A Framework for Large Scale Use of Scanner Data in the Dutch CPI

Jan de Haan[a]

3 April 2015

Abstract: Statistics Netherlands is planning to use scanner data on a large scale for the compilation of the CPI, covering supermarkets, department stores, do-it-yourself stores, etc. Ideally, to make the production process as efficient as possible, a limited number of fully or semi-automated methods would be used. The purpose of this paper is to propose a framework supporting these plans. Our basic aggregation formula is what we refer to as a "quality-adjusted unit value index", which is equal to the value index divided by a quantity index that is defined as the ratio of quality-adjusted or standardized quantities. Time dummy regression models play an important role in the estimation of the quality-adjustment (standardization) factors. There are two extreme cases. If information on all relevant item characteristics is available, then the use of time dummy hedonic models is preferred. When characteristics information is lacking, the use of time-product dummy (fixed effects) models is proposed. We also discuss a number of issues that need to be resolved before our ideas can be implemented in CPI production, such as the definition of homogeneity, treatment of revisions and choice of estimation window.

Keywords: fixed effects, hedonic regression, quality adjustment, unit values.

JEL Classification: C43, E31.

1. Introduction

More than twenty years ago, Saglio (1994) presented a paper at the first Ottawa Group meeting on the use of scanner data to construct unit values and price indexes. Today, a large body of literature is available on CPI measurement issues related to scanner data, covering topics such as item sampling, choice of index number formula at the lower and upper aggregation level, quality adjustment, and treatment of sales.[1]

In spite of all the research that has been done, so far only a handful of countries actually implemented scanner data into their CPI. Australia, the Netherlands, Norway, Sweden, and Switzerland have included scanner data from supermarkets, using different methods and practices (Dalén, 2014). The current Dutch method is described in van der Grient and de Haan (2010) and compared to an alternative method in van der Grient and de Haan (2011). New Zealand recently introduced scanner data for consumer electronics using a method developed by de Haan and Krsinich (2014a); see Statistics New Zealand (2014).

Statistics Netherlands wants to expand the use of scanner data in the CPI beyond supermarkets and to cover also department stores, do-it-yourself stores, etc. Ideally, to make production as efficient as possible, a limited number of fully or semi-automated methods would be used. The purpose of this paper is to propose a framework supporting these plans. Our basic aggregation formula is the "quality-adjusted unit value index", which is equal to the value index divided by a quantity index that is defined as the ratio of quality-adjusted or standardized quantities. Dalén (1998) seems to have been the first to describe a quality-adjusted unit value index. De Haan (2004a) suggested a slightly different formulation. De Haan and Krsinich (2014b) showed how time dummy hedonic regressions can be used to estimate the quality adjustment factors. Their work forms the basis for the methods proposed in the present paper.

---

[1] A large part of this research was presented at various meetings of the Ottawa Group. Without trying to be exhaustive, we mention the following studies. Early studies on potential uses of scanner data in the CPI, in particular on the construction of elementary indexes, are Dalén (1997), Hawkes (1997) (1998), de Haan and Opperdoes (1997a,b), Bradley et al. (1997), Reinsdorf (1999), Jain and Caddy (2001), Jain and Abello (2001), Richardson (2001), and de Haan (2002). De Haan, Schut and Opperdoes (1999) looked into sampling issues. Ioannides and Silver (1997), Silver, Ioannides and Haworth (1997), Okamoto and Salou (2001), and Silver and Heravi (2005) addressed hedonic quality adjustments. Ivancic, Diewert and Fox (2011) and de Haan and van der Grient (2011) investigated methods for eliminating chain drift due to promotional sales. Note that some of the references are publications in journals; preliminary versions were presented at Ottawa Group meetings.

The paper is structured as follows. In section 2, we start with the decomposition of a value change into a price index and a quantity index, and then we outline the idea behind a quality-adjusted unit value index for broadly comparable items. This is done for two time periods in the static case, i.e. in a situation with matched items only. We also address the estimation of the quality adjustment factors using a basic economic method, leading to standard Laspeyres and Paasche indexes, and the econometric time dummy hedonic method mentioned above.

Section 3 deals with the two period dynamic situation where the set of broadly comparable items changes over time as a result of new items appearing on the market and "old" items disappearing.

In section 4, the approach is extended to more than two time periods. We explain our preference for multilateral methods that pool data across the whole sample period. It is shown that, if expenditure shares are used as regression weights, the multilateral time dummy index is the geometric counterpart of the quality-adjusted unit value index. We argue that the two indexes are likely to be very similar in practice.

Hedonic regressions require information on product characteristics. In section 5 we propose using time-product dummy (fixed effects) models if this information is not available. An issue here is whether or not these models produce truly quality-adjusted price indexes.

In section 6, we discuss a number of issues that need to be taken care of before our ideas can be implemented in CPI production, such as the definition of homogeneity, treatment of revisions and choice of estimation window. Revisions arise from the fact that the results for earlier periods will change when new data is added to the sample and the models are re-estimated. We suggest using a rolling window approach, discuss two possible splicing methods, and address the choice of window length.

In the final section 7, we describe our framework for large scale use of scanner data. The framework exists of choosing the quality-adjusted unit value index as the sole aggregation formula and a five stage procedure for incorporating scanner data into the CPI. We also examine how prices extracted from websites, where quantity information is lacking, would fit into the framework. We end with a research proposal. It includes exploring the potential use of "text mining" and "machine learning" to try and retrieve item characteristics from product descriptions in order to match comparable items in the absence of detailed characteristics information.

2. The two period static case

2.1 The index number problem and the quality-adjusted unit value index

Suppose we have price and quantity data for all items belonging to a particular product. These items or product varieties are broadly similar in that their quality can be described by the same set of characteristics, albeit in different quantities. We assume that a single hedonic model applies to broadly similar items. In this section, we examine the static case with a fixed population or universe of items, as in standard index number theory. This will serve as an introduction to the dynamic situation with new and disappearing items dealt with in section 3. In the present section, and in section 3, we consider two time periods. The many periods case will be addressed in section 4.

Our notation is as follows. U is the fixed set of broadly comparable items. The prices of item i in the base period 0 and the comparison period 1 are denoted by p_i^0 and p_i^1, respectively; q_i^0 and q_i^1 are the corresponding quantities purchased. Both prices and quantities are strictly positive. The total values in the two periods are $V^0 = \sum_{i \in U} p_i^0 q_i^0$ and $V^1 = \sum_{i \in U} p_i^1 q_i^1$. Our aim is to decompose the value ratio $V^{01} = V^1 / V^0$ into a price index P^{01} and a quantity index Q^{01}:

$$V^{01} = P^{01} \times Q^{01}. \tag{1}$$

The choice of P^{01} and Q^{01} is known as the *index number problem*. Computing two out of the three indexes will suffice as the third one can be derived from (1). Usual practice is to construct a price index and then deflate the value index to obtain a quantity index. In some cases, for example health and education in the Dutch national accounts, quantity indexes – or "volume indexes" in their language – are constructed and implicit price indexes are obtained via (1). Our approach is in the latter spirit.

Standard quantity index number formulas are needed since adding up quantities of heterogeneous goods is not meaningful. Yet, for broadly comparable items, adding up standardized or *quality-adjusted quantities* is an appealing approach. Using quality adjustment (standardization) factors, we may be able to express the quantities of all the items in units of a base item and then simply add them up. The ratio of these aggregate quality-adjusted quantities in both periods is a measure of quantity change. Importantly, as we will demonstrate in section 3, this approach can be easily extended to the dynamic situation.

The above can be formalized as follows. Suppose quality adjustment factors $\lambda_{i/b}^0$ and $\lambda_{i/b}^1$ exist which express the quantity purchased of item i in periods 0 and 1, q_i^0 and q_i^1, in $\lambda_{i/b}^0 q_i^0$ and $\lambda_{i/b}^1 q_i^1$ units of an arbitrary item b ($\lambda_{b/b}^0 = \lambda_{b/b}^1 = 1$). Summing over all items, this standardization leads to $\sum_{i \in U} \lambda_{i/b}^0 q_i^0$ and $\sum_{i \in U} \lambda_{i/b}^1 q_i^1$ equivalent units of b. The base item b could be any $i \in U$ or be defined by average characteristics. The ratio of the quality-adjusted quantities in periods 0 and 1 is

$$Q^{01} = \frac{\sum_{i \in U} \lambda_{i/b}^1 q_i^1}{\sum_{i \in U} \lambda_{i/b}^0 q_i^0}. \tag{2}$$

There are two potential issues with the quantity index (2), both resulting from the fact that the quality adjustment factors are not fixed across time but pertain to period 0 in the denominator of (2) and to period 1 in the numerator. The first issue is that the quantity index violates the *identity test*, which is regarded as one of the most important tests any index number formula should satisfy. That is, if the quantities purchased of all the items stay the same ($q_i^0 = q_i^1$ for all $i \in U$), then the quantity index is not necessarily equal to 1. The second issue is that, since the quality adjustment factors can be viewed as "taste parameters", the quantity index (as well as the implicit price index) would be affected by changes in taste. Some people may find this undesirable, at least in the short run.

