Scanner Data in the CPI: The Imputation CCDI Index Revisited

Jan de Haan[a] **and Jacco Daalmans**[b]

24 March 2019

Abstract: The imputation CCDI index combines the multilateral GEKS-Törnqvist, or CCDI, method with hedonic imputations for the "missing prices" of unmatched new and disappearing items. This index is free of chain drift, uses all of the matches in the data and is explicitly quality-adjusted. We discuss different variants of the imputation CCDI index and show how they can be decomposed into the matched-item CCDI price index and a quality-adjustment factor. Using scanner data on TVs provided by a large retailer in the Netherlands, the hedonic imputation CCDI indexes are illustrated and compared with the monthly chained matched-item Törnqvist price index and the expenditure-share weighted time dummy hedonic index.

Keywords: hedonic quality adjustment, multilateral index number methods, new and disappearing items, Törnqvist price index.

JEL Classification: C43, E31.

---

[a] Division of Process Development, IT and Methodology, Statistics Netherlands; j.dehaan@cbs.nl.

[b] Division of Process Development, IT and Methodology, Statistics Netherlands; j.daalmans@cbs.nl.

The authors would like to thank Johan Verburg for assistance with the empirical work and participants at the EMG workshop, 29-30 November 2018, UNSW, Sydney for helpful comments on an earlier version of this paper. The views expressed in this paper are those of the authors and do not necessarily reflect the views of Statistics Netherlands. Note that the scanner data utilized are confidential and cannot be shared with others.

1. Introduction

Over the past decade, several statistical agencies have implemented barcode scanning data, or scanner data for short, in the CPI. Scanner data sets typically contain a census of items sold within a store or retail chain, and it seems worthwhile using all the data rather than taking samples. Scanner data also makes it possible to calculate unit values, i.e. average transaction prices, at the item level, which are preferable to the traditionally observed shelf prices.

The use of scanner data in the CPI poses several challenges (Van Kints, De Haan and Webster, 2019). An important issue that has not been fully resolved is how to deal with item churn. Item churn can be substantial, especially when items are identified by barcode. There are two possible situations: new items, i.e. newly introduced barcodes, are either comparable with disappearing/existing items or they differ in terms of price-determining characteristics. In the first situation, a standard matched-item price index, or maximum-overlap price index as it is also referred to, does not measure any price change during product relaunches when prices and barcodes change at the same time. In the second situation, a matched-item price index would not explicitly adjust for quality change.

Conventional index number theory suggests two basic rules for good practice in price measurement. First, we must *compare like with like*. That is, as much as possible, the prices of the exact same items ought to be tracked over time. Second, an *appropriate index number formula* should be used. The availability of both prices and quantities in scanner data sets enables us to use a superlative index number formula (Diewert, 1976). Superlative indexes have good properties from the viewpoint of the economic approach to index number theory. The Fisher index is “best” in terms of the axiomatic approach, but the Törnqvist index usually approximates the Fisher very well.

To maximize the number of matches in the data, period-on-period chaining of a superlative index was recommended in the CPI Manual (ILO et al., 2004). This is good advice as long as prices and quantities exhibit smooth trends, but in scanner data these trends are often not smooth at all. During promotional sales or discounts, when prices are temporarily lowered, quantities tend to spike, and a chained superlative index then is likely to suffer from (usually downward) chain drift (Diewert, 2018a). A solution is the use of a multilateral rather than bilateral index method.

Multilateral methods were originally developed for spatial price comparisons; for overviews of the different methods, see Balk (1996; 2008) or Diewert (1999). When adapted to the time dimension, multilateral methods estimate indexes simultaneously for all periods within a given "window" and lead to transitive indexes that are free of chain drift. A few statistical agencies already implemented multilateral methods and others are considering doing so.

From the perspective of conventional index number theory, the GEKS method can be viewed as the preferred multilateral method, although Diewert (2018a) advocated similarity linking as an alternative to GEKS. The standard GEKS method uses bilateral matched-item Fisher price indexes as input in the procedure to attain transitivity, and this was the approach taken in the influential paper by Ivancic, Diewert and Fox (2011). Instead of bilateral Fisher indexes, bilateral Törnqvist indexes can be used. This GEKS-Törnqvist or CCDI approach was followed by De Haan and Van der Grient (2011) and has recently been implemented by the Australian Bureau of Statistics (ABS, 2017).

Matched-item GEKS and CCDI price indexes do not explicitly account for new and disappearing items and can suffer from quality change bias or bias due to a change in barcode during a product relaunch. What we would like to have is a GEKS or CCDI index that is explicitly quality-adjusted through imputation of the unmatched new and disappearing items' "missing prices". The present paper revisits the imputation CCDI method where the imputations are based on hedonic regressions. It follows up on De Haan and Krsinich (2014) who developed a specific type of hedonic imputation CCDI index. Statistics New Zealand implemented their method in the CPI for scanner data on consumer electronics goods (Statistics New Zealand, 2014).

The rest of the paper is structured as follows. Section 2 explains the imputation Törnqvist price index. In Section 3, we propose running a separate hedonic regression for each period to estimate the unmatched items' missing prices. The obvious method is "single" imputation, but "double" imputation, where the observed prices of unmatched items are replaced by predicted values from the hedonic regressions, could help reduce omitted variables bias. Section 4 describes the imputation CCDI index and shows that it can be written as the product of the matched-item CCDI index and a quality-adjustment factor. Section 5 addresses the issue of product relaunches and the impact of how items are identified. Section 6 provides an example using Dutch scanner data on TVs. Section 7 discusses our findings and suggests future work.

2. The imputation Törnqvist price index

With scanner data, prices and quantities purchased are available for all product varieties, or items as we will call them, belonging to a product category. This means that the use of a superlative index number formula to calculate aggregate price change is feasible. We use the Törnqvist formula because its geometric structure assists the decompositions outlined below.

Suppose that the set of items U is fixed across time periods $t = 0,...,T$. Denoting the price of item i in the base period 0 and the comparison period t (>0) by p_i^0 and p_i^t, and the corresponding expenditure shares by s_i^0 and s_i^t, the bilateral Törnqvist price index going directly from period 0 to period t is

$$P_T^{0t} = \prod_{i \in U} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{s_i^0 + s_i^t}{2}} . \qquad (1)$$

In reality, significant churn is typically observed at the item level, depending, as will be explained later on, on how items are identified; there are likely to be many new and disappearing items across the entire sample period (or estimation window) $0,...,T$. We use the following notation. U^0 and U^t denote the set of items purchased in periods 0 and t. The subset of U^0 that is not purchased in period t is denoted by U_D^{0t}, and the subset of U^t that is not purchased in period 0 is denoted by U_N^{0t}. $U_M^{0t} = U^0 \cap U^t$ is the subset of items purchased in period 0 as well as in period t, so that $U_M^{0t} \cup U_D^{0t} = U^0$ and $U_M^{0t} \cup U_N^{0t} = U^t$.

Every item purchased in period 0 and/or period t should in principle be included in a price (and quantity) comparison between 0 and t – the price index must be defined on the union $U^0 \cup U^t = U_M^{0t} \cup U_D^{0t} \cup U_N^{0t}$. The period t prices for $i \in U_D^{0t}$ and the period 0 prices for $i \in U_N^{0t}$ are unobservable or "missing" and have to be imputed by $\hat{p}_i^t$ and $\hat{p}_i^0$. By definition, we have $s_i^t = 0$ for $i \in U_D^{0t}$ and $s_i^0 = 0$ for $i \in U_N^{0t}$. Thus, the (single) *imputation Törnqvist price index* defined on the union becomes

$$P_{IT}^{0t} = \prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{s_i^0 + s_i^t}{2}} \prod_{i \in U_D^{0t}} \left(\frac{\hat{p}_i^t}{p_i^0} \right)^{\frac{s_i^0}{2}} \prod_{i \in U_N^{0t}} \left(\frac{p_i^t}{\hat{p}_i^0} \right)^{\frac{s_i^t}{2}}$$

$$= \left[\prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{s_i^0} \prod_{i \in U_D^{0t}} \left(\frac{\hat{p}_i^t}{p_i^0} \right)^{s_i^0} \right]^{\frac{1}{2}} \left[\prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{s_i^t} \prod_{i \in U_N^{0t}} \left(\frac{p_i^t}{\hat{p}_i^0} \right)^{s_i^t} \right]^{\frac{1}{2}}$$

$$= \left[\left[\prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{s_{iM(0t)}^0} \right]^{1-s_{D(0t)}^0} \left[\prod_{i \in U_D^{0t}} \left(\frac{\hat{p}_i^t}{p_i^0} \right)^{s_{iD(0t)}^0} \right]^{s_{D(0t)}^0} \right]^{\frac{1}{2}}$$

