels tends to be negative, while a negative correlation leads to positive bias. It was argued that

[2] Examples include Levenberg-Marquardt, Gauss-Newton, (L-)BFGS, and Nelder-Mead (see for example Kelley, 1999, for a comprehensive overview of such algorithms).

this is true not only for the TPD regression but also for the GEKS approach and the GK approach. It would be desirable to know whether the magnitude of the bias differs between the various approaches. A sound empirical analysis of this question requires a meaningful benchmark for quantifying the bias of the various multilateral index methods.

6.1 Benchmark

Simply *defining* the results of the NLTPD regression (or of any other multilateral index method) as the benchmark, would be a prejudiced approach. It is preferable to have a benchmark against which all multilateral index methods, including the NLTPD regression, can be compared. Thus, a dataset with a known general price level change is required. Any deviation between this benchmark and the general price change computed by the investigated multilateral price index can be interpreted as bias.

In practice, the available price and quantity data are generally characterized by individual price trends and shifts in expenditure weights. Consequently, such data do not possess a known general price level change and, therefore, no unassailable benchmark, unless possibly biased value judgements are used to define such a benchmark (e.g., stipulating the results of the NLTPD regression as the benchmark). A promising alternative to such data are carefully designed simulation studies where the purchased quantities of the various products reflect the typical features of actual purchasing behaviour.

One typical feature is stockpiling. During sales, households not only increase their consumption but also build up stocks. Thus, the purchased quantities exceed the level of consumption. After the sale, the households start to reduce their inventories. Thus, during the after-sales period, the purchased quantities fall below the level of consumption. A second feature of actual purchasing behaviour is delayed adjustments to price changes. Adjustment costs are an important cause for such consumption smoothing. Auer (2024a, pp. 6-7) denotes the quantity effects caused by stockpiling as "overshooting quantities" and the quantity effects caused by consumption smoothing as "sticky quantities".

The present study develops a simulation framework that allows for overshooting quantities caused by stockpiling as well as for sticky quantities caused by consumption smoothing. To get an unassailable benchmark, household behaviour is represented by two CES subutility functions (one representing consumption and the other stockpiling) that are embedded in a Cobb-Douglas utility function with exponents adding to one. Details are provided in Auer (2024b, Appendix C, pp. 24-29). From the perspective of utility maximizing households, the products are symmetric in the sense that the purchased quantities of any pair of products equalize when their prices are identical and constant. Given this symmetric appreciation of the available products, it is possible to construct a natural benchmark for the performance of the various multilateral index methods.

As an illustration, Table 2 lists the prices and purchased quantities of three products during periods r and t. In addition, the table lists a hypothetical price-quantity scenario s with prices p_i^s and quantities x_i^s. The period r prices of Products A to C prevailed also during the previous periods. In other words, the quantities x_i^r represent the households' optimal "long-run" reaction to the prices p_i^r. The hypothetical price-quantity scenario s is merely a permutation of the period r scenario. Thus, for households with a symmetric appreciation of the available products, any multilateral price index should indicate that no price change occurred between period r and the hypothetical price-quantity scenario s.[3]

	p_i^r	x_i^r	p_i^s	x_i^s	p_i^t	x_i^t
Product A:	20	6	*40*	*3*	44	3
Product B:	40	3	*10*	*9*	11	9
Product C:	10	9	*20*	*6*	22	6

Table 2: Prices and quantities of three products during periods r and t.

Multiplying the prices p_i^s by the factor 1.1, yields the prices p_i^t. In analogy to period r, it is assumed that the period t prices prevailed also in the preceding periods. Thus, the purchased quantities x_i^t are the households' optimal "long-run" reaction to the prices p_i^t. They are identical to the purchased quantities during period s. Since from period s to period t all prices increase by 10 per cent and all quantities remain constant, the change in the average price level is also 10 per cent.[4]

In sum, the price change between periods r and t can be decomposed into the price change between period r and the hypothetical price-quantity-scenario s and the price change between the hypothetical price-quantity-scenario s and period t. Between period r and scenario s, the price level remains constant, while between scenario s and period t, the price level increases by 10 per cent. Thus, the price level change between periods r and t is also 10 per cent. This number is the benchmark for any multilateral index method comparing the prices of periods r and t. If a multilateral index method yields a price change that deviates from 10 per cent, this deviation can be considered as bias. This is the basic construction principle of the simulation approach developed in the present study.