It can therefore be argued that the quality adjustment factors must be kept fixed across time,[2] and we write the quantity index in generic form as

$$Q^{01} = \frac{\sum_{i \in U} \lambda_{i/b} q_i^1}{\sum_{i \in U} \lambda_{i/b} q_i^0}, \tag{3}$$

where $\lambda_{i/b}$ equals $\lambda_{i/b}^0$, $\lambda_{i/b}^1$, or some average value. Dividing V^{01} by (3) gives

$$P_{QAUV}^{01} = \frac{V^{01}}{Q^{01}} = \frac{\sum_{i \in U} p_i^1 q_i^1 \Big/ \sum_{i \in U} \lambda_{i/b} q_i^1}{\sum_{i \in U} p_i^0 q_i^0 \Big/ \sum_{i \in U} \lambda_{i/b} q_i^0} = \frac{\sum_{i \in U} p_i^1 q_i^1 \Big/ \sum_{i \in U} \tilde{q}_i^1}{\sum_{i \in U} p_i^0 q_i^0 \Big/ \sum_{i \in U} \tilde{q}_i^0}, \tag{4}$$

---

[2] It can also be argued that the above issues are not relevant in the present context. Once the quantities of the items are expressed in units of the base item and "perfect homogeneity" is attained, the axiomatic or test approach to index number theory may no longer seem important. Also, in the longer run we do want taste changes to affect the index. Nevertheless, as we will see below, index number theory (and economic theory) is needed to give us some guidance when it comes to estimating the quality adjustment factors.

where $\tilde{q}_i^0 = \lambda_{i/b} q_i^0$ and $\tilde{q}_i^1 = \lambda_{i/b} q_i^1$ are the quantities purchased of item i in period 0 and 1 measured in equivalent units of the base item b. The price index given by (4) is a ratio of quality-adjusted unit values $\sum_{i \in U} p_i^0 q_i^0 / \sum_{i \in U} \tilde{q}_i^0$ and $\sum_{i \in U} p_i^1 q_i^1 / \sum_{i \in U} \tilde{q}_i^1$. We refer to this price index as a *quality-adjusted unit value index*.

An alternative way to write the quality-adjusted unit value index is

$$P_{QAUV}^{01} = \frac{\sum_{i \in U} (p_i^1 / \lambda_{i/b}) \lambda_{i/b} q_i^1 \Big/ \sum_{i \in U} \tilde{q}_i^1}{\sum_{i \in U} (p_i^0 / \lambda_{i/b}) \lambda_{i/b} q_i^0 \Big/ \sum_{i \in U} \tilde{q}_i^0} = \frac{\sum_{i \in U} \tilde{p}_i^1 \tilde{q}_i^1 \Big/ \sum_{i \in U} \tilde{q}_i^1}{\sum_{i \in U} \tilde{p}_i^0 \tilde{q}_i^0 \Big/ \sum_{i \in U} \tilde{q}_i^0} = \frac{\sum_{i \in U} w_i^1 \tilde{p}_i^1}{\sum_{i \in U} w_i^0 \tilde{p}_i^0}, \tag{5}$$

where $\tilde{p}_i^0 = p_i^0 / \lambda_{i/b}$ and $\tilde{p}_i^1 = p_i^1 / \lambda_{i/b}$ denote the *quality-adjusted prices* of item i (with respect to item b) in periods 0 and 1; $w_i^0 = \tilde{q}_i^0 / \sum_{i \in U} \tilde{q}_i^0$ and $w_i^1 = \tilde{q}_i^1 / \sum_{i \in U} \tilde{q}_i^1$. So the quality-adjusted unit value index can also be viewed as the ratio of weighted quality-adjusted prices in which the quality-adjusted quantities serve as weights.

The second alternative way to write the quality-adjusted unit value index is

$$P_{QAUV}^{01} = \frac{\left[\sum_{i \in U} \lambda_{i/b} q_i^1 \Big/ \sum_{i \in U} p_i^1 q_i^1 \right]^{-1}}{\left[\sum_{i \in U} \lambda_{i/b} q_i^0 \Big/ \sum_{i \in U} p_i^0 q_i^0 \right]^{-1}} = \frac{\left[\sum_{i \in U} s_i^1 (\tilde{p}_i^1)^{-1} \right]^{-1}}{\left[\sum_{i \in U} s_i^0 (\tilde{p}_i^0)^{-1} \right]^{-1}}, \tag{6}$$

where $s_i^0 = p_i^0 q_i^0 / \sum_{i \in U} p_i^0 q_i^0$ and $s_i^1 = p_i^1 q_i^1 / \sum_{i \in U} p_i^1 q_i^1$ denote the expenditure shares of item i. So the quality-adjusted unit value index equals the ratio of weighted harmonic averages of quality-adjusted prices with expenditure shares serving as weights.

An interesting situation arises when all quality adjustment factors are the same. Since each item then is of the "same quality" as the base item b, we have $\lambda_{i/b} = 1$ for all $i \in U$, hence $\tilde{q}_i^0 = q_i^0$ and $\tilde{q}_i^1 = q_i^1$ in equations (4) and (5) and $\tilde{p}_i^0 = p_i^0$ and $\tilde{p}_i^1 = p_i^1$ in equations (5) and (6). In this case, the quality-adjusted unit value index simplifies to the ordinary unit value index:[3]

$$P_{UV}^{01} = \frac{\sum_{i \in U} p_i^1 q_i^1 \Big/ \sum_{i \in U} q_i^1}{\sum_{i \in U} p_i^0 q_i^0 \Big/ \sum_{i \in U} q_i^0} = \frac{\left[\sum_{i \in U} s_i^1 (p_i^1)^{-1} \right]^{-1}}{\left[\sum_{i \in U} s_i^1 (p_i^0)^{-1} \right]^{-1}}. \tag{7}$$

---

[3] As far as we know, de Haan (2004a) was the first to write the unit value index as a ratio of expenditure share weighted harmonic average prices (and the quality-adjusted unit value index as the quality-adjusted counterpart).

2.2 Estimating the quality adjustment factors

In this section, we discuss ways to estimate the quality adjustment factors. An important condition is that the quantity index (3) and the quality-adjusted unit value index (4) be *invariant to the choice of base item b*. We investigate two estimation methods: a basic economic method and an econometric/hedonic method. It is the latter method we will be building upon in subsequent sections.

The basic method

Economic theory suggests that, under certain conditions, the difference in price between a pair of broadly comparable items reflects the market value of the quality difference. Accordingly, the ratio of the period t prices of item i and base item b is a useful measure of the quality adjustment factor in period t $(t = 0,1)$. Setting $\lambda_{i/b} = p_i^0 / p_b^0$ for all $i \in U$ in (3) yields

$$Q^{01} = \frac{\sum_{i \in U}(p_i^0 / p_b^0)q_i^1}{\sum_{i \in U}(p_i^0 / p_b^0)q_i^0} = \frac{\sum_{i \in U} p_i^0 q_i^1}{\sum_{i \in U} p_i^0 q_i^0} = Q_L^{01}, \tag{8}$$

which is the Laspeyres quantity index. Alternatively, setting $\lambda_{i/b} = p_i^1 / p_b^1$ gives

$$Q^{01} = \frac{\sum_{i \in U}(p_i^1 / p_b^1)q_i^1}{\sum_{i \in U}(p_i^1 / p_b^1)q_i^0} = \frac{\sum_{i \in U} p_i^1 q_i^1}{\sum_{i \in U} p_i^1 q_i^0} = Q_P^{01}, \tag{9}$$

which is the Paasche quantity index. Both indexes are obviously invariant to the choice of base item. It is well known that the corresponding price indexes – or quality-adjusted unit value indexes in our language – that satisfy equation (1) are the Paasche price index and the Laspeyres price index, respectively:[4]

$$P_{QAUV}^{01} = \frac{V^{01}}{Q_L^{01}} = \frac{\sum_{i \in U} p_i^1 q_i^1}{\sum_{i \in U} p_i^0 q_i^1} = P_P^{01}; \tag{10}$$

$$P_{QAUV}^{01} = \frac{V^{01}}{Q_P^{01}} = \frac{\sum_{i \in U} p_i^1 q_i^0}{\sum_{i \in U} p_i^0 q_i^0} = P_L^{01}. \tag{11}$$

---

[4] For the static case, von Auer (2014) showed that many standard price index number formulas, including those of Laspeyres and Paasche, belong to a whole family of generalized unit value indexes.

For broadly comparable items, we expect the price ratios p_i^0 / p_b^0 and p_i^1 / p_b^1 to be more or less the same unless the two periods are far apart. Yet, there will most likely be disturbances so that the Laspeyres and Paasche quantity indexes, hence the Paasche and Laspeyres price indexes, will generally differ. Taking geometric means, which leads to Fisher quantity and price indexes, is a natural solution. However, the Fisher quantity index cannot be written in the form of (3).