$$\times \left[\left[\prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{s_{iM(0t)}^t} \right]^{1-s_{N(0t)}^t} \left[\prod_{i \in U_N^{0t}} \left(\frac{p_i^t}{\hat{p}_i^0} \right)^{s_{iN(0t)}^t} \right]^{s_{N(0t)}^t} \right]^{\frac{1}{2}}, \qquad (2)$$

where $s_{iM(0t)}^0$ and $s_{iM(0t)}^t$ denote the expenditure shares of item i with respect to the set U_M^{0t} in period 0 and period t, respectively; $s_{iD(0t)}^0$ is the period 0 expenditure share of i with respect to the set U_D^{0t} of disappearing items, and $s_{iN(0t)}^t$ is the period t expenditure share of i with respect to the set U_N^{0t} of new items; $s_{D(0t)}^0 = \sum_{i \in U_D^{0t}} s_i^0$ is the aggregate period 0 expenditure share of disappearing items, and $s_{N(0t)}^t = \sum_{i \in U_N^{0t}} s_i^t$ is the aggregate period t expenditure share of new items. Notice that $1 - s_{D(0t)}^0 = s_{M(0t)}^0 = \sum_{i \in U_M^{0t}} s_i^0$ and $1 - s_{N(0t)}^t = s_{M(0t)}^t = \sum_{i \in U_M^{0t}} s_i^t$ are the matched items' aggregate shares.

The second expression of (2) writes the imputation Törnqvist price index as the geometric average of the imputation geometric Laspeyres price index, which is defined on the period 0 set U^0, and the imputation geometric Paasche price index, defined on the period t set U^t. Because $U_D^{t0} = U_N^{0t}$ and $U_N^{t0} = U_D^{0t}$, P_{IT}^{0t} satisfies the time reversal test $(P_{IT}^{t0} = 1/P_{IT}^{0t})$ when the same imputed values enter the calculation of the price index going backwards.

From the third expression of (2) it is easy to verify that the imputation Törnqvist price index can be decomposed as

$$P_{IT}^{0t} = \prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{s_{iM(0t)}^0 + s_{iM(0t)}^t}{2}} \left[\frac{\prod_{i \in U_D^{0t}} \left(\frac{\hat{p}_i^t}{p_i^0} \right)^{s_{iD(0t)}^0}}{\prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{s_{iM(0t)}^0}} \right]^{\frac{s_{D(0t)}^0}{2}} \left[\frac{\prod_{i \in U_N^{0t}} \left(\frac{p_i^t}{\hat{p}_i^0} \right)^{s_{iN(0t)}^t}}{\prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{s_{iM(0t)}^t}} \right]^{\frac{s_{N(0t)}^t}{2}} = P_{MT}^{0t} D^{0t} N^{0t}. \qquad (3)$$

Decomposition (3) was discussed earlier by De Haan and Krsinich (2014). Diewert, Fox and Schreyer (2018) derived the same decomposition in a slightly different manner. The first component in (3), P_{MT}^{0t}, is the matched-item Törnqvist price index. The second component, D^{0t}, is equal to the ratio, raised to the power of $s_{D(0t)}^0 / 2$, of the imputation geometric Laspeyres price index for the disappearing items and the geometric Laspeyres price index for the matched items. The third component, N^{0t}, equals the ratio, raised to the power of $s^t_{N(0t)}/2$, of the imputation geometric Paasche price index for the new items and the geometric Paasche price index for the matched items. D^{0t} and N^{0t} are set equal to 1 if U^{0t}_D and U^{0t}_N are empty. The use of P^{0t}_{MT} to measure price change implicitly assumes that D^{0t} and N^{0t} cancel out or that the missing prices have been imputed such that both components are equal to 1. These are strong assumptions.

3. The use of hedonic regression

Imputation normally requires some sort of modeling. Our goal is to construct imputation price indexes at the product level, where a product consists of different varieties (items), such as different TV models. Ohta and Griliches (1976, p. 326) mentioned that "What the hedonic approach attempted was to provide a tool for estimating "missing prices", prices of particular bundles not observed in the original or later periods. [.....] Because of its focus on price explanation and its purpose of "predicting" the price of unobserved variants of a commodity in particular periods, the hedonic hypothesis can be viewed as asserting the existence of a reduced-form relationship between prices and the various characteristics of the commodity." We agree with their view and consider hedonics to be the preferred approach to estimating missing prices at the item level.

Hedonic regression requires information on the price determining characteristics. We assume that characteristics information is observed by the statistical agency but that some relevant characteristics may not be available, in which case an omitted variables problem arises. We only examine log-linear models. These models usually perform well in empirical applications.

The estimating equation for the log-linear hedonic model in period t $(t = 0,...,T)$ can be written as

$$\ln p^t_i = \alpha^t + \sum_{k=1}^{K} \beta^t_k z_{ik} + \varepsilon^t_i, \tag{4}$$

where z_{ik} is the quantity of the k-th characteristic $(k = 0,...,K)$ for item i and β^t_k the corresponding parameter; α^t is the intercept and ε^t_i is an error term with a zero mean. The quantities z_{ik} are assumed time invariant. This means we do not include life cycle effects (see Melser and Syed, 2016) or other time-dependent variables. For most newly produced consumer goods this is appropriate and in line with almost all hedonic studies in the literature.

We assume that equation (4) is estimated on the data of each period separately by Weighted Least Squares (WLS) regression; to reflect the economic importance of the items we follow Diewert, Heravi and Silver (2009) and set the regression weights equal to the items' expenditure shares in each period. The estimated parameters are $\hat{\alpha}^t$ and $\hat{\beta}_k^t$ $(t = 0,...,T; k = 1,..., K)$, and the predicted prices are $\hat{p}_i^t = \exp(\hat{\alpha}^t)\exp[\sum_{k=1}^{K} \hat{\beta}_k^t z_{ik}]$. The predicted prices for the disappearing and the new items, given their characteristics, serve as imputations in (3). This so-called *single imputation* method is a natural choice as it restricts imputations to the missing prices and leaves unaffected all the observed prices, both for unmatched and matched items.

De Haan and Krsinich (2014) implemented a different single imputation method. They estimated the hedonic model on the pooled data of the two periods compared. The estimating equation for the bilateral Time Dummy Hedonic (TDH) method, where the characteristics parameters are now constrained to be the same in periods 0 and *t*, is

$$\ln p_i^t = \alpha + \delta^t D_i^{0t} + \sum_{k=1}^{K} \beta_k z_{ik} + \varepsilon_i^t, \tag{5}$$

where D_i^{0t} is a dummy variable that has the value 1 if the observation is from period *t* and 0 otherwise. The TDH index is found by exponentiating the estimated time dummy parameter $\hat{\delta}^t$, i.e. $P_{TDH}^{0t} = \exp(\hat{\delta}^t)$. De Haan (2004a) showed that if (5) is estimated by WLS regression with weights $(s_i^0 + s_i^t)/2$ for $i \in U_M^{0t}$, $s_i^0/2$ for $i \in U_D^{0t}$, and $s_i^t/2$ for $i \in U_N^{0t}$, P_{TDH}^{0t} can be written as a single hedonic imputation Törnqvist price index where the missing prices are equal to the predicted prices $\hat{p}_i^t = \exp(\hat{\alpha})\exp(\hat{\delta}^t)\exp[\sum_{k=1}^{K} \hat{\beta}_k z_{ik}]$ for $i \in U_D^{0t}$ and $\hat{p}_i^0 = \exp(\hat{\alpha})\exp[\sum_{k=1}^{K} \hat{\beta}_k z_{ik}]$ for $i \in U_N^{0t}$ from the pooled regression.

Pooling data preserves degrees of freedom and yields more efficient estimates, but constraining characteristics parameters to be fixed over time may be too restrictive (Pakes, 2003) and lead to bias in the index. Unless statistical tests consistently support parameter fixity, it would be preferable to estimate the model for each period separately, data permitting.