6.2 Framework

In the simulation approach, $N = 30$ different products exist and the number of periods is $T = 50$. The exogenously given periodic income of the households increases over time. In

[3] Krtscha (1979, p. 69) and Auer (2002, p. 534) formulate an analogous postulate in the context of bilateral axiomatic index theory.

[4] In the context of bilateral axiomatic index theory, the Proportionality axiom makes the same postulate even without imposing the restriction of constant quantities.

period $t = 50$, it is 10 per cent larger than in period $t = 1$. A third of the households are stockpiling households and all other households smooth consumption.

The prices of period $t = 1$ are randomly drawn from the interval $[15, 25]$. The basic pattern of the prices and quantities of periods $t = 1$ and $t = 50$ is like in Table 2. More specifically, the quantities in period $t = 50$ are a permutation of the quantities of period $t = 1$. The prices represent the same permutation, the only difference being that in period $t = 50$ all prices are 10 per cent higher than in period $t = 1$. Thus, total expenditures during period $t = 50$ are 10 per cent larger than during period $t = 1$. After period $t = 47$, all prices remain constant. Thus, the quantities in period $t = 50$ can be considered as long-term reactions to the prices prevailing during periods $t = 48$ to $t = 50$.

The price trend of each product i is linear with an additive stochastic component. The slope of the price trends varies between the products. Some products have downward trending prices. The linear trends are possibly interrupted by stochastic sales. The discounts randomly vary up to 20 per cent. To ensure smooth trends of the multilateral price indices, the average number of sales is not uniform but first increasing and then decreasing over time.

As an illustration, Figure 2 shows a typical price pattern and the corresponding quantities. The price trend depicted in the upper panel resembles a random process with drift. Consumption smoothing leads to sticky quantities, that is, the quantities fluctuate less than the prices. The upward trend of the prices is interrupted by sales in periods $t = 18$ and $t = 27$ and $t = 33$. The corresponding overshooting quantity reactions of stockpiling households are shown in the middle panel. The bottom panel shows the sticky quantity reactions of households that pursue consumption smoothing.

In the simulations, the price level of period $t = 1$ is set to 100: $P^1 = 100$. Thus, the benchmark for the estimated price level $\widehat{P}^{50}$ of the various multilateral index methods is the value 110. Any deviation from this value represents bias.

Four different cases are considered. Case 1 assumes that the data are complete. Thus, all multilateral index methods should yield unbiased $\widehat{P}^{50}$-values. To examine this conjecture, for each multilateral index method, 2000 samples are computed (e.g., 2000 TPD regressions) and the average value of the 2000 $\widehat{P}^{50}$-values is calculated. An average value of 110 would say that the multilateral index method is unbiased. This procedure is conducted for the TPD regression, the NLTPD regression, the GK method, and the GEKS method.

Cases 2 to 4 allow for data gaps. A gap is an item-period combination for which the price and quantity are not recorded. The gaps occur randomly. On average, the probability of a data gap is 5 per cent. Regardless of the gaps, the price-quantity combinations of period $t = 1$ prevail in permuted form in period $t = 50$, the only difference being, that the prices during period $t = 50$ are 10 per cent higher than in period $t = 1$.