The econometric method

This method makes use of a hedonic model, which explains the price of items in terms of a set of product characteristics. More precisely, we use the log-linear "time dummy model". A log-linear model specification usually fits the data better than a strictly linear specification. The characteristics parameters are constrained to be fixed over time. This constraint can be questioned, but it allows us to pool the data of periods 0 and 1 in order to increase degrees of freedom.

The estimating equation for the two perod time dummy hedonic model is

$$\ln p_i^t = \delta^0 + \delta^1 D_i^1 + \sum_{k=1}^{K} \beta_k z_{ik} + \varepsilon_i^t, \tag{12}$$

where z_{ik} is the quantity of the k-th characteristic ($k = 0,...,K$) for item i and β_k^t the corresponding parameter; D_i^1 is a dummy variable that has the value 1 if i is purchased in period 1 and 0 otherwise; δ^0 is the intercept and ε_i^t is an error term with an expected value of zero. Estimating (12) by least squares regression produces coefficients $\hat{\delta}^0$, $\hat{\delta}^1$ and $\hat{\beta}_k$. Thus, the predicted prices in periods 0 and 1 are $\hat{p}_i^0 = \exp(\hat{\delta}^0)\exp[\sum_{k=1}^{K} \hat{\beta}_k z_{ik}]$ and $\hat{p}_i^1 = \exp(\delta^0)\exp(\hat{\delta}^1)\exp[\sum_{k=1}^{K} \hat{\beta}_k z_{ik}]$.[5]

Replacing the quality adjustment factors in equation (3) by the estimated period 0 price ratio $\hat{p}_i^0 / \hat{p}_b^0$ gives

$$Q^{01} = \frac{\sum_{i \in U}(\hat{p}_i^0 / \hat{p}_b^0)q_i^1}{\sum_{i \in U}(\hat{p}_i^0 / \hat{p}_b^0)q_i^0} = \frac{\sum_{i \in U}\hat{p}_i^0 q_i^1}{\sum_{i \in U}\hat{p}_i^0 q_i^0} = \hat{Q}_P^{01}, \tag{13}$$

and replacing them by the estimated period 1 price ratio $\hat{p}_i^1 / \hat{p}_b^1$ gives

---

[5] Taking antilogs is a nonlinear transformation and so the predicted prices are not unbiased. Kennedy (1981) and van Garderen and Shah (2002) suggested adjustments for this type of small sample bias. We assume that the number of observations is large enough to ignore the bias.

$$Q^{01} = \frac{\sum_{i \in U} (\hat{p}_i^1 / \hat{p}_b^1) q_i^1}{\sum_{i \in U} (\hat{p}_i^1 / \hat{p}_b^1) q_i^0} = \frac{\sum_{i \in U} \hat{p}_i^1 q_i^1}{\sum_{i \in U} \hat{p}_i^1 q_i^0} = \hat{Q}_L^{01}.$$ (14)

It seems as if we have obtained estimators of Paasche and Laspeyres quantity indexes. As mentioned earlier, we expect the price ratios p_i^0 / p_b^0 and p_i^1 / p_b^1 to be roughly the same. The time dummy method constrains the estimated ratios to be exactly the same, i.e., $\hat{p}_i^0 / \hat{p}_b^0 = \hat{p}_i^1 / \hat{p}_b^1 (= \exp[\sum_{k=1}^{K} \hat{\beta}_k (z_{ik} - z_{bk})])$. It follows that $\hat{Q}_P^{01}$ and $\hat{Q}_L^{01}$ coincide, which makes it difficult to interpret them as Paasche and Laspeyres indexes. As we will see, the resulting index number formula crucially depends on the weights applied in the regression.

Given that $\hat{p}_i^0 / \hat{p}_b^0 = \hat{p}_i^1 / \hat{p}_b^1$, we can use $\hat{p}_i^1 / \hat{p}_b^1$ as the estimator of the quality adjustment factor in the numerator of equation (3) and $\hat{p}_i^0 / \hat{p}_b^0$ in the denominator to get the following alternative expression for the quantity index:

$$Q^{01} = \frac{\sum_{i \in U} (\hat{p}_i^1 / \hat{p}_b^1) q_i^1}{\sum_{i \in U} (\hat{p}_i^0 / \hat{p}_b^0) q_i^0} = \frac{\sum_{i \in U} \hat{p}_i^1 q_i^1}{\sum_{i \in U} \hat{p}_i^0 q_i^0} \frac{1}{(\hat{p}_b^1 / \hat{p}_b^0)} = \frac{\sum_{i \in U} \hat{p}_i^1 q_i^1}{\sum_{i \in U} \hat{p}_i^0 q_i^0} \frac{1}{P_{TD}^{01}},$$ (15)

where $P_{TD}^{01} = \hat{p}_b^1 / \hat{p}_b^0 = \exp(\hat{\delta}^1)$ is the time dummy price index. Dividing the value index by (15) yields

$$P_{QAUV}^{01} = \left[\frac{\sum_{i \in U} \hat{p}_i^0 q_i^0 \sum_{i \in U} p_i^1 q_i^1}{\sum_{i \in U} p_i^0 q_i^0 \sum_{i \in U} \hat{p}_i^1 q_i^1} \right] P_{TD}^{01} = \left[\frac{\sum_{i \in U} s_i^0 \exp(u_i^0)}{\sum_{i \in U} s_i^1 \exp(u_i^1)} \right] P_{TD}^{01},$$ (16)

where s_i^0 and s_i^1 are the expenditure shares in periods 0 and 1, and $u_i^0 = \ln(\hat{p}_i^0 / p_i^0)$ and $u_i^1 = \ln(\hat{p}_i^1 / p_i^1)$ are the regression residuals in the two periods.

From an econometric point of view, Ordinary Least Squares (OLS) regression will suffice if the variance of the errors is constant. In the static/matched item case, this would produce the unweighted geometric average of price ratios, i.e. the Jevons price index. But from an index number point of view, items should be weighted according to their economic importance to obtain a weighted index. A useful measure of economic importance in this context is expenditure shares. Diewert (2004) showed that the use of average expenditure shares in the two periods, $(s_i^0 + s_i^1)/2$, as weights in a Weighted Least Squares (WLS) regression makes the time dummy index equal to the superlative Törnqvist price index $P_T^{01} = \prod_{i \in U} (p_i^1 / p_i^0)^{(s_i^0 + s_i^1)/2}$. We then find

$$P_{QAUV}^{01} = \left[\frac{\sum\limits_{i \in U} s_i^0 \exp(u_i^0)}{\sum\limits_{i \in U} s_i^1 \exp(u_i^1)} \right] P_T^{01} .$$

(17)

The bracketed factor in equation (17) is a ratio of expenditure share weighted averages of exponentiated residuals. We expect this factor to be close to 1, hence $P_{QAUV}^{01} \cong P_T^{01}$.[6]

3. The two period dynamic case

In section 2, the set of items was kept fixed with the aim of explaining our ideas and pointing towards similarities and differences with standard index number theory. The present section deals with the more interesting dynamic situation where the set of items changes across time as a result of new items appearing on the market and "old" items disappearing. We are still comparing two periods. In section 4 below, we address the realistic situation with many periods and a constantly changing item universe. This will affect the methods we can employ.

3.1 The dynamic universe

The sets of available items in periods 0 and 1 are denoted by U^0 and U^1. It is important to realize that for making price and quantity comparisons between these two periods, we should not look at U^0 and U^1 separately but rather at the *union* $U^{01} = U^0 \cup U^1$. This makes it possible to derive imputation price and quantity indexes. A subset of items is usually purchased in both time periods. This matched set or intersection is denoted by $U_M = U^0 \cap U^1$. The set of disappearing items, i.e. items purchased in period 0 but not in period 1, is denoted by U_D^0, while the set of new items, i.e. items purchased in period 1 but not in period 0, is denoted by U_N^1. Note that $U_D^0 \cup U_M = U^0$, $U_N^1 \cup U_M = U^1$, and $U^{01} = U^0 \cup U^1 = U_M \cup U_D^0 \cup U_N^1$. Prices are again strictly positive, but quantities are now non-negative: in the two period dynamic case, quantities are either positive or zero in one of the periods (or zero in both periods, but that is irrelevant as those items do not belong to the union U^{01}).

---

[6] Needless to say that conventional index number theory would favor the WLS time dummy index over the quality-adjusted unit value index: the Törnqvist index belongs to the class of superlative indexes, and so multiplying the time dummy index by the bracketed factor would only "bias" the result. We will return to this issue in section 4.2 on the choice of regression weights in the many periods case.