When relevant characteristics are omitted from the hedonic model, the predicted prices may be biased. To counteract bias in the price index, De Haan (2004b) and Hill and Melser (2008) proposed *double imputation* where, in addition to the prediction of the missing prices, the observed prices for the new and disappearing items are replaced by their predicted values from the regressions. The double imputation Törnqvist price index is defined as

$$P_{DIT}^{0t} = \prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{s_i^0 + s_i^t}{2}} \prod_{i \in U_D^{0t}} \left(\frac{\hat{p}_i^t}{\hat{p}_i^0} \right)^{\frac{s_i^0}{2}} \prod_{i \in U_N^{0t}} \left(\frac{\hat{p}_i^t}{\hat{p}_i^0} \right)^{\frac{s_i^t}{2}}, \tag{6}$$

which can be decomposed, similar to (3), as

$$P_{DIT}^{0t} = \prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{s_{iM(0t)}^0 + s_{iM(0t)}^t}{2}} \left[\frac{\prod_{i \in U_D^{0t}} \left(\frac{\hat{p}_i^t}{\hat{p}_i^0} \right)^{s_{iD(0t)}^0}}{\prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{s_{iM(0t)}^0}} \right]^{\frac{s_{D(0t)}^0}{2}} \left[\frac{\prod_{i \in U_N^{0t}} \left(\frac{\hat{p}_i^t}{\hat{p}_i^0} \right)^{s_{iN(0t)}^t}}{\prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{s_{iM(0t)}^t}} \right]^{\frac{s_{N(0t)}^t}{2}} = P_{MT}^{0t} D_{DI}^{0t} N_{DI}^{0t}. \tag{7}$$

The idea behind double imputation is that the (omitted variables) biases of the predicted prices in the numerator and denominator of the price relatives for the unmatched new and disappearing items are expected to cancel out, at least partially. Note that the double imputation Törnqvist price index also satisfies the time reversal test, again provided that the same imputed values are used to calculate the index going backwards.

There is a straightforward relationship between the single and double imputation indexes. The second and third components of (7) can be written as

$$\left[\frac{\prod_{i \in U_D^{0t}} \left(\frac{\hat{p}_i^t}{p_i^0} \right)^{s_{iD(0t)}^0}}{\prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{s_{iM(0t)}^0}} \right]^{\frac{s_{D(0t)}^0}{2}} \left[\prod_{i \in U_D^{0t}} \left(\frac{p_i^0}{\hat{p}_i^0} \right)^{s_{iD(0t)}^0} \right]^{\frac{s_{D(0t)}^0}{2}}; \tag{8}$$

$$\left[\frac{\prod_{i \in U_N^{0t}} \left(\frac{p_i^t}{\hat{p}_i^0} \right)^{s_{iN(0t)}^t}}{\prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{s_{iM(0t)}^t}} \right]^{\frac{s_{N(0t)}^t}{2}} \left[\prod_{i \in U_N^{0t}} \left(\frac{\hat{p}_i^t}{p_i^t} \right)^{s_{iN(0t)}^t} \right]^{\frac{s_{N(0t)}^t}{2}}. \tag{9}$$

Using (3), it follows that

$$P_{DIT}^{0t} = \left[\prod_{i \in U_D^{0t}} \left(\frac{p_i^0}{\hat{p}_i^0} \right)^{s_{iD(0t)}^0} \right]^{\frac{s_{D(0t)}^0}{2}} \left[\prod_{i \in U_N^{0t}} \left(\frac{\hat{p}_i^t}{p_i^t} \right)^{s_{iN(0t)}^t} \right]^{\frac{s_{N(0t)}^t}{2}} P_{IT}^{0t}. \tag{10}$$

The weighted averages of the regression residuals $e_i^0 = \ln(p_i^0 / \hat{p}_i^0)$ and $e_i^t = \ln(p_i^t / \hat{p}_i^t)$ for the disappearing items in period 0 and for the new items in period t, respectively, are given by $\bar{e}^0_{D(0t)} = \sum_{i \in U^{0t}_D} s^0_{iD(0t)} e^0_i$ and $\bar{e}^t_{N(0t)} = \sum_{i \in U^{0t}_N} s^t_{iN(0t)} e^t_i$. Thus equation (10) can also be written as

$$P^{0t}_{DIT} = \exp\left[\frac{s^0_{D(0t)}}{2}\bar{e}^0_{D(0t)} - \frac{s^t_{N(0t)}}{2}\bar{e}^t_{N(0t)}\right] P^{0t}_{IT}. \qquad (11)$$

In Section 6 we will argue that under certain pricing strategies of the retailers, a difference may be found between the residuals of new and disappearing items, and (11) shows how this generates a difference between P^{0t}_{IT} and P^{0t}_{DIT}. For example, if $\bar{e}^t_{N(0t)} > 0$ and $\bar{e}^0_{D(0t)} < 0$, we find $P^{0t}_{DIT} < P^{0t}_{IT}$. When there are no omitted variables in the hedonic regressions, single imputation seems to be the preferred approach, and in this example P^{0t}_{DIT} would then have a downward bias.

In practice there will almost always be one or more omitted variables, however. Although newness as such is not a characteristic to control for when measuring quality-adjusted price change, new items can have additional features which consumers value but which a statistical agency does not observe. Part of the relatively high prices of new items in period *t*, giving rise to $\bar{e}^t_{N(0t)} > 0$, could be the result of not controlling for these unobserved features. To some extent at least, the bracketed factor in (11) adjusts for this effect, and so P^{0t}_{DIT} might be a better choice than P^{0t}_{IT}.

The importance of omitted variables is uncertain in many cases and making an informed choice between P^{0t}_{DIT} and P^{0t}_{IT} can be difficult. When in doubt, we could take the geometric average of the two imputation indexes:

$$\left(P^{0t}_{DIT} P^{0t}_{IT}\right)^{\frac{1}{2}} = \exp\left[\frac{s^0_{D(0t)}}{4}\bar{e}^0_{D(0t)} - \frac{s^t_{N(0t)}}{4}\bar{e}^t_{N(0t)}\right] P^{0t}_{IT}. \qquad (12)$$

4. The imputation CCDI index

A disadvantage of the bilateral imputation Törnqvist index is that the result will become increasingly reliant on imputed prices since the number of matches in the data between the base period 0 and the comparison period *t* likely diminishes over time. That is, the index becomes increasingly model-based and ignores many matched items that may be available across the window $0,...,T$. An alternative approach would be to calculate the period-on-period chained (single or double) imputation Törnqvist price index. However, empirical research revealed that high-frequency chaining of matched-item price indexes, including superlative price indexes, can lead to significant drift (Feenstra and Shapiro, 2003; Ivancic, 2007).

To deal with the chain drift problem, Ivancic, Diewert and Fox (2011) proposed using a multilateral method, in particular GEKS (Gini, 1931; Eltetö and Köves, 1964; Szulc, 1964). De Haan and Van der Grient (2011) followed up on their work and used bilateral Törnqvist instead of Fisher indexes as building blocks in the GEKS procedure. The matched-item GEKS-Törnqvist index is also known as CCDI (Caves, Christensen and Diewert, 1982; Inklaar and Diewert, 2016) index.

The CCDI index equals the geometric mean of the ratios of all possible bilateral matched-item Törnqvist price indexes, where each link period l $(0 \leq l \leq T)$ serves as the base. The CCDI index going from the starting period 0 of the time series to comparison period t $(t = 1,...,T)$ can be expressed as

$$P_{CCDI}^{0t} = \prod_{l=0}^{T} \left[P_{MT}^{0l} / P_{MT}^{tl} \right]^{\frac{1}{T+1}} = \prod_{l=0}^{T} \left[P_{MT}^{0l} P_{MT}^{lt} \right]^{\frac{1}{T+1}} , \tag{13}$$

where P_{MT}^{0l} and P_{MT}^{lt} are given by (1) for matched items; note that l, the "base period" in P_{MT}^{lt}, can be greater than t. Taking the mean across all possible link periods ensures that the index will be independent of the choice of base period, and it can be shown that this implies path-independency or *transitivity*. Since bilateral matched-item Törnqvist price indexes satisfy the time reversal test, the CCDI index also satisfies this test.

In general, $P_{CCDI}^{0t} \neq P_{MT}^{0t}$. If all the items were matched across the entire window, chaining would be unnecessary and the bilateral (direct) index P_{MT}^{0t} would be preferable to P_{CCDI}^{0t}. Notice, however, that for $l = 0$ and $l = t$ in (13) we have $P_{MT}^{00} P_{MT}^{0t} = P_{MT}^{0t}$ and $P_{MT}^{0t} P_{MT}^{tt} = P_{MT}^{0t}$. "Double counting" of P_{MT}^{0t}, which is required to make P_{CCDI}^{0t} transitive, means that the direct comparison between 0 and t weights twice as much as each of the indirect comparisons.

The CCDI method can be extended to include unmatched new and disappearing items by using bilateral imputation rather than matched-item Törnqvist indexes. We use the acronym ICCDI for single Imputation CCDI and DICCDI for the Double Imputation variant. Thus, with P_{IT}^{0l} and P_{IT}^{lt}, given by (2), the ICCDI index can be expressed as

$$P_{ICCDI}^{0t} = \prod_{l=0}^{T} \left[P_{IT}^{0l} / P_{IT}^{tl} \right]^{\frac{1}{T+1}} = \prod_{l=0}^{T} \left[P_{IT}^{0l} P_{IT}^{lt} \right]^{\frac{1}{T+1}} . \tag{14}$$

The DICCDI index is found by replacing P_{IT}^{0l} and P_{IT}^{lt} by P_{DIT}^{0l} and P_{DIT}^{lt}, given by (6). Just like the CCDI index, ICCDI and DICCDI indexes satisfy the time reversal test and are transitive.