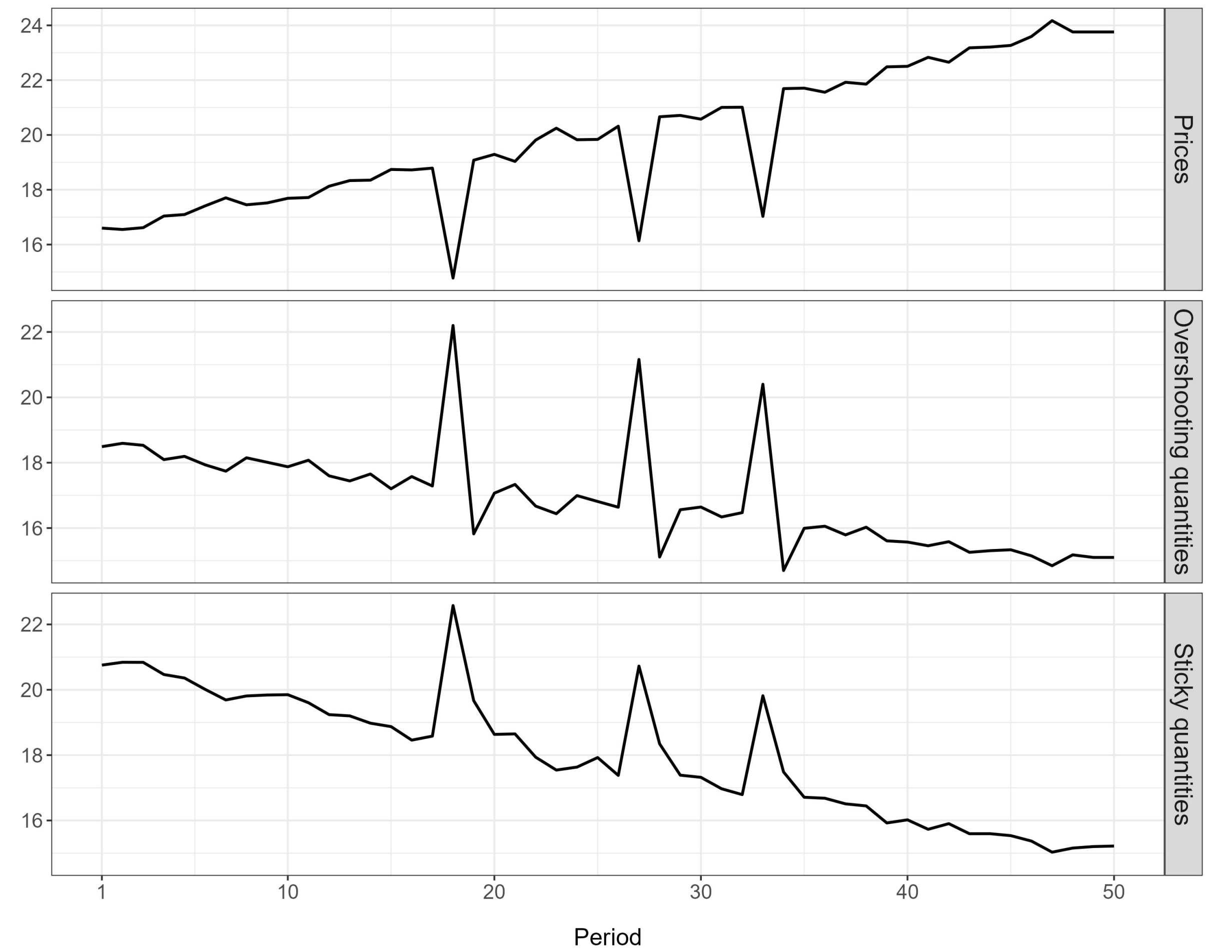

Figure 2: Simulated prices (upper panel) and quantities (middle and bottom panel) of one product.

In Case 2, the probability of a data gap of some product i is independent of its δ_i-value. Therefore, regardless of the applied multilateral index method, the 2000 samples should yield $\widehat{P}^{50}$-values that average to 110. In Case 3, the δ_i-value of some product i and its number of data gaps are negatively correlated. Thus, except for the NLTPD method, the average of the $\widehat{P}^{50}$-values are expected to exceed 110. In Case 4, a positive correlation is considered. Thus, the opposite direction of bias should arise.

6.3 Results

Figure 3 depicts the four cases. For of the four multilateral index approaches, the figure depicts the average price level of each period. As expected, in Cases 1 (complete data) and 2 (gaps completely at random), each of the four multilateral index approaches yields for period $t = 50$ general price levels, $\widehat{P}^{50}$, that closely approximate the benchmark 110.