Defining the aggregate value ratio on U^{01} is straightforward:

$$V^{01} = \frac{V^1}{V^0} = \frac{\sum_{i \in U^{01}} p_i^1 q_i^1}{\sum_{i \in U^{01}} p_i^0 q_i^0} = \frac{\sum_{i \in U^1} p_i^1 q_i^1}{\sum_{i \in U^0} p_i^0 q_i^0}. \tag{18}$$

Again, our aim is to decompose the value ratio into a quantity index and a price index, or in our case a quantity index and a quality-adjusted unit value index. We start with a generic quantity index similar to (3), but now defined on the set U^{01}:

$$Q^{01} = \frac{\sum_{i \in U^{01}} \lambda_{i/b} q_i^1}{\sum_{i \in U^{01}} \lambda_{i/b} q_i^0} = \frac{\sum_{i \in U_M} \lambda_{i/b} q_i^1 + \sum_{i \in U_D^0} \lambda_{i/b} q_i^1 + \sum_{i \in U_N^1} \lambda_{i/b} q_i^1}{\sum_{i \in U_M} \lambda_{i/b} q_i^0 + \sum_{i \in U_D^0} \lambda_{i/b} q_i^0 + \sum_{i \in U_N^1} \lambda_{i/b} q_i^0}. \tag{19}$$

Since $q_i^1 = 0$ for $i \in U_D^0$ and $q_i^0 = 0$ for $i \in U_N^1$, equation (19) simplifies to

$$Q^{01} = \frac{\sum_{i \in U_M} \lambda_{i/b} q_i^1 + \sum_{i \in U_N^1} \lambda_{i/b} q_i^1}{\sum_{i \in U_M} \lambda_{i/b} q_i^0 + \sum_{i \in U_D^0} \lambda_{i/b} q_i^0} = \frac{\sum_{i \in U^1} \lambda_{i/b} q_i^1}{\sum_{i \in U^0} \lambda_{i/b} q_i^0}, \tag{20}$$

which is the ratio of the sum of quality-adjusted quantities. Dividing the value index by the quantity index (20) yields the dynamic counterpart of the quality-adjusted unit value index given by equations (4) and (6):

$$P_{QAUV}^{01} = \frac{\sum_{i \in U^1} p_i^1 q_i^1 \Big/ \sum_{i \in U^1} \lambda_{i/b} q_i^1}{\sum_{i \in U^0} p_i^0 q_i^0 \Big/ \sum_{i \in U^0} \lambda_{i/b} q_i^0} = \left[\frac{\sum_{i \in U^1} s_i^1 (\tilde{p}_i^1)^{-1}}{\sum_{i \in U^0} s_i^0 (\tilde{p}_i^0)^{-1}} \right]^{-1}, \tag{21}$$

where $\tilde{p}_i^0 = p_i^0 / \lambda_{i/b}$ and $\tilde{p}_i^1 = p_i^1 / \lambda_{i/b}$ are quality-adjusted prices, as before.

Notice that, just like the value index (18), the quantity index (20) is effectively based on two *different* sets of items, the period 0 set (in the denominator) and the period 1 set (in the numerator). This may seem unusual, but the resulting quality-adjusted unit value index (21) turns out to be consistent with standard imputation price indexes.

3.2 Estimating the quality adjustment factors

We now return to the two methods to estimate the quality adjustment factors, the basic method and the econometric method, and discuss the implications for the quantity index (20) and the quality-adjusted unit value index (21).

The basic method

This method would use $\lambda_{i/b} = p_i^0 / p_b^0$ in equation (20) for the quantity index. However, base period prices for $i \in U_N^1$ cannot be directly observed; they are "missing" and have to be imputed by $\hat{p}_i^0$. Assuming that the base item b belongs to the matched set U_M, we find

$$Q^{01} = \frac{\sum_{i \in U_M}(p_i^0 / p_b^0)q_i^1 + \sum_{i \in U_N^1}(\hat{p}_i^0 / p_b^0)q_i^1}{\sum_{i \in U_M}(p_i^0 / p_b^0)q_i^0 + \sum_{i \in U_D^0}(p_i^0 / p_b^0)q_i^0} = \frac{\sum_{i \in U_M}p_i^0 q_i^1 + \sum_{i \in U_N^1}\hat{p}_i^0 q_i^1}{\sum_{i \in U_M}p_i^0 q_i^0 + \sum_{i \in U_D^0}p_i^0 q_i^0} = Q_{L,SI}^{01} . \tag{22}$$

We refer to (22) as the *single imputation* Laspeyres quantity index. Note that prices are imputed, not quantities – imputing quantities makes no sense. We require some kind of imputation method. If hedonic imputation is used, the basic method becomes partially econometric.

Alternatively, we can use $\lambda_{i/b} = p_i^1 / p_b^1$ in (20). Since period 1 prices for $i \in U_D^0$ are unobservable, they must be imputed by $\hat{p}_i^1$, yielding the single imputation Paasche quantity index

$$Q^{01} = \frac{\sum_{i \in U_M}(p_i^1 / p_b^1)q_i^1 + \sum_{i \in U_N^1}(p_i^1 / p_b^1)q_i^1}{\sum_{i \in U_M}(p_i^1 / p_b^1)q_i^0 + \sum_{i \in U_D^0}(\hat{p}_i^1 / p_b^1)q_i^0} = \frac{\sum_{i \in U_M}p_i^1 q_i^1 + \sum_{i \in U_N^1}p_i^1 q_i^1}{\sum_{i \in U_M}p_i^1 q_i^0 + \sum_{i \in U_D^0}\hat{p}_i^1 q_i^0} = Q_{P,SI}^{01} . \tag{23}$$

The corresponding quality-adjusted unit value indexes are obtained by dividing the value index by the above single imputation quantity indexes. This gives

$$P_{QAUV}^{01} = \frac{\sum_{i \in U_M}p_i^1 q_i^1 + \sum_{i \in U_N^1}p_i^1 q_i^1}{\sum_{i \in U_M}p_i^0 q_i^1 + \sum_{i \in U_N^1}\hat{p}_i^0 q_i^1} = P_{P,SI}^{01} , \tag{24}$$

which is known as the single imputation Paasche price index, and

$$P_{QAUV}^{01} = \frac{\sum_{i \in U_M}p_i^1 q_i^0 + \sum_{i \in U_D^0}\hat{p}_i^1 q_i^0}{\sum_{i \in U_M}p_i^0 q_i^0 + \sum_{i \in U_D^0}p_i^0 q_i^0} = P_{L,SI}^{01} , \tag{25}$$

which is the single imputation Laspeyres price index. As expected, the imputation price indexes are based on a *single set* of items: U^1 in case of the Paasche index and U^0 in case of the Laspeyres index.

By taking the geometric mean of expressions (22) and (23) and expressions (24) and (25), respectively, single imputation Fisher quantity and price indexes are obtained:

$$
O_{F,SI}^{01} = \left[\frac{\sum_{i \in U_M} p_i^0 q_i^1 + \sum_{i \in U_N^1} \hat{p}_i^0 q_i^1}{\sum_{i \in U_M} p_i^0 q_i^0 + \sum_{i \in U_D^0} p_i^0 q_i^0} \frac{\sum_{i \in U_M} p_i^1 q_i^1 + \sum_{i \in U_N^1} p_i^1 q_i^1}{\sum_{i \in U_M} p_i^1 q_i^0 + \sum_{i \in U_D^0} \hat{p}_i^1 q_i^0} \right]^{1/2} ; \tag{26}
$$

$$
P_{F,SI}^{01} = \left[\frac{\sum_{i \in U_M} p_i^1 q_i^0 + \sum_{i \in U_D^0} \hat{p}_i^1 q_i^0}{\sum_{i \in U_M} p_i^0 q_i^0 + \sum_{i \in U_D^0} p_i^0 q_i^0} \frac{\sum_{i \in U_M} p_i^1 q_i^1 + \sum_{i \in U_N^1} p_i^1 q_i^1}{\sum_{i \in U_M} p_i^0 q_i^1 + \sum_{i \in U_N^1} \hat{p}_i^0 q_i^1} \right]^{1/2} . \tag{27}
$$

While being a measure of quality-adjusted price change, the single imputation Fisher price index (27) cannot be interpreted as a quality-adjusted unit value index and is not fit for our purpose.[7]

The econometric method

With a few minor changes, the econometric method for the static case discussed earlier carries over to the dynamic case. The time dummy model (12) remains our estimating hedonic equation but a slightly modified WLS regression will be needed and explained below. The predicted period 0 and period 1 prices are $\hat{p}_i^0 = \exp(\hat{\delta}^0) \exp[\sum_{k=1}^K \hat{\beta}_k z_{ik}]$ and $\hat{p}_i^1 = \exp(\hat{\delta}^0) \exp(\hat{\delta}^1) \exp[\sum_{k=1}^K \hat{\beta}_k z_{ik}]$, as before, and the time dummy index is given by $P_{TD}^{01} = \hat{p}_i^1 / \hat{p}_i^0 = \exp(\hat{\delta}^1)$.