Substituting the second expression of decomposition (3) for P_{IT}^{0l} and P_{IT}^{lt} into the second expression of (14) yields the following possible decomposition:

$$P_{ICCDI}^{0t} = P_{CCDI}^{0t} D_{SI}^{0t} N_{SI}^{0t}, \tag{15}$$

with $D_{SI}^{0t} = \prod_{l=0}^{T} [D^{0l} D^{lt}]^{1/(T+1)}$ and $N_{SI}^{0t} = \prod_{l=0}^{T} [N^{0l} N^{lt}]^{1/(T+1)}$. In a multilateral context, where we are not comparing 0 and *t* directly, the notions of "new" and "disappearing" become somewhat blurred. This impedes the interpretation of the components D_{SI}^{0t} and N_{SI}^{0t}. We therefore refrain from making a distinction between the effects of new items and disappearing items and decompose P_{ICCDI}^{0t} as

$$P_{ICCDI}^{0t} = P_{CCDI}^{0t} \Omega_{SI}^{0t}, \tag{16}$$

where $\Omega_{SI}^{0t} = \prod_{l=0}^{T} [D^{0l} N^{0l} D^{lt} N^{lt}]^{1/(T+1)}$ is a *quality-adjustment factor* which measures the impact of the *unmatched* items across the estimation window $0,...,T$. Recall that the constituent elements D^{0l}, D^{lt}, N^{0l} and N^{lt} are equal to 1 when the corresponding sets of disappearing and new items (U_D^{0l}, U_D^{lt}, U_N^{0l} and U_N^{lt}) are empty. For our purpose it is not necessary to estimate the various elements; the quality-adjustment factor can simply be derived as $\Omega_{SI}^{0t} = P_{ICCDI}^{0t} / P_{CCDI}^{0t}$.

Similarly, the DICCDI index can be decomposed as

$$P_{DICCDI}^{0t} = P_{CCDI}^{0t} \Omega_{DI}^{0t}, \tag{17}$$

where $\Omega_{DI}^{0t} = \prod_{l=0}^{T} [D_{DI}^{0l} N_{DI}^{0l} D_{DI}^{lt} N_{DI}^{lt}]^{1/(T+1)}$ is again a quality-adjustment factor which can be derived implicitly as $\Omega_{DI}^{0t} = P_{DICCDI}^{0t} / P_{CCDI}^{0t}$. A comparison of the time series for Ω_{SI}^{0t} and Ω_{DI}^{0t} shows the impact of single versus double imputation. In Section 3 we argued that in some cases it may be useful to take the geometric mean of the bilateral single and double imputation Törnqvist price indexes. With $(P_{IT}^{0l} P_{DIT}^{0l})^{1/2}$ and $(P_{IT}^{lt} P_{DIT}^{lt})^{1/2}$ feeding into the GEKS procedure, which of course leads to the geometric average of the ICCDI and DICCDI indexes, we have

$$\prod_{l=0}^{T} \left[\left(P_{IT}^{0l} P_{DIT}^{0l} \right)^{\frac{1}{2}} \left(P_{IT}^{lt} P_{DIT}^{lt} \right)^{\frac{1}{2}} \right]^{\frac{1}{T+1}} = \left(P_{ICCDI}^{0t} P_{DICCDI}^{0t} \right)^{\frac{1}{2}} = P_{CCDI}^{0t} \left(\Omega_{SI}^{0t} \Omega_{DI}^{0t} \right)^{\frac{1}{2}}, \tag{18}$$

where the quality-adjustment factor is equal to the geometric average of Ω_{SI}^{0t} and Ω_{DI}^{0t}.

Given that the bilateral Törnqvist price indexes satisfy the time reversal test and are equal to 1 when the time periods "compared" are identical, it can be shown that for the sample period $0,...,T$, which has a window length of $T+1$ periods, we only need to estimate $T(T+1)/2$ different bilateral matched-item or imputation Törnqvist indexes to construct the CCDI, ICCDI and DICCDI indexes. Note that in De Haan and Krsinich's (2014) ITGEKS method also $T(T+1)/2$ (bilateral time dummy) hedonic regressions must be run whereas in the present (D)ICCDI context only $T+1$ regressions have to be run.

Empirical work has shown that matched-item CCDI indexes (and other matched-item multilateral indexes) can be sensitive to the choice of window length (ABS, 2017). One of the causes seemed to be clearance sales, i.e. disappearing items that are sold at an usually low price and with a relatively large expenditure share. We expect (D)ICCDI indexes to be less sensitive to clearance sales, and to the choice of window length, than their matched-item counterparts due to the imputations made for the missing prices, but ultimately this is an empirical matter.

To be able to include strongly seasonal items, the window should be at least 13 months long (or 5 quarters in case of a quarterly CPI). However, in multilateral indexes, past price movements affect measured recent price changes so that recent price changes become less "characteristic" as the window length grows. In order to restrict the loss of characteristicity, the window should not be too long, perhaps not exceeding 25 months (9 quarters).

A disadvantage of multilateral methods is that when new data becomes available and the index is estimated on the larger data set, previous estimates will be revised. A number of extension methods have been proposed to deal with revisions. The Appendix gives a brief overview of these methods, drawing heavily from Van Kints, De Haan and Webster (2019). We are in favor of a rolling-window approach with a mean splice, but other approaches can be considered.

The construction of (rolling window) matched-item CCDI indexes should not be a problem. Several agencies have already experimented with these indexes and may be willing to share their experiences and software with others. An R package (IndexNumR) for calculating a variety of matched-item price indexes, including CCDI, is available at https://CRAN.R-project.org/package=/IndexNumR. It should be easy to modify and re-use the software for the estimation of ICCDI and DICCDI indexes.

5. Item definition and relaunches

The CPI Manual (ILO et al., 2004) recommends the use of unit values as prices and unit value indexes as price indexes for homogeneous goods; see also Balk (1998). Diewert, Fox and De Haan (2016) argued that the length of the time periods for which unit values are calculated should be in line with the publication frequency of the CPI, i.e. a monthly CPI requires monthly unit values whereas a quarterly CPI requires quarterly unit values. Controlling for store type can be important to attain homogeneity if the service provided along with the purchase of the good differs across outlets (Ivancic and Fox, 2013).

Many product varieties have a barcode and a corresponding Global Trade Item Number (GTIN). Barcodes relate to the products that consumers actually purchase and, combined with store type, define homogeneous items. Since GTIN is always available in scanner data sets, calculating unit values at the barcode level for a single store, store type or retail chain is straightforward.

A potential problem is that barcodes may change even if the items stay the same from the consumers' perspective, for instance when there is a slight change in the type of packaging. Price changes that occur during such product *relaunches* will be excluded from matched-item price indexes if items are identified by barcode (Reinsdorf, 1999; De Haan, 2002; Chessa, 2016). Clothing is a famous example: due to seasonality, many items are unavailable for some time before re-appearing on the market, often with new barcodes. Matched barcode price indexes for clothing will have a downward bias due to the continuous price declines observed for individual barcodes during their availability. This is why Greenlees and McClelland (2010) found the matched-item GEKS(-Fisher) index for women's tops to be severely downward biased.

Chessa (2016) did not identify items by barcode but cross-classified a number of observable characteristics and calculated unit values across all the barcodes belonging to the strata or "groups". If enough characteristics are available in the data set, then this stratification approach defines relatively homogeneous items. The use of Stock Keeping Units (SKUs) in for example Australia (ABS, 2017) is in fact a detailed group method as SKUs normally comprise different barcodes that represent very similar goods. But if characteristics information is limited, stratification can lead to heterogeneity. The "true" fraction of matched items will then be overstated and *unit value bias* likely occurs; see also Dalén (2017).

Relaunches point to a potential problem with matched-item price indexes: these indexes do not explicitly account for new and disappearing items. In a symmetric price index with imputations for the missing prices, such as the imputation Törnqvist index, hence in the (D)ICCDI index, relaunches will be captured. In this case there is no need for relying on a group approach, and the use of barcodes to define items is appropriate, provided of course that the imputations lead to sensible estimated price relatives for the unmatched items.