In Case 3, the δ_i-value of a product is negatively correlated with this product's number of gaps. As a consequence, only the $\widehat{P}^{50}$-values of the NLTPD regression reach on average the benchmark 110. By contrast, the price levels of the TPD regression and the GK are

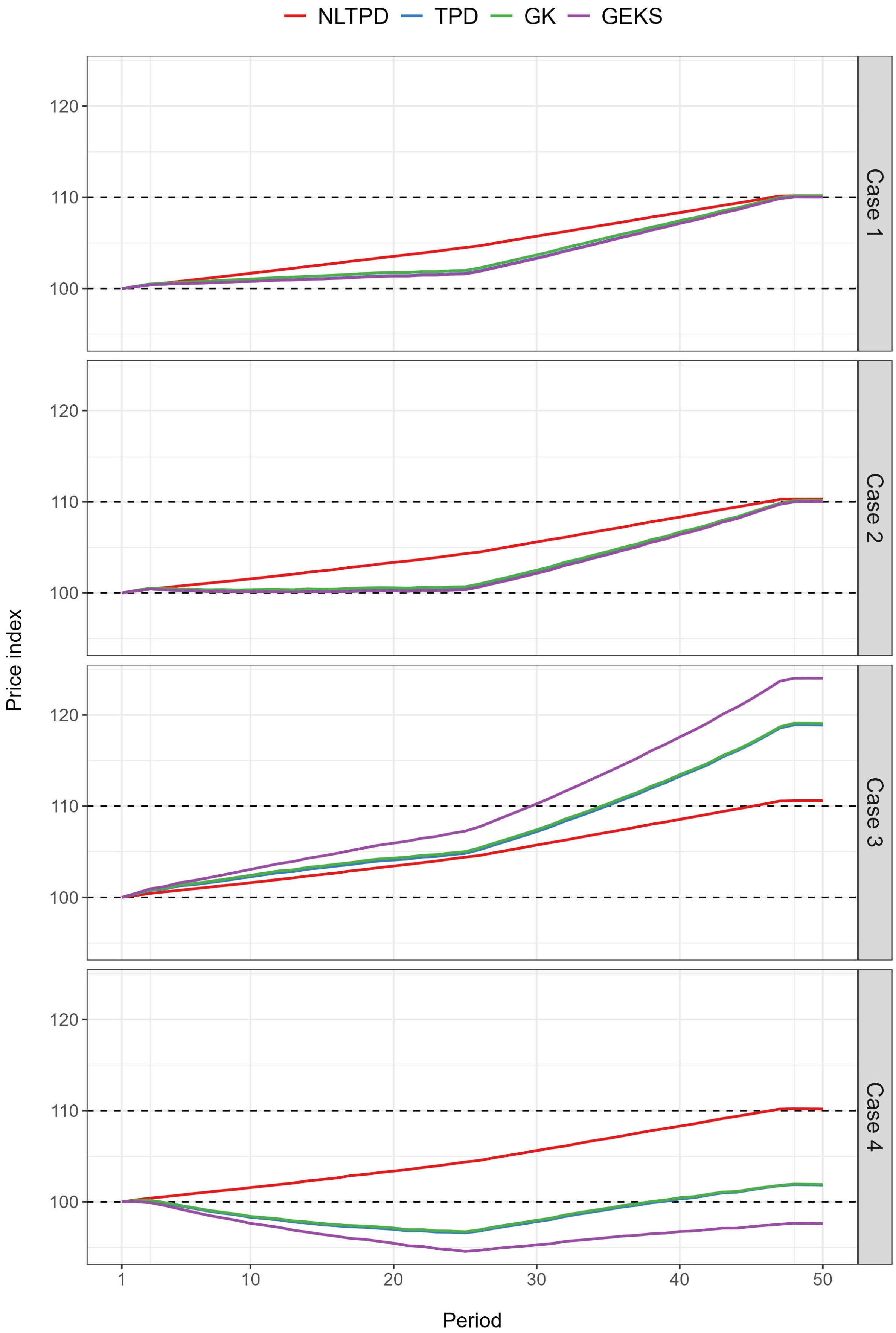

Figure 3: Average of price level estimates of 2000 samples for Cases 1 to 4 and for each of the multilateral price index approaches (NLTPD, TPD, GK, and GEKS).

substantially upward biased. Interestingly, the GEKS approach is even more biased. Case 4 examines a positive correlation between the δ_i-values and the number of gaps. This reverses the direction of the bias. Again, the GEKS approach performs worst.