Note that $\hat{p}_i^0 / \hat{p}_b^0 = \hat{p}_i^1 / \hat{p}_b^1 (= \exp[\sum_{k=1}^K \hat{\beta}_k (z_{ik} - z_{bk})])$. Using $\hat{p}_i^1 / \hat{p}_b^1$ and $\hat{p}_i^0 / \hat{p}_b^0$ as estimates of the quality adjustment factors in the numerator and denominator of (20), respectively, the following expression for the quantity index is obtained:

$$
Q^{01} = \frac{\sum_{i \in U^1} (\hat{p}_i^1 / \hat{p}_b^1) q_i^1}{\sum_{i \in U^0} (\hat{p}_i^0 / \hat{p}_b^0) q_i^0} = \frac{\sum_{i \in U^1} \hat{p}_i^1 q_i^1}{\sum_{i \in U^0} \hat{p}_i^0 q_i^0} \frac{1}{(\hat{p}_b^1 / \hat{p}_b^0)} = \frac{\sum_{i \in U^1} \hat{p}_i^1 q_i^1}{\sum_{i \in U^0} \hat{p}_i^0 q_i^0} \frac{1}{P_{TD}^{01}} . \tag{28}
$$

Dividing the value index by (28) gives rise to the following quality-adjusted unit value index:

$$
P_{QAUV}^{01} = \left[\frac{\sum_{i \in U^0} \hat{p}_i^0 q_i^0}{\sum_{i \in U^0} p_i^0 q_i^0} \frac{\sum_{i \in U^1} p_i^1 q_i^1}{\sum_{i \in U^1} \hat{p}_i^1 q_i^1} \right] P_{TD}^{01} = \left[\frac{\sum_{i \in U^0} s_i^0 \exp(u_i^0)}{\sum_{i \in U^1} s_i^1 \exp(u_i^1)} \right] P_{TD}^{01}, \tag{29}
$$

----

[7] De Haan (2002) referred to (27) as a generalized Fisher price index.

where $u_i^0 = \ln(\hat{p}_i^0 / p_i^0)$ and $u_i^1 = \ln(\hat{p}_i^1 / p_i^1)$ are the regression residuals.

For the choice of weights in the WLS regression, we apply a result derived by de Haan (2004b) saying that if regression weights $(s_i^0 + s_i^1)/2$ for $i \in U_M$, $s_i^0 / 2$ for $i \in U_D^0$ and $s_i^1 / 2$ for $i \in U_N^1$ are used, the time dummy index equals[8]

$$P_{TD}^{01} = \prod_{i \in U_M}\left(\frac{p_i^1}{p_i^0}\right)^{\frac{(s_i^0 + s_i^1)}{2}} \prod_{i \in U_D^0}\left(\frac{\hat{p}_i^1}{p_i^0}\right)^{\frac{s_i^0}{2}} \prod_{i \in U_N^1}\left(\frac{p_i^1}{\hat{p}_i^0}\right)^{\frac{s_i^1}{2}} = P_{T,SI}^{01}. \tag{30}$$

Expression (30) is a single imputation Törnqvist price index where the "missing prices" are imputed by the predicted values from the WLS time dummy regression. The quality-adjusted unit value index (29) can therefore be written as

$$P_{QAUV}^{01} = \left[\frac{\sum_{i \in U^0} s_i^0 \exp(u_i^0)}{\sum_{i \in U^1} s_i^1 \exp(u_i^1)}\right] P_{T,SI}^{01}. \tag{31}$$

Expression (31) resembles expression (16) for the quality-adjusted unit value index in the static case; if all items are matched ($U^0 = U^1 = U$) and no imputations for "missing prices" are needed, (31) reduces to (16).

4. The many periods case

In this section, we extend our approach to the realistic case with multiple time periods. There are now three different options: estimating direct (bilateral) indexes, calculating chained period-on-period indexes, or estimating multilateral indexes. Below, we explain our preference for multilateral indexes and apply the econometric method to estimate the quality adjustment factors. We also discuss the choice of weights in the pooled time dummy regression.

4.1 Transitivity and quality-adjusted unit value indexes

The estimation of direct (bilateral) indexes in the many periods case is a straightforward extension of the two period case described in section 3. Suppose we want to estimate

---

[8] The derivation can also be found in the Appendix to de Haan and Krsinich (2014a). Notice that the regression weights are equal to the weights of the items in the single imputation Törnqvist price index (30).

price indexes for periods $t = 1,...,T$. We can just compare each period t directly with the base period and estimate quality-adjusted unit value indexes according to equation (29) using period 0 and period t (rather than period 1) data. This method has two drawbacks. First, scanner data typically show a high rate of churn in terms of new and disappearing items;[9] the matched part of the imputation Törnqvist index (30) diminishes rapidly so that we would be increasingly relying on model based imputations. It would be better to implement methods that make maximum use of all the matches in the data. Second, the assumption of fixed characteristics parameters may be justifiable in the short run but becomes debatable when the sample period grows. Thus, adhering to direct indexes is problematic and we would have to find ways of dealing with this problem; in practice some kind of chaining will be necessary.

Period-on-period chaining might seem a promising approach because it makes use of the matched data for all pairs of adjacent periods. Indeed, the CPI Manual (ILO et al., 2004) advocates the use of chained superlative price indexes. However, empirical studies during the last decade provided evidence of significant chain drift in period-on-period chained superlative indexes.[10] The lesson is that high frequency chaining should not be used.

Multilateral index number approaches to price comparisons across time relate to more than two periods and generate transitive indexes; the price changes between any two time periods are independent of the choice of base period. Transitivity implies that the index can be written in chained form and by construction does not suffer from chain drift. When applied to pooled data of three or more periods, the time dummy hedonic method is a multilateral approach that yields transitive quality-adjusted price indexes. In our opinion, this method is most appropriate when dealing with scanner data on a large scale.

The relevant set in the many periods case exists of all items purchased in one or more periods during the sample period. Yet, for a quantity index that compares period t

---

[9] A high rate of churn can be partly the result of the way in which we define homogenous items. We will address the homogeneity issue in section 6.

[10] Downward drift in chained superlative price indexes for goods sold in supermarkets is documented in e.g., Ivancic (2007), Ivancic, Diewert and Fox (2011), and de Haan and van der Grient (2011). The drift is mainly due to quantities spiking when storable goods are on sale. De Haan and Krsinich (2014a) found drift in chained superlative price indexes for consumer electronic products due to seasonality in the prices and quantities sold.

to period 0, we are only interested in those items purchased in period 0, period t, or in both periods, i.e. in the union $U^{0t} = U^0 \cup U^t$. Similar to what we did in section 3.1, we define a matched set $U_M^{0t} = U^0 \cap U^t$ of items purchased in both periods, a set of items $U_D^{0(t)}$ of items purchased in period 0 (and perhaps in some other periods as well) but not in period t, and a set $U_N^{t(0)}$ of items purchased in period t (and perhaps in other periods also) but not in period 0. Note that $U_D^{0(t)} \cup U_M^{0t} = U^0$ and $U_N^{t(0)} \cup U_M^{0t} = U^t$. The quantity index is

$$Q^{0t} = \frac{\sum\limits_{i \in U_M^{0t}} \lambda_{i/b} q_i^t + \sum\limits_{i \in U_D^{0(t)}} \lambda_{i/b} q_i^t + \sum\limits_{i \in U_N^{t(0)}} \lambda_{i/b} q_i^t}{\sum\limits_{i \in U_M^{0t}} \lambda_{i/b} q_i^0 + \sum\limits_{i \in U_D^{0(t)}} \lambda_{i/b} q_i^0 + \sum\limits_{i \in U_N^{t(0)}} \lambda_{i/b} q_i^0} = \frac{\sum\limits_{i \in U^t} \lambda_{i/b} q_i^t}{\sum\limits_{i \in U^0} \lambda_{i/b} q_i^0}. \tag{32}$$

The second expression of (32) holds true because $q_i^t = 0$ for $i \in U_D^{0(t)}$ and $q_i^0 = 0$ for $i \in U_N^{t(0)}$.

Quantity index (32) is transitive.[11] The value index is also transitive, and so the quality-adjusted unit value index

$$P_{QAUV}^{0t} = \frac{V^{0t}}{Q^{0t}} = \frac{\sum\limits_{i \in U^t} p_i^t q_i^t \Big/ \sum\limits_{i \in U^1} \lambda_{i/b} q_i^t}{\sum\limits_{i \in U^0} p_i^0 q_i^0 \Big/ \sum\limits_{i \in U^0} \lambda_{i/b} q_i^0} = \left[\frac{\sum\limits_{i \in U^t} s_i^t (\tilde{p}_i^t)^{-1}}{\sum\limits_{i \in U^0} s_i^0 (\tilde{p}_i^0)^{-1}} \right]^{-1}, \tag{33}$$

with quality-adjusted prices $\tilde{p}_i^0 = p_i^0 / \lambda_{i/b}$ and $\tilde{p}_i^t = p_i^t / \lambda_{i/b}$, is transitive. In practice of course we have to estimate the quality adjustment factors $\lambda_{i/b}$. The pooled time dummy method preserves transitivity by producing estimates of $\lambda_{i/b}$ which are fixed across the whole sample period $t = 0, ..., T$.