Now suppose we define items by barcode. If a relaunch occurs where item *i* that leaves the sample in period 0 at a price p_i^0 is replaced in period *t* by a fully comparable item *j* that enters at a price p_j^t, it is easily understood from equation (6) that the price relatives for *i* and *j* should both be equal to p_j^t / p_i^0 because *i* and *j* are then effectively treated as a single matched item. Suppose furthermore that a number of characteristics are available to run separate hedonic regressions in each period. Given that *i* and *j* have identical characteristics, we find $\hat{p}_i^t / \hat{p}_i^0 = \hat{p}_j^t / \hat{p}_j^0$ in the double imputation Törnqvist and DICCDI indexes. Note that with single imputation, the estimated price relatives for *i* and *j* will generally not be the same.

Even with double imputation there is of course no guarantee that the estimated price relatives are equal to p_j^t / p_i^0. Yet, we expect the double imputation approach to work satisfactorily. Expenditure-share weighting ensures that items *i* and *j* affect the regression results, and thus the predicted prices, according to economic importance. So when relaunches are important in terms of expenditure shares, the predicted prices will reflect that.

Why then did Chessa (2016) propose a stratification approach? This was related to the choice of multilateral method: Statistics Netherlands has recently implemented the GK (Geary, 1958; Khamis, 1972) method for the treatment of scanner data. As the GK method does not depend on imputations for the missing prices of unmatched items (Diewert, 2018), broadening the item definition by grouping GTINs or SKUs that are similar in terms of the available characteristics is a straightforward way of incorporating characteristics information to deal with the problem of relaunches. Note that the same applies to the multilateral TPD (Time Product Dummy) method, which is adapted from Summers' (1973) Country Product Dummy method. Both Krsinich (2016) and De Haan (2015) proposed using TPD when characteristics information is lacking, albeit partly for different reasons.

Diewert and Fox (2017) used the economic approach to index number theory to assess multilateral methods and argued that restrictive assumptions about consumers' preferences are necessary to defend GK and TPD. More specifically, the GK method is *additive*, and the TPD method is approximately additive, while economic theory does in general not support additivity. On the other hand, most empirical research found little differences between GK, TPD and matched-item CCDI indexes if items were identified by GTIN or SKU. This may be somewhat surprising, but it is reassuring for statistical agencies that apply GK or TPD.

6. An empirical example

For our empirical example, we use 17 months of scanner data on TVs sold by a Dutch retail chain; note that online sales are excluded. Items are identified by European Article Number, the European version of GTIN, and item prices are calculated as unit values across all the stores belonging to this retail chain. The following categorical variables serve as explanatory variables in the hedonic regressions: brand (6 categories, including a category for "low quality"), screen size (7 categories), screen resolution (3 categories), curved screen (yes/no), screen type (2 categories), processor type (4 categories), energy class (4 categories), Internet access (yes/no), video on demand (yes/no), 3D (yes/no), DLNA (yes/no), and satellite receiver (yes/no).

Figure 1 contains six price indexes: the chained Törnqvist and five multilateral indexes, i.e. the matched-item CCDI index, the single imputation CCDI (ICCDI) index, the ITGEKS index, which is De Haan and Krsinich's (2014) single imputation CCDI variant, the double imputation CCDI (DICCDI) index, and the weighted Time Dummy Hedonic (TDH) index. All the multilateral indexes are estimated simultaneously on the data for the entire sample period. The TDH index follows the standard approach where a model with dummy variables for the different periods (except period 0) is estimated on the pooled data, in this case by expenditure-share weighted regression. Just like the ICCDI, DICCDI and ITGEKS indexes, the TDH index is explicitly adjusted for quality changes and transitive. R squared for the multi-period TDH regression is 0.943. The R squared values for the monthly regressions to construct the ICCDI and DICCDI indexes are even higher, between 0.949 and 0.981. The regression results can be obtained from the authors on request.

Figure 1: Price indexes for TVs (January 2015= 100)

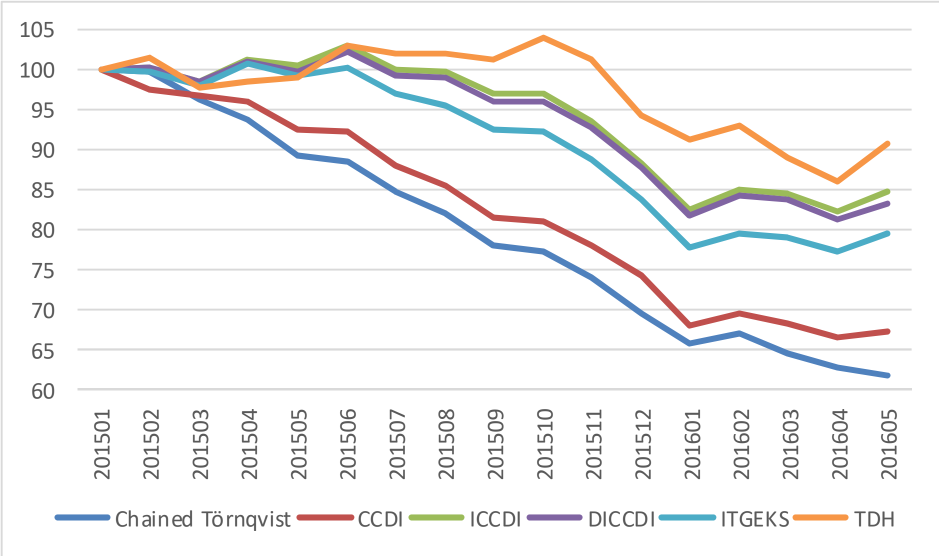


Figure 1 shows that different methods lead to very different results. The chained Törnqvist index measures a price decrease of almost 40% in just 17 months and seems to suffer from chain drift as it sits below the CCDI index. The most interesting result is the big impact of the imputations for the missing prices of new and disappearing items. Compared with the ICCDI index, the matched-item CCDI index has a large downward bias. How can this result be explained?

The retailer's pricing strategy could be a major driver. Since prices of high-tech consumer goods at the model/barcode level often have a downward trend, matched-item price indexes typically decrease. Suppose retailers follow a price skimming strategy for new models and an inventory cleaning strategy for obsolete models. The prices of new models will then be relatively high given their characteristics and those of old models relatively low. In other words, the residuals from the hedonic regressions are likely to be positive for new models and negative for old models (Silver and Heravi, 2005; De Haan, Hendriks and Scholz, 2016).

Our empirical findings are consistent with these pricing strategies. The easiest way to illustrate this, in particular for new items, is to start with the single imputation Törnqvist price index given by (3). Due to the relatively high prices of new TV models, the geometric Paasche price index in the numerator of the third component is likely to exceed the geometric Paasche price index for the matched models in the denominator; N^{0t} is expected to be greater than 1. So, with price skimming the introduction of new

TV models typically has an upward effect on the imputation Törnqvist price index. We expect to observe this upward effect for many bilateral comparisons across the sample period so that it carries over to the ICCDI index. Note though that, as was mentioned in Section 3, part of this upward effect might be due to a lack of adjusting for unobserved features embodied in new models.

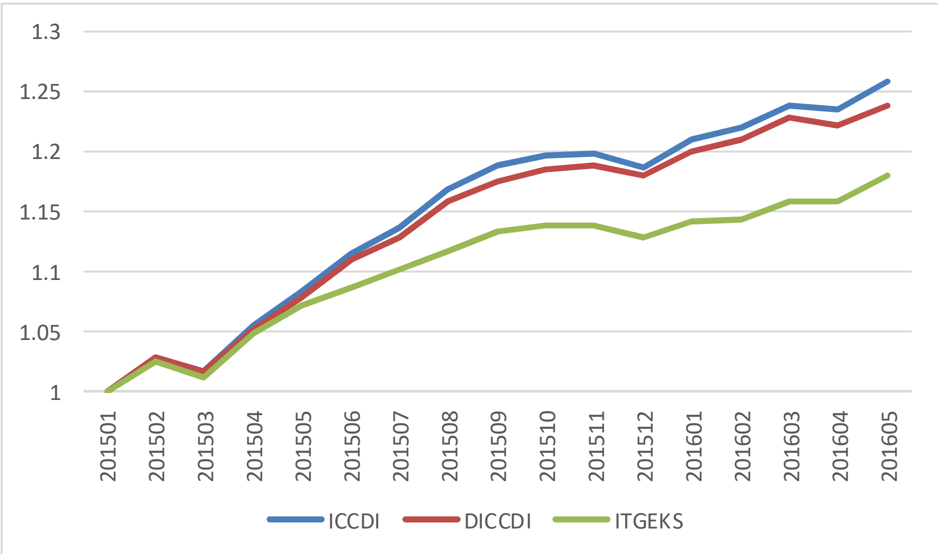
The difference between the ICCDI index and the DICCDI index in Figure 1 is also consistent with a price skimming and inventory cleaning strategy. Equation (11) in Section 3 tells us that in this case the bilateral double imputation Törnqvist index will sit below the single imputation Törnqvist index if new items generally had relatively high prices and disappearing items had relatively low prices. Again, this effect carries over to the ICCDI index and the DICCDI index. Yet the difference between the two indexes in our example is very small; the hedonic model fits the data so well that the predicted values for the unmatched items are close to the observed prices. This also suggests that the effect of any unobserved features is limited.