The price levels of the TPD regression and the GK approach are hardly distinguishable. The price levels of the NLTPD regression are similar for all four cases. In other words, data gaps, be they correlated with the price level elasticities or not, do not have a major influence on the measured price levels. By contrast, the price levels generated by the TPD and GK approaches substantially vary from case to case. The strongest impact of the data gaps arises for the price levels of the GEKS approach.

Table 3 lists the index numbers $\widehat{P}^{50}$ depicted in Figure 3. In Case 1, all multilateral index approaches generate a deviation from 110 that is less than 0.16 percentage points. In Case 2, the deviation is not much larger. In Cases 3 and 4, only the NLTPD regression performs well. In Case 3, the upward bias of the GEKS approach amounts to more than 14 percentage points, while the TPD regression and the GK approach generate an upward bias of around 9 percentage points, respectively. The downward bias in Case 4 is around 8 percentage points (TPD and GK) and more than 12 percentage points (GEKS).

	NLTPD	TPD	GK	GEKS
Case 1	0.158	0.007	0.144	0.015
Case 2	0.291	0.008	0.147	0.010
Case 3	0.598	8.894	9.066	14.025
Case 4	0.182	-8.160	-8.089	-12.375

Table 3: Bias (in index points) of indices in the last simulation period.

The NLTPD regression is not only unbiased, but its standard deviation is also below those of the other multilateral index approaches. The only exception is Case 1 (see Table 4).

	NLTPD	TPD	GK	GEKS
Case 1	0.314	0.002	0.050	0.004
Case 2	0.579	2.519	2.518	2.623
Case 3	0.839	2.937	2.914	3.167
Case 4	0.801	2.475	2.460	2.476

Table 4: Standard deviation (in index points) of indices in the last simulation period.

7 Concluding Remarks

Multilateral price index methods have been advocated as an effective remedy against chain drift bias. However, they have the potential to mitigate also other problems in intertemporal price comparisons. One such problem is data gaps in conjunction with products that exhibit different price level elasticities, that is, different elasticities of the product prices with respect to the general price level. If the gaps are more frequent for products with a small or negative price level elasticity than for products with a large positive price level elasticity, the bilateral index number comparing some price comparison period to some price reference period is upward biased. Downward bias arises when the gaps predominantly relate to products with large positive price level elasticities.

The problem could be cured if it were known whether the gaps are mostly related to products with large price level elasticities or to products with small and negative price level elasticities. Thus, information is required from periods in which the missing products are offered. The most widely used multilateral index methods (GEKS, TPD, and GK) draw on such information. Nevertheless, they are unable to avoid the bias present in the bilateral index because they do not identify the products' price level elasticities. For example, the GEKS approach merely processes bilateral price indices and the latter are biased because of the data gaps. Therefore, the GEKS approach yields biased results, too.

Therefore, this study introduced a novel multilateral index method that amends the TPD regression by an estimation of the price level elasticity of each product. This regression approach is denoted as NLTPD regression because the regression model is non-linear in the parameters to be estimated. In a comprehensive simulation study, it was demonstrated that the NLTPD regression avoids the bias present in index methods that do not take account of heterogeneous price level elasticities.

References

Auer, L. v. (2002). "Spurious inflation: The legacy of Laspeyres and others". In: *The Quarterly Review of Economics and Finance*, 42 (3), pp. 529–542.

— (2024a). "Inflation measurement in the presence of stockpiling and consumption smoothing". In: *Review of Income and Wealth*. forthcoming.

— (2024b). *Inflation measurement in the presence of stockpiling and consumption smoothing*. Research Papers in Economics 02/24. Universität Trier.