The estimating equation for the multilateral time dummy model is

$$\ln p_i^t = \delta^0 + \sum_{t=1}^{T} \delta^t D_i^t + \sum_{k=1}^{K} \beta_k z_{ik} + \varepsilon_i^t, \tag{34}$$

where D_i^t is a dummy variable that has the value 1 if item i is purchased in period t and 0 otherwise. The predicted period 0 and period t prices from a least squares regression of (34) are $\hat{p}_i^0 = \exp(\hat{\delta}^0) \exp[\sum_{k=1}^{K} \hat{\beta}_k z_{ik}]$ and $\hat{p}_i^t = \exp(\delta^0) \exp(\hat{\delta}^t) \exp[\sum_{k=1}^{K} \hat{\beta}_k z_{ik}]$; the time dummy index is given by $P_{TD}^{0t} = \hat{p}_i^t / \hat{p}_i^0 = \exp(\hat{\delta}^t)$. The multilateral time dummy index is transitive because the regression results are independent of the choice of base period (which in model (34) is the starting period 0).

---

[11] The index would even be transitive for time-dependent quality adjustment factors.

We have $\hat{p}_i^0 / \hat{p}_b^0 = \hat{p}_i^t / \hat{p}_b^t (= \exp[\sum_{k=1}^{K} \hat{\beta}_k (z_{ik} - z_{bk})])$. Using $\hat{p}_i^t / \hat{p}_b^t$ and $\hat{p}_i^0 / \hat{p}_b^0$ as estimates of the quality adjustment factors in the numerator and denominator of (32), respectively, we obtain

$$Q^{0t} = \frac{\sum_{i \in U^t} (\hat{p}_i^t / \hat{p}_b^t) q_i^t}{\sum_{i \in U^0} (\hat{p}_i^0 / \hat{p}_b^0) q_i^0} = \frac{\sum_{i \in U^t} \hat{p}_i^t q_i^t}{\sum_{i \in U^0} \hat{p}_i^0 q_i^0} \frac{1}{(\hat{p}_b^t / \hat{p}_b^0)} = \frac{\sum_{i \in U^t} \hat{p}_i^t q_i^t}{\sum_{i \in U^0} \hat{p}_i^0 q_i^0} \frac{1}{P_{TD}^{0t}} . \tag{35}$$

The multilateral quality-adjusted unit value index becomes

$$P_{QAUV}^{0t} = \frac{V^{0t}}{Q^{0t}} = \left[\frac{\sum_{i \in U^0} \hat{p}_i^0 q_i^0}{\sum_{i \in U^0} p_i^0 q_i^0} \frac{\sum_{i \in U^t} p_i^t q_i^t}{\sum_{i \in U^t} \hat{p}_i^t q_i^t} \right] P_{TD}^{0t} = \left[\frac{\sum_{i \in U^0} s_i^0 \exp(u_i^0)}{\sum_{i \in U^t} s_i^t \exp(u_i^t)} \right] P_{TD}^{0t} , \tag{36}$$

where $u_i^0 = \ln(\hat{p}_i^0 / p_i^0)$ and $u_i^t = \ln(\hat{p}_i^t / p_i^t)$ are the regression residuals in periods 0 and t $(t = 1,...,T)$. The quality-adjusted unit value index (36) and the quantity index (35) are independent of the choice of base item b, as required. In the extreme situation when the quantities of all the characteristics z_{ik} happen to be the same for all items in the pooled data set, R^2 and the factor between square brackets in (36) equal 1, so that the quality-adjusted unit value index simplifies to the ordinary unit value index. This is one of the strong points of the time dummy hedonic approach to estimating the quality adjustment factors.

4.2 The choice of regression weights

An important question is: what regression weights should be used when estimating the multilateral time dummy hedonic model (34)? The results found by Diewert (2004) and de Haan (2004) for the two period static and dynamic cases do not simply carry over to the many periods case. For instance, it will not be possible to obtain a multilateral time dummy index that exactly equals the matched item Törnqvist index – we know that the time dummy index is transitive but the Törnqvist is not. Still we want to weight items according to their economic importance. Diewert (2004) suggested using expenditure shares pertaining to the periods the items are actually observed, i.e. s_i^0 for $i \in U^0$ and s_i^t for $i \in U^t$ $(t = 1,...,T)$. These weights have been used in a number of studies, e.g., by Ivancic and Fox (2013) and de Haan and Krsinich (2014a) (2014b).

De Haan and Krsinich (2014b) showed that the weighted time dummy index can be written as

$$P_{TD}^{0t} = \frac{\prod_{i \in U^t} (\hat{\tilde{p}}_i^t)^{s_i^t}}{\prod_{i \in U^0} (\hat{\tilde{p}}_i^0)^{s_i^0}}, \qquad (37)$$

where $\hat{\tilde{p}}_i^0 = p_i^0 / \hat{\lambda}_{i/b}$ and $\hat{\tilde{p}}_i^t = p_i^t / \hat{\lambda}_{i/b}$ are the estimated quality-adjusted prices, with $\hat{\lambda}_{i/b} = \hat{p}_i^0 / \hat{p}_b^0 = \hat{p}_i^t / \hat{p}_b^t$. According to (37), the (WLS) time dummy index is equal to the ratio of expenditure share weighted *geometric* means of the estimated quality-adjusted prices. Applying the same $\hat{\lambda}_{i/b}$ as estimates of the $\lambda_{i/b}$ in (33), the quality-adjusted unit value index is equal to the ratio of expenditure share weighted *harmonic* means of the estimated quality-adjusted prices:

$$P_{QAUV}^{0t} = \left[\frac{\sum_{i \in U^t} s_i^t (\hat{\tilde{p}}_i^t)^{-1}}{\sum_{i \in U^0} s_i^0 (\hat{\tilde{p}}_i^0)^{-1}} \right]^{-1}. \qquad (38)$$

The bracketed factor in expression (36) measures the gap between the quality-adjusted unit value index and the time dummy index. De Haan and Krsinich (2014b) derived the following result:

$$\frac{\sum_{i \in U^0} s_i^0 \exp(u_i^0)}{\sum_{i \in U^t} s_i^t \exp(u_i^t)} \cong \frac{1 + \frac{1}{2}(\sigma^0)^2}{1 + \frac{1}{2}(\sigma^t)^2}, \qquad (39)$$

where $(\sigma^0)^2 = \sum_{i \in U^0} s_i^0 (u_i^0)^2$ and $(\sigma^t)^2 = \sum_{i \in U^t} s_i^t (u_i^t)^2$ denote the weighted variances of the residuals from the WLS regression in periods 0 and *t*. So the (weighted) variance of the regression residuals or, equivalently, the dispersion of the quality-adjusted prices, is the main driver of the difference between the two indexes.[12]

Expressions (36) and (39) indicate that the quality-adjusted unit value index will sit below (above) the time dummy index when the variance of the residuals increases (decreases) over time. Due to the logarithmic functional form for the hedonic model, this type of heteroskedasticity is unlikely to occur. In a linear hedonic model with price rather than log of price as the dependent variable, the absolute errors tend to grow over time when there is inflation. The logarithmic transformation neutralizes this tendency, as pointed out by Diewert (2004). So we would expect the two indexes to have similar

---

[12] For a discussion on the difference between unweighted price indexes at the elementary level in terms of price dispersion and product heterogeneity, see Silver and Heravi (2007b).

trends and volatility. This was confirmed by empirical work by de Haan and Krsinich (2014b) on New Zealand scanner data for a number of consumer electronics products; the differences between the two types of index were negligible.

Even though the time series are likely to be very similar, we favor the quality-adjusted unit value index over the time dummy index because the former reduces to the ordinary unit value index if all items had the same quantities of characteristics while the time dummy hedonic index produces the geometric counterpart of the unit value index. More generally, we could think of (39) as a factor that changes the "geometric quality-adjusted unit value index" into the desired arithmetic version.

Estimating the (arithmetic) quality-adjusted unit value index from scanner data is easy: calculate the expenditure shares for each item in each time period, run a WLS regression of the time dummy model on the pooled data of periods $t = 0,...,T$ and save the residuals, then calculate the time dummy index $P_{TD}^{0t} = \exp(\hat{\delta}^t)$ and factor (39) for each period t, and finally multiply P_{TD}^{0t} by that factor.

There are two further issues. The first one is how to deal with revisions that arise when the sample period is extended and new data is added. This issue will be discussed in section 6. The second issue is how to proceed when information on characteristics is not available. This is the topic discussed in section 5 below.

5. Fixed effects: the time-product dummy method

Aizcorbe, Corrado and Doms (2003) claimed that quality-adjusted price indexes can be constructed without observing item characteristics. They used a regression model which only includes dummy variables for the items plus dummy variables for time periods. De Haan and Krsinich (2014b) named it the *Time-Product Dummy* (TPD) method because it adapts the Country-Product Dummy (CPD) method in order to measure price change across time rather than space.[13] Silver and Heravi (2005) argued that in the many period situation, the TPD index "will have a tendency to follow the chained matched model results." But that cannot be true in general because high frequency chaining of weighted indexes can lead to significant drift while the TPD method yields transitive, hence drift free indexes.