Another interesting finding in Figure 1 is the difference between the ICCDI (and DICCDI) index and the ITGEKS index. De Haan and Krsinich’s (2014) ITGEKS index is a variant of the single imputation CCDI method where the missing prices are imputed using hedonic regressions where the data for the bilateral comparisons is pooled instead of using a separate regression in each month; see Section 3. The reason for the ITGEKS index to sit below the ICCDI index is, as we have checked, that the imputed prices for the new items based on the pooled (bilateral) regressions are generally higher than the imputed prices based on the monthly regressions and those for the disappearing items lower. This is most likely to stem from the different regression weights used. Compared with the monthly ICCDI regressions, the impact of new and disappearing items in the pooled ITGEKS regressions is reduced because only half of their expenditure shares are used as weights. Since the observed prices of new items are relatively high and those of disappearing items relatively low, plus the fact that prices typically decrease over time, we would indeed expect a downward effect on the ITGEKS index.

Constraining the characteristics’ parameters to be fixed over time in the ITGEKS regressions can impact on the difference between the ITGEKS index and ICCDI index as well. The coefficients from the monthly regressions are, however, quite stable and in line with the coefficients from the ITGEKS regressions, and so we believe the impact is small. Note that, because the parameters are fixed in each bilateral comparison, we are essentially assuming parameter fixity across the entire window. The same assumption underlies the multi-period TDH method. As it turns out, the expenditure-share weighted TDH index seems to have a significant upward bias. This is not the place to discuss the issue in any detail, but the bias must be related to the index number formula behind the weighted TDH method.

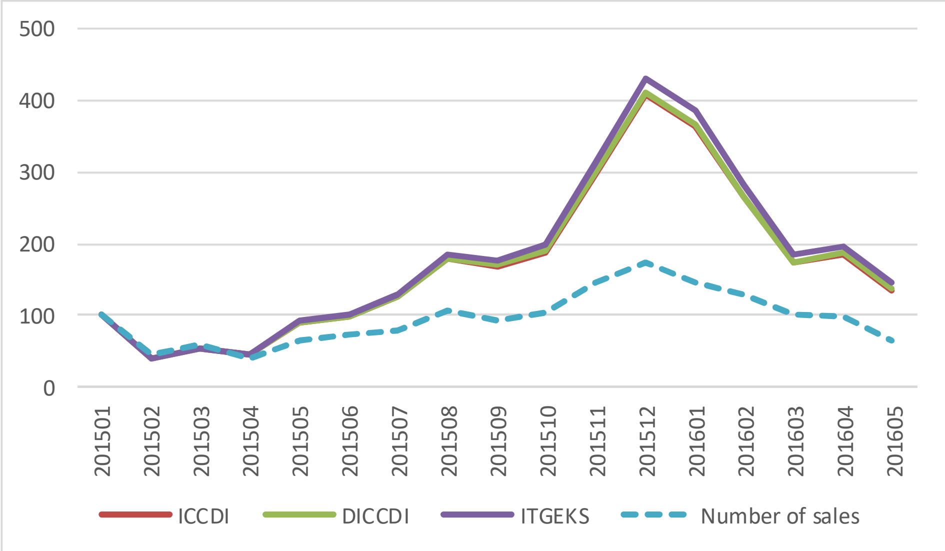
Figure 2 summarizes the impact of the imputations for the unmatched items by comparing the quality-adjustment factors for the ICCDI, DICCDI and ITGEKS indexes, which were defined in Section 4 as the ratio of these indexes and their matched-item counterpart, the CCDI index. As changes in the quality mix are already incorporated in the CCDI index, the quality-adjustment factor does not tell us whether average quality of TVs sold has improved or deteriorated. A value greater than 1 merely means that the implicit quantity index – the ratio of the value index and the price index – will be lower than the quantity index obtained when using the CCDI index.

Figure 2: Quality-adjustment factors



The implicit quantity indexes for IDDCI, DICCDI and ITGEKS shown in Figure 3 all point to a huge quantity increase in November and December of 2015. The dotted line, showing the index of total number of TVs sold, can be viewed as a quantity index that is not adjusted for quality (mix) changes. Thus, Figure 3 suggests there has been a large improvement in average quality of TVs sold, due to compositional change as well as the introduction of new items and the disappearance of “old” items.

Figure 3: Implicit quantity indexes (January 2015= 100)



From the second expression on the right-hand side of equation (14) it follows that the ICCDI index in this empirical example, with $T=16$, is equal to the unweighted geometric mean of 17 (linked Törnqvist) price indexes going from month 0 to month t given by $P_{IT,l}^{0t}=P_{IT}^{0l}P_{IT}^{lt}$ with link months $l=0,...,16$, i.e. $P_{ICCDI}^{0t}=\prod_{l=0}^{16}\left(P_{IT,l}^{0t}\right)^{1/17}$. Recall that $P_{IT,0}^{0t}=P_{IT,t}^{0t}=P_{IT}^{0t}$; in each month the direct single imputation Törnqvist price index "counts twice" in the calculation of the ICCDI index. The 17 constituent price indexes for ICCDI are plotted in Figure 4. The spread of these indexes is substantial; in the final month ($T=16$, May 2016) the index numbers range from 82.70 for $l=12$ to 100.54 for $l=0$ and $l=16$. The pattern of most of the indexes is broadly similar though, and there seems to be no systematic tendency in index levels to increase or decrease when the link month l goes from 0 to 16.

We do not show the 17 constituent indexes for DICCDI, because they are almost identical to those for ICCDI. This result confirms that double imputation does not add much when the hedonic model fits the data very well, i.e. when the R squared values of the monthly regressions are high. In Figure 5 we do show a similar graph for CCDI. The spread of the 17 constituent price indexes – in this case linked matched-item Törnqvist price indexes $P_{MT,l}^{0t}=P_{MT}^{0l}P_{MT}^{lt}$ – for the link periods $l=0,...,15$ appears to be relatively small. The unusual behavior of $P_{MT,16}^{0t}$ is due to an increase of clearance sales in $T=16$, i.e. items sold at very low prices which disappear in $T=17$ (June 2016; not included in the data set utilized).

Figure 4: Constituent indexes of ICCDI (January 2015= 100)

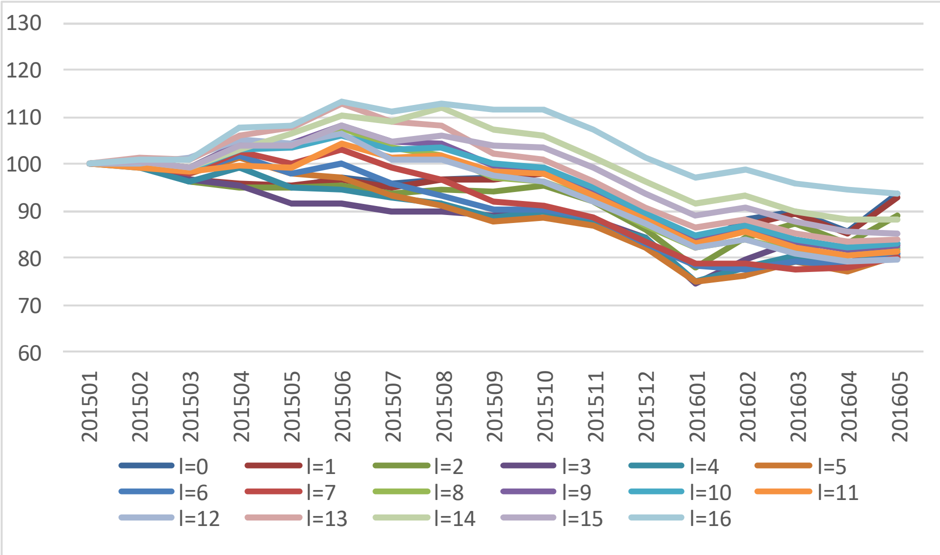
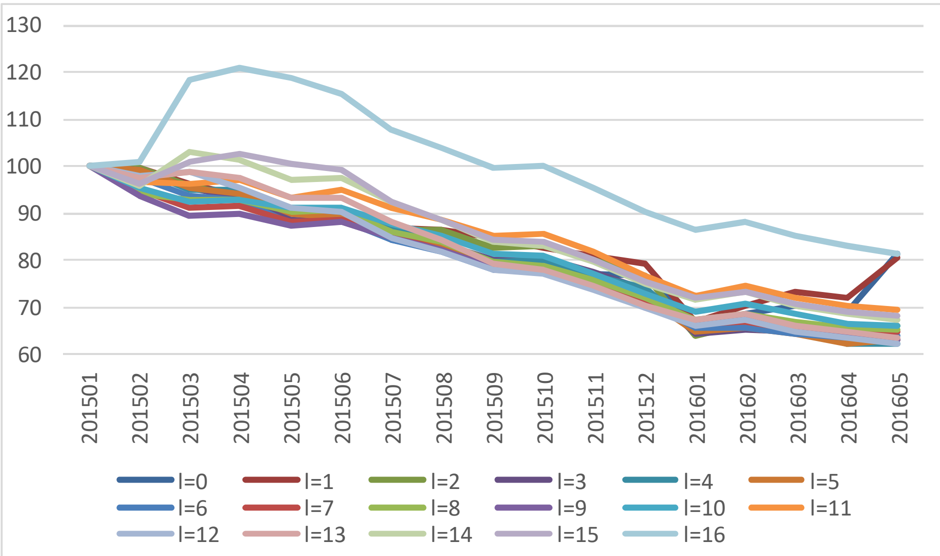