Auer, L. v. and S. Weinand (2022). *A Nonlinear Generalization of the Country-Product-Dummy Method*. Discussion Paper 2022/45. Deutsche Bundesbank.

Balk, B. M. (1980). "A Method for Constructing Price Indices for Seasonal Commodities". In: *Journal of the Royal Statistical Society. Series A (General)*, 143 (1), pp. 68–75.

— (1981). "A Simple Method for Constructing Price Indices for Seasonal Commodities". In: *Statistische Hefte*, 22 (1), pp. 72–78.

Chessa, A. G. (2019). "A comparison of index extension methods formultilateral methods". In: *Paper presented at the 16th Meeting of the Ottawa Group, Rio de Janeiro, Brazil, 2019.*

Chessa, A. G., J. Verburg, and L. Willenborg (2017). "A comparison of price index methods for scanner data". In: *Paper presented at the 15th Meeting of the Ottawa Group, Eltville, Germany, 2017.*

Crompton, P. (2000). "Extending the Stochastic Approach to Index Numbers". In: *Applied Economics Letters*, 7 (6), pp. 367–371.

Diewert, W. E. and K. J. Fox (2022). "Substitution Bias in Multilateral Methods for CPI Construction". In: *Journal of Business & Economic Statistics*, 40 (1), pp. 355–369.

Eltetö, Ö. and P. Köves (1964). "On a Problem of Index Number Computation Relating to International Comparison". In: *Statisztikai Szemle*, 42, pp. 507–518.

Geary, R. C. (1958). "A Note on the Comparison of Exchange Rates and Purchasing Power Between Countries". In: *Journal of the Royal Statistical Society. Series A (General)*, 121 (1), pp. 97–99.

Gini, C. (1924). "Quelques Considerations au Sujet de la Construction des Nombres Indices des Prix et des Questions Analogues". In: *Mentron*, 4 (1), pp. 3–162.

— (1931). "On the Circular Test of Index Numbers". In: *International Statistical Review*, 9 (2), pp. 3–25.

Haan, J. de, R. Hendriks, and M. Scholz (2021). "Price Measurement Using Scanner Data: Time-Product Dummy Versus Time Dummy Hedonic Indexes". In: *Review of Income and Wealth*, 67 (2), pp. 394–417.

Haan, J. de and F. Krsinich (2014). "Scanner Data and the Treatment of Quality Change in Nonrevisable Price Indexes". In: *Journal of Business & Economic Statistics*, 32 (3), pp. 341–358.

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 (1), pp. 24–35.

Kelley, C. T. (1999). *Iterative Methods for Optimization*. Society for Industrial and Applied Mathematics.

Khamis, S. H. (1972). "A New System of Index Numbers for National and International Purposes". In: *Journal of the Royal Statistical Society. Series A (General)*, 135 (1), pp. 96–121.

Kokoski, M. F., B. R. Moulton, and K. D. Zieschang (1999). "Interarea Price Comparisons for Heterogeneous Goods and Several Levels of Commodity Aggregation".

In: *International and Interarea Comparisons of Income, Output, and Prices*. Ed. by A. Heston and R. E. Lipsey. University of Chicago Press, pp. 123–169.

Krtscha, M. (1979). “Über Preisindices und deren axiomatische Charakterisierung”. PhD thesis. Universität Karlsruhe.

Maddison, A. and D. S. P. Rao (1996). *A generalized approach to international comparison of agricultural output and productivity*. Working Paper. s.n.

Rao, D. S. P. (1990). “A System of Log-Change Index Numbers for Multilateral Comparisons”. In: *Contributions to Economic Analysis*. Elsevier, pp. 127–139.

— (2005). “On the Equivalence of Weighted Country-Product-Dummy (CPD) Method and the Rao-System for Multilateral Price Comparisons”. In: *Review of Income and Wealth*, 51 (4), pp. 571–580.

Summers, R. (1973). “International Price Comparisons Based upon Incomplete Data”. In: *Review of Income and Wealth*, 19 (1), pp. 1–16.

Szulc, B. J. (1964). “Indices for Multiregional Comparisons”. In: *Przeglad Statystyczny*, 3, pp. 239–254.