---

[13] The CPD method dates back to Summers (1973). Diewert (1999) and Balk (2001) reviewed the various approaches to international price comparisons.

In the time dummy hedonic model (34), the characteristics of each item and the corresponding parameters are assumed fixed over time. This implies that their combined effect on the log of price is also constant across time. Therefore, if information on item characteristics is not available, it seems natural to replace the unobservable "constant" hedonic effects $\sum_{k=1}^{K} \beta_k z_{ik}$ by the item specific fixed values γ_i. This is what the TPD method does. If N different items are observed across the entire sample period $0,...,T$, most of which will typically not be purchased in every time period, then the estimating equation for the TPD or *fixed effects* model is

$$\ln p_i^t = \delta^0 + \sum_{t=1}^{T} \delta^t D_i^t + \sum_{i=1}^{N-1} \gamma_i D_i + \varepsilon_i^t, \tag{40}$$

where D_i is a dummy variable that has the value of 1 if the observation relates to item i and 0 otherwise. A dummy for an arbitrary item N is not included ($\gamma_N = 0$) to identify the model. The least squares estimates are $\hat{\delta}^0$, $\hat{\delta}^t$ ($t = 1,...,T$) and $\hat{\gamma}_i$ ($i = 1,...,N-1$), and we set $\hat{\gamma}_N = 0$. Note that while items with identical characteristics are assumed to have identical fixed effects γ_i, the estimates $\hat{\gamma}_i$ will generally not be exactly the same. The predicted prices are $\hat{p}_i^0 = \exp(\hat{\delta}^0)\exp(\hat{\gamma}_i)$ and $\hat{p}_i^t = \exp(\hat{\delta}^0)\exp(\hat{\delta}^t)\exp(\hat{\gamma}_i)$ for all i. Similar to the time dummy hedonic index, the TPD index for period t is calculated as $P_{TPD}^{0t} = \hat{p}_i^t / \hat{p}_i^0 = \exp(\hat{\delta}^t)$.

The TPD method is a non-hedonic variant of the time dummy method. In many scanner data sets, only limited information on characteristics is available. It would have been nice if we could combine the hedonic and non-hedonic methods by including both the available item characteristics and item dummy variables in the time dummy model, but that is not possible because the model would then no longer be identified: the vector of values for any characteristic can be written as a linear combination of the N-1 vectors for the product dummies and the intercept.

The TPD method has been used in a number of studies. Aizcorbe, Corrado and Doms (2003) estimated TPD indexes for computers using OLS. More than two decades earlier, Balk (1980) proposed a WLS version for constructing price indexes for seasonal products. Ivancic, Fox and Diewert (2009) adopted an expenditure share weighted TPD approach to estimating price indexes for products sold in Australian supermarkets. Their WLS TPD approach was applied by Krsinich (2011) (2013) (2014) and de Haan and Krsinich (2014a) to consumer electronics products sold in New Zealand. As explained in section 4, we favor the use of expenditure shares as weights in pooled time dummy

regressions. Since the TPD model is an instance of the time dummy model, our choice of weights remains the same.

From a statistical perspective, the TPD method is less efficient than the hedonic time dummy method because much more parameters have to be estimated. On the other hand, the TPD method is cost efficient in that there is no need to collect information on item characteristics. But a more important issue is whether this method really produces price indexes which are adjusted for quality changes, as claimed by Aizcorbe, Corrado and Doms (2003) and recently also by Krsinich (2013) (2014). We will try to shed some light on this issue.

One should recognize that (average) quality change has two components: quality mix change due to changes over time in the quantities purchased of existing or matched items, and quality change due to new and disappearing items. When superlative indexes can be constructed, as they can with scanner data, the first component does not pose any problems. Superlative price indexes treat the base period and the comparison period in a symmetric fashion, ensuring that relative quantity changes, hence quality mix changes, are handled appropriately. Thus, the main issue is how well the TPD method accounts for the effect of new and disappearing items.

The TPD method is essentially a matched item (or panel) method: an item must be observed at least two times during the sample period to be non-trivially included. In the two period case, the resulting matched item index differs from the index found by estimating the model without new and disappearing items only in that the matched items are weighted slightly differently (unless the weights have been normalized; see de Haan and Krsinich, 2014a). In the context of two countries, Diewert (2004) mentioned that the method "can deal with situations where say item n* has transactions in one country but not the other" and that "the prices of item n* will be zeroed out".[14]

The fact that items with a single observation are zeroed out, carries over to the many period case. Thus, in contrast to the multilateral time dummy hedonic index, the multilateral TPD index does not account for the effects of *all* unmatched (i.e., new and disappearing) items. Nevertheless, the resulting index will usually differ from a chained matched model index. This is because items which are new or disappearing in adjacent period comparisons are often observed multiple times during the whole sample period,

---

[14] Although the items will be zeroed out, their fixed effects can still be estimated.

and are not zeroed out. These items contain longitudinal information on price changes that is used in a multilateral time-product dummy model whereas they are ignored in a chained matched model index.

We will now formally show what drives the difference between the expenditure share weighted multilateral TPD index and the chained matched model Törnqvist index. The TPD index, being a special case of the time dummy hedonic index, can be written in the same form as equation (37):

$$
P_{TPD}^{0t} = \frac{\prod_{i \in U^t} (\hat{\tilde{p}}_i^t)^{s_i^t}}{\prod_{i \in U^0} (\hat{\tilde{p}}_i^0)^{s_i^0}}, \tag{41}
$$

where $\hat{\tilde{p}}_i^0 = p_i^0 / (\hat{p}_i^0 / \hat{p}_b^0) = p_i^0 / \exp(\hat{\gamma}_i - \hat{\gamma}_b)$ and $\hat{\tilde{p}}_i^t = p_i^t / (\hat{p}_i^t / \hat{p}_b^t) = p_i^t / \exp(\hat{\gamma}_i - \hat{\gamma}_b)$; b is again an arbitrary base item. If the estimated fixed effects $\hat{\gamma}_i$ and $\hat{\gamma}_b$ approximate the hedonic price effects $\sum_{k=1}^K \hat{\beta}_k z_{ik}$ and $\sum_{k=1}^K \hat{\beta}_k z_{bk}$ reasonably well, then $\hat{\tilde{p}}_i^0$ and $\hat{\tilde{p}}_i^t$ are quality-adjusted prices. Without loss of generality, we can set N in model (40) equal to b, yielding $\hat{\gamma}_b = 0$, $\hat{\tilde{p}}_i^0 = p_i^0 / \exp(\hat{\gamma}_i)$ and $\hat{\tilde{p}}_i^t = p_i^t / \exp(\hat{\gamma}_i)$. Due to transitivity, the TPD index can be written as a period-on-period chained index:

$$
P_{TPD}^{0t} = \prod_{\tau=1}^t \left[\frac{\prod_{i \in U^\tau} (\hat{\tilde{p}}_i^\tau)^{s_i^\tau}}{\prod_{i \in U^{\tau-1}} (\hat{\tilde{p}}_i^{\tau-1})^{s_i^{\tau-1}}} \right]. \tag{42}
$$

Let us focus on a single chain link, i.e. the bracketed factor in (42). De Haan and Hendriks (2013) showed that each chain link can be decomposed into four components, as follows

$$
\frac{P_{TPD}^{0t}}{P_{TPD}^{0,t-1}} = \prod_{i \in U_M^{t-1,t}} \left(\frac{p_i^t}{p_i^{t-1}} \right)^{\frac{s_{iM}^{t-1} + s_{iM}^t}{2}} \left[\frac{\prod_{i \in U_N^{t-1,t}} (\hat{\tilde{p}}_i^t)^{s_{iN}^t}}{\prod_{i \in U_M^{t-1,t}} (\hat{\tilde{p}}_i^t)^{s_{iM}^t}} \right]^{s_N^t} \left[\frac{\prod_{i \in U_D^{t-1,t}} (\hat{\tilde{p}}_i^{t-1})^{s_{iD}^{t-1}}}{\prod_{i \in U_M^{t-1,t}} (\hat{\tilde{p}}_i^{t-1})^{s_{iM}^{t-1}}} \right]^{-s_D^{t-1}} \frac{\prod_{i \in U_M^{t-1,t}} (\hat{\tilde{p}}_i^t)^{\frac{s_{iM}^t - s_{iM}^{t-1}}{2}}}{\prod_{i \in U_M^{t-1,t}} (\hat{\tilde{p}}_i^{t-1})^{\frac{s_{iM}^{t-1} - s_{iM}^t}{2}}}. \tag{43}
$$