Figure 5: Constituent indexes of CCDI (January 2015= 100)



7. Discussion, conclusions and future work

The appropriate way to identify "homogeneous items" for the treatment of scanner data has been a topic for discussion in Europe, particularly in relation to the implementation of multilateral index number methods in the European Harmonized Index of Consumer Prices. What can be learned from this discussion? First of all, it is important to realize that homogeneity is defined by the consumers' perception of what constitutes individual items; homogeneity obviously does not depend on the choice of multilateral method, or any method for that matter. The issue of how to differentiate between items, as well as the proper concept of price, is as old as price and quantity measurement itself and has little to do with the use of scanner data, the implementation of new index methods, or quality adjustment (through hedonic regression or otherwise).

Of course in statistical practice we need to operationalize the way in which items are identified. Most people will agree that differentiating between items should be done by looking at the total set of characteristics that consumers value. The practical problem is that the "true" set of characteristics is unknown and that – even if we knew the "true' set – it is unlikely that all of the relevant characteristics information will be available to the statistical agency. The discussion in Europe has focused on trying to strike a balance between homogeneity and matching over time. The idea is to find an optimal solution to the *trade-off* between an increase in heterogeneity when using fewer characteristics than required to attain full homogeneity and a loss of matches in the data when using many characteristics. For this purpose, Chessa (2018) developed a method (Match Adjusted R Squared) which, starting from the set of available characteristics, derives the "optimal" grouping of GTINs into items to be used in index compilation.

Grouping GTINs is understandable because Statistics Netherlands decided to use the GK method: this method cannot explicitly adjust for quality change in the sense that it is incompatible with the imputation of missing prices, and so to deal with churn at the barcode level, groups are formed. However, the choice of method now affects the item definition while items are ideally defined independently. Also, stratification that yields heterogeneous items is at odds with the principle of comparing like with like. From an index number point of view, we would like to have a transitive quality-adjusted price index where the above trade-off is not an issue.

In this paper we revisited such a price index: the multilateral CCDI index with explicit imputations for the missing prices of unmatched new and disappearing items. In order to compare like with like and avoid unit value bias, we propose identifying items by GTIN, which relates to the products that are actually offered for sale, or SKU, which is a highly detailed group approach. A high churn rate resulting from using GTIN/SKU as item identifier should not be a problem, because the missing prices are imputed using hedonic regression.

Not everyone will endorse the use of hedonic regression. Diewert and Feenstra (2017), Diewert, Fox and Schreyer (2017) and Diewert (2018a) argued that the missing prices should be interpreted as Hicksian *reservation prices*, the unobservable prices that would drive down demand for the items to zero. We are hesitant to apply the reservation prices approach to the level of product varieties and tend to agree with Reinsdorf and Schreyer (2017) that this approach is meant for completely new goods rather than new variants of an existing product. The hedonic imputation approach does not assume that the *demand* for a new variety in the earlier period was zero but that it was unavailable due to *supply* restrictions, which could be technological constraints, strategic choices of manufacturers (such as delaying the introduction of producible models), models being temporarily out of stock, etc. This approach tries to estimate the price of the new variety in the earlier period as if it had already been available, and this price will not be very different from the prices of broadly comparable varieties.

The choice between the two concepts can have far reaching implications for the price change measured. The reservation prices approach typically measures a substantial price decrease for new items whereas the imputations approach could lead to a modest price decrease or, as in our empirical example for TVs, even a price increase.

Even if reservation prices were useful in the case of broadly comparable items, it is difficult to see how this approach would fit into official price measurement given the complexity of the econometric estimation. Diewert (2018a) mentioned the same issue and proposed a simple non-hedonic imputations approach, perhaps inspired by recent work suggesting that – compared with traditional quality adjustment methods – hedonic regression can lead to inaccurate results (Adams and Klayman, 2018). Diewert proposed carrying forward the last observed price for a disappearing item and carrying backward the first observed price for a new item and then adjusting these prices for inflation. This is an example of traditional methods to impute prices for *temporarily unavailable items*. It can also be thought of as "implicit" quality adjustment as opposed to explicit quality adjustment where item characteristics play a role.

Diewert's (2018a) proposal is dependent on the choice of measure to adjust the carried forward and carried backward prices for inflation. To some extent, this choice is arbitrary. For the bilateral imputation Törnqvist price index, the obvious choice seems to be the matched-item Törnqvist price index, but other choices are possible and not necessarily worse. Admittedly, hedonic imputations are somewhat arbitrary too as they depend on the choice of functional form and the characteristics included in the model. Yet, in contrast to Diewert's method, hedonic imputations explicitly adjust for quality change and are likely to significantly reduce the problem of relaunches when items are identified by barcode.

In our empirical example we suggested that the results for the hedonic indexes for TVs were related to dumping and price skimming. But these pricing strategies are not confined to goods where technical progress is important; they have also been found for products sold in grocery stores (Melser and Syed, 2016) and for clothing and other fashion goods. Product relaunching can be viewed as an extreme case. Relaunches seem to occur frequently for e.g. personal care items sold in drugstores with the prices of the replacement items often being higher than those of the replaced items (Chessa, 2016). The retailers may have some degree of market power and consumers are unable, or may be unwilling, to substitute away from the replacement items because of brand loyalty or transaction costs if they have to go to another store.

Data on characteristics permitting, future empirical work could examine whether the (D)CCDI method does indeed appropriately deal with the problem of relaunches for products such as personal care items, as we believe it does. It would also be interesting to compare the (D)CCDI index with the *quality-adjusted unit value index* examined by De Haan and Krsinich (2018). The latter index can be viewed as the hedonic counterpart to the GK index, where the quality-adjustment parameters are explicitly estimated using a multi-period TDH hedonic regression model rather than implicitly derived, as is the case in the GK method (Van Kints, De Haan and Webster, 2019). De Haan and Krsinich (2018) argued that the weighted TDH index produces an accurate approximation to the quality-adjusted unit value index.

An important topic for future research is the treatment of new characteristics. An additional regression must be run to estimate the rolling-window (D)ICCDI index when data for the next period becomes available. Suppose in period T+1 a new characteristic is introduced which should be included in the hedonic model. Like any imputation price index, the (D)ICCDI method faces the problem that, since separate hedonic regressions are run for all periods, the missing prices of the (new) items with the new characteristic cannot be estimated (Crawford and Neary, 2019). Omitting the new characteristic from the period T+1 regression likely produces biased imputations and it is uncertain whether double imputation would remove most of the bias in the index.

Appendix: The treatment of revisions

This Appendix provides a very brief overview of the methods that have been proposed to deal with revisions in multilateral price indexes when the time series is extended.

Rolling window methods estimate the indexes on a window with a fixed length. The estimation window is shifted forwards each period, and the results from the latest window are then spliced onto the existing time series. This can for example be done by splicing the most recent period-on-period index movement onto the latest index number, suggested by Ivancic, Diewert and Fox (2011). An alternative to this movement splice is Krsinich's (2016) window splice, which splices the most recent index change across the entire window onto the index of T-1 periods ago.

These extension methods splice price movements onto a single period. Diewert and Fox (2017) argued that, as all periods are equally valid, it seems preferable to use a rolling-window *mean splice* by taking the geometric mean of the price indexes obtained by using every possible splicing period. This makes the result independent of the choice of splicing period.

The annually chained *direct extension* method (Chessa, 2016) constructs short-term multilateral index series of, say, 13 months, starting in e.g. December and ending in December of the next year, and chain links them in December of each year to arrive at a long-term time series. Note that the length of the estimation window for the short-term indexes is extended each month – the index for January in the short-term series is estimated on two months of data, and so forth, until in December thirteen months of data is used.

References

Adams, B. and A. Klayman (2018), "Cross-Validation of Quality-Adjustment Methods for Price Indexes", *Monthly Labor Review*, June 2018, Bureau of Labor Statistics, US.

Australian Bureau of Statistics (ABS) (2017), "An Implementation Plan to Maximise the Use of Transactions Data in the CPI", Information Paper 6401.0.60.004, ABS, Canberra, Australia.

Balk, B.M. (1996), "A Comparison of Ten Methods for Multilateral International Price and Volume Comparison", *Journal of Official Statistics* 12, 199-222.

Balk, B.M. (1998), "On the Use of Unit Value Indices as Consumer Price Sub-indices", Paper presented at the 4th meeting of the Ottawa Group, 22-24 April 1998, Washington DC, US.

Balk, B.M. (2008), *Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference*. New York: Cambridge University Press.

Caves, D.W., L.R. Christensen and W.E. Diewert (1982), "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity", *Econometrica* 50, 1393-1414.

Chessa, A.G. (2016), "A New Methodology for Processing Scanner Data in the Dutch CPI", *Eurona* 1, 49-69.

Chessa, A.G. (2018), "MARS: A Method for Defining Products and Linking Barcodes of Item Relaunches", Research Paper, Statistics Netherlands, The Hague.

Crawford, I. and P. Neary (2019), "New Characteristics and Hedonic Price Index Numbers", Discussion Paper DP13445, Centre for Economic Policy Research, London, UK.

Dalén, J. (2017), "Unit Values and Aggregation in Scanner Data – Towards a Best Practice", Paper presented at the 15th meeting of the Ottawa Group, 10-12 May 2017, Altville am Rhein, Germany.

Diewert, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics* 4, 115-145.

Diewert, W.E. (1999), "Axiomatic and Economic Approaches to International Comparisons", pp. 13-87 in A. Heston and R.E. Lipsey (eds.), *International and Interarea Comparisons of Income, Output and Prices*. Chicago: University of Chicago Press.

Diewert, W.E. (2018a), "Scanner Data, Elementary Price Indexes and the Chain Drift Problem", Discussion Paper 18-06, Vancouver School of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W.E. (2018b), "Quality Adjustment and Hedonics: A Unified Approach", Discussion Paper 18-11, Vancouver School of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W.E., S. Heravi and M. Silver (2009), "Hedonic Imputation versus Time Dummy Hedonic Indexes", pp. 87-116 in W.E. Diewert, J. Greenlees and C. Hulten (eds.), *Price Index Concepts and Measurement*. Chicago: University of Chicago Press.

Diewert, W.E., K.J. Fox and J. de Haan (2016), "A Newly Identified Source of Potential CPI Bias: Weekly versus Monthly Unit Value Price Indexes", *Economics Letters* 141, 169-172.

Diewert, W.E. and K.J. Fox (2017), "Substitution Bias in Multilateral Methods for CPI Construction Using Scanner Data", Discussion Paper 17-02, Vancouver School of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W.E., K.J. Fox and P. Schreyer (2018), "The Digital Economy, New Products and Consumer Welfare", ESCoE Discussion Paper 2018-16, Economic Statistics Centre of Excellence, National Institute of Economic and Social Research, London, UK.

Eltetö, Ö. and P. Köves (1964), "On an Index Computation Problem in International Comparisons" [in Hungarian], *Statiztikai Szemle* 42, 507-518.

Feenstra, R.C. and M.D. Shapiro (2003), "High-Frequency Substitution and the Measurement of Price Indexes", pp. 123-150 in R.C. Feenstra and M.D. Shapiro (eds.), *Scanner Data and Price Indexes*. Chicago: University of Chicago Press.

Geary, R.C. (1958), "A Note on the Comparison of Exchange Rates and Purchasing Power between Countries", *Journal of the Royal Statistical Society A* 121, 97-99.

Gini C. (1931), "On the Circular Test of Index Numbers", *International Review of Statistics* 9, 3-25.

Greenlees, J. and R. McClelland (2010), "Superlative and Regression-Based Consumer Price Indexes for Apparel Using U.S. Scanner Data", Paper presented at the General Conference of the IARIW, 27 August 2010, St. Gallen, Switzerland.

de Haan, J. (2002), "Generalised Fisher Price Indexes and the Use of Scanner Data in the Consumer Price Index", *Journal of Official Statistics* 18, 61-85.

de Haan, J. (2004a), "The Time Dummy Index as a Special Case of the Imputation Törnqvist Index", Paper presented at the 8th meeting of the Ottawa Group, 23-25 August 2004, Helsinki, Finland.

de Haan, J. (2004b), "Direct and Indirect Time Dummy Approaches to Hedonic Price Measurement", *Journal of Economic and Social Measurement* 29, 427-443.

de Haan, J. (2010), "Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and 'Re-pricing' Methods", *Journal of Economics and Statistics* 230, 772-791.

de Haan, J. (2015), "A Framework for Large Scale Use of Scanner Data in the Dutch CPI", Paper presented at the 14th meeting of the Ottawa Group, 20-22 May 2015, Tokyo, Japan.

de Haan, J. and H.A. van der Grient (2011), "Eliminating Chain Drift in Price Indexes Based on Scanner Data", *Journal of Econometrics* 161, 36-46.

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

de Haan, J., R. Hendriks and M. Scholz (2016), "A Comparison of Weighted Time-Product Dummy and Time Dummy Hedonic Indexes", Graz Economic Papers 2016-13, Department of Economics, University of Graz, Austria.

de Haan, J. and F. Krsinich (2018), "Time Dummy Hedonic and Quality-Adjusted Unit Value Indexes: Do They Really Differ?", *Review of Income and Wealth* 64, 757-776.

Hill, R. and D. Melser (2008), "Hedonic Imputation and the Price Index Problem: An Application to Housing", *Economic Inquiry* 46, 593-609.

ILO/IMF/OECD/UNECE/Eurostat/The World Bank (2004), *Consumer Price Index Manual: Theory and Practice*. ILO Publications, Geneva.

Inklaar, R. and W.E. Diewert (2016), "Measuring Industry Productivity and Cross Country Convergence", *Journal of Econometrics* 192, 426-433.

Ivancic, L. (2007), "Scanner Data and the Construction of Price Indices", PhD thesis, School of Economics, The University of New South Wales, Sydney.

Ivancic, L., W.E. Diewert and K.J. Fox (2011), "Scanner Data, Time Aggregation and the Construction of Price Indexes", *Journal of Econometrics* 161, 24-35.

Ivancic, L. and K.J. Fox (2013), "Understanding Price Variation across Stores and Supermarket Chains: Some Implications for CPI Aggregation Methods", *Review of Income and Wealth* 59, 629-647.

Khamis, S.H. (1972), "A New System of Index Numbers for National and International Purposes", *Journal of the Royal Statistical Society A* 135, 96-121.

van Kints, M., J. de Haan and M. Webster (2019), "Utilising Big Data and Multilateral Index Methods to Produce the Australian CPI: Theory, Implementation and Empirical Results", *Statistical Journal of the IAOS* (forthcoming).

Krsinich, F. (2016), "The FEWS Index: Fixed Effects with a Window Splice", *Journal of Official Statistics* 32, 375-404.

Melser, D. and I.A. Syed (2016), "Life Cycle Price Trends and Product Replacement: Implications for the Measurement of Inflation", *Review of Income and Wealth* 62, 509-533.

Ohta, M. and Z. Griliches (1976), "Automobile Prices Revisited: Extensions of the Hedonic Hypothesis", pp. 325-390 in *Household Production and Consumption*, N. Terleckyj (ed.). New York: Columbia University Press.

Pakes, A. (2003), "A Reconsideration of Hedonic Price Indexes with an Application to PCs", *American Economic Review* 93, 1578-1596.

Reinsdorf, M.B. (1999), "Using Scanner Data to Construct CPI Basic Component Indexes", *Journal of Business & Economic Statistics* 17, 152-160.

Reinsdorf, M. and P. Schreyer (2017), "Measuring Consumer Inflation in a Digital Economy", Paper presented at the ESCoE Conference, 16-17 May 2018, London, UK.

Silver, M. and S. Heravi (2005), "A Failure in the Measurement of Inflation – Results from a Hedonic and Matched Experiment Using Scanner Data", *Journal of Business & Economic Statistics* 23, 152-160.

Statistics New Zealand (2014), "Measuring Price Change for Consumer Electronics Using Scanner Data". Available from www.stats.govt.nz.

Summers, R. (1973), "International Price Comparisons Based Upon Incomplete Data", *Review of Income and Wealth* 19, 1-16.

Szulc, B. (1964), "Index Numbers of Multilateral Regional Comparisons" [in Polish], *Przeglad Statysticzny* 3, 239-254.