In (43), $U_M^{t-1,t} = U^{t-1} \cap U^t$ is the set of (matched) items purchased in both period t-1 and period t, $U_D^{t-1,t}$ is the set of (disappearing) items purchased in period t-1, and possibly in other periods as well, but not in period t, and $U_D^{t-1,t}$ is the set of (new) items purchased in period t, and perhaps also in other periods, but not in period t-1; $s_M^{t-1} = \sum_{i \in U_M^{t-1,t}} s_i^{t-1}$ and $s_M^t = \sum_{i \in U_M^{t-1,t}} s_i^t$ denote the aggregate expenditure shares of the matched items in

periods $t-1$ and t, and $s_D^{t-1} = \sum_{i \in U_D^{t-1,t}} s_i^{t-1} = 1 - s_M^{t-1}$ and $s_N^t = \sum_{i \in U_N^{t-1,t}} s_i^t = 1 - s_M^t$ denote the aggregate expenditure shares of the disappearing items and new items. The item specific expenditure shares in equation (43) have been normalized: $s_{iM}^{t-1} = s_i^{t-1} / s_M^{t-1}$ and $s_{iM}^t = s_i^t / s_M^t$ are the matched items' normalized shares in t-1 and t, and $s_{iD}^{t-1} = s_i^{t-1} / s_D^{t-1}$ and $s_{iN}^t = s_i^t / s_N^t$ are the normalized shares for the unmatched (new and disappearing) items, with $\sum_{i \in S_M^{t-1,t}} s_{iM}^{t-1} = \sum_{i \in S_M^{t-1,t}} s_{iM}^t = \sum_{i \in U_D^{t-1,t}} s_{iD}^{t-1} = \sum_{i \in U_N^{t-1,t}} s_{iN}^t = 1$.

The first component of (43), $\prod_{i \in S_M^{t-1,t}} (p_i^t / p_i^{t-1})^{(s_{iM}^{t-1} + s_{Mi}^t)/2}$, is the adjacent-period matched item Törnqvist index. The second and third components describe the effects of disappearing and new items, respectively. When there are no new or disappearing items between periods $t-1$ and t, then $s_N^t = s_D^{t-1} = 0$ and the chain link is equal to the product of the matched model Törnqvist index and the fourth component of (43). Because we know that the TPD index is transitive but the Törnqvist index is not, we might interpret the fourth component as a factor that eliminates potential drift in the chained Törnqvist index.

A similar decomposition holds true for the time dummy hedonic index, the only difference being that the quality-adjusted prices are different from those obtained for the TPD index. It is likely that the quality-adjusted prices from the TPD model approximate the quality-adjusted prices from the hedonic model better as the sample period grows and the number of matches for a particular item in the data increases. On the other hand, we do not want the sample period to become very long because this conflicts with the underlying assumption of fixed characteristics parameters. So there is a trade-off, but it is difficult to tell what the optimal sample period would be.

The difference between the TPD index and the time dummy hedonic index also depends on the specification of the hedonic model. Krsinich (2014) showed that the two indexes coincide when *i*) all the characteristics are (modelled as) categorical, *ii*) not only the main effects are included but all second and higher order interaction terms as well, and *iii*) the items are identified by their characteristics. This is an interesting, although perhaps not very surprising, finding. Based on this result, Krsinich (2014) interprets the TPD model as the most general version of a hedonic model. She even seems to prefer it to standard models in which most or all interaction terms are excluded. Another, and in our opinion more appealing, interpretation is that the TPD model is a degenerated case of hedonics.

As mentioned before, most scanner data sets do not contain a lot of information on item characteristics. This makes it impossible to define items by their characteristics, which is condition *iii)* above. Moreover, if we had data on all of the characteristics, then obviously we would not use the TPD model but the hedonic model instead. In practice therefore, the TPD model will typically be used for data sets where items are necessarily identified by model numbers or barcodes. This is a practical solution, as the identifiers are readily available. However, the use of such detailed identifiers can create problems which have sometimes been overlooked, as we will argue in section 6.

The TPD model is a special case, or approximation, of the time dummy hedonic model, and so the expenditure share weighted TPD index has a similar structure as the expenditure share weighted time dummy index. Thus, the use of the WLS TPD method to calculate a quality-adjusted unit value index yields a non-hedonic expression similar to equation (36):

$$P_{QAUV}^{0t} = \left[\frac{\sum_{i \in U^0} s_i^0 \exp(u_i^0)}{\sum_{i \in U^t} s_i^t \exp(u_i^t)} \right] P_{TPD}^{0t}, \tag{44}$$

where $u_i^0 = \ln(\hat{p}_i^0 / p_i^0)$ and $u_i^t = \ln(\hat{p}_i^t / p_i^t)$ are now the regression residuals in periods 0 and t $(t = 1,...,T)$ from the TPD model. Just like its hedonic counterpart, the quality-adjusted unit value index (44) is transitive and independent of the choice of base item b, as required.

We are still left with the problem that items that are observed only once during the sample period are zeroed out because the predicted prices are equal to the observed prices, i.e. because the observations lie exactly on the regression surface. Consequently, the impact of entirely new items – items which have not been purchased before – will be ignored. When time passes and the sample period grows, these items will be included in a non-trivial fashion (unless they are not purchased again, which is unlikely). But new items are constantly being introduced on the market, and extending the sample period does not help much in this respect.

Extending the sample period and adding data has another consequence: revision of previously estimated indexes. Revisions apply to any multilateral approach, including TPD and time dummy hedonic approaches. In section 6 below, we will discuss different ways of handling this problem. The most promising method seems to be what Krsinich (2014) calls a window splice.

6. Revisions, homogeneity and lack of matching

Two issues are discussed in this section that need to be resolved before implementing the framework put forward in section 7: revisions and a lack of matching of items. The second issue is particularly relevant when sufficient information on item characteristics is not available and TPD indexes rather than time dummy hedonic indexes are inputs in the quality-adjusted unit value indexes.

6.1 The treatment of revisions

Statistical agencies do not revise their CPI. Once the CPI has been published, it is kept unchanged, even if new data became available that would improve previously published figures.[15] This no revisions policy poses problems for the indexes we want to construct. When data for the next period (period $T+1$ in our case) is added and the multilateral time dummy hedonic or TPD models are re-estimated, the results for all the previous periods $(1,...,T)$ change. A *rolling window approach* overcomes the revisions problem. The models are estimated on the data of an estimation window with fixed length which is shifted forward each period. This procedure raises two questions. What is the optimal window length? And how should the estimates from the most recent window be linked to the existing time series?

Given the underlying assumption of fixed parameters for the characteristics, the window should be as short as possible but at the same time long enough to be able to handle seasonal goods. For time dummy hedonic models, a window length of 13 months thus seems to be a natural choice, assuming that the CPI is published monthly. For TPD models, on the other hand, a window length of 13 months can be too short. The optimal window length depends partly on how the time series is updated, i.e. on the splicing method.

Before going into different methods of splicing, we should mention that every splicing method impairs the transitivity property of multilateral price indexes. So chain drift in the linked time series cannot be completely ruled out. As long as the estimation window is longer than a year, however, it is most unlikely that chain drift becomes a major problem.

---

[15] Many statistical agencies indicate that the latest figure is provisional but no agency accepts continuing revisions of the CPI.

Broadly speaking, two splicing methods can be distinguished; the conventional *movement splice* method and an alternative *window splice* method. To illustrate the two methods, suppose that the window length is 13 months. The standard method works as follows: after moving forward the window one month and re-estimating the model, the most recently estimated month-on-month movement of the index is spliced on to the existing time series. The window splice method, as proposed by Krsinich (2014), splices the entire newly estimated 13-month series on to the index level pertaining to 12 months ago.

The two methods resolve the revisions problem quite differently. The standard approach gives priority to short term changes. Each month-to-month index movement to update the existing time series is based on a single estimation window and is easy to interpret. Longer term changes of the updated time series, such as annual changes where the index of a particular month is compared to the index of 12 months earlier, are based on different estimation windows. This means some chain drift in annual changes might arise. Using the window splice approach, each month-on-month movement is based on adjacent estimation windows and therefore more difficult to interpret. In the Appendix, the two splicing methods are compared.

According to Krsinich (2014), the biggest problem with the standard approach is that "the revised movement for back periods is not incorporated into the longer term index movement" while her method "maintains the integrity of the index over the longer term".[16] Although Krsinich's (2014) splicing method can be applied to time dummy hedonic indexes as well, it seems particularly suited for TPD indexes. The method "is a form of implicit revision, incorporating not only the implicit price movements of new products being introduced, but also enables the fixed effects estimates to be updated as more prices are observed for each product". This is indeed a strong point of window splicing for TPD indexes.

A potential problem with Krsinich's (2014) method is that it does not revise for items that are only observed in the first month of the estimation window; just like newly introduced items in the last period of the window, these disappearing items are zeroed

---

[16] Ivancic, Diewert and Fox (2011) and de Haan and van der Grient (2011) employed a rolling year standard splicing approach to updating GEKS indexes. The GEKS method is an alternative multilateral method. Krsinich (2014) noted that her window splice approach is a simplified version of a suggestion by Melser (2011) for improving the splicing of the rolling year GEKS.

out. In other words, the method is asymmetrical and does not satisfy the multi period time reversal test. To resolve this problem, we could extend the estimation window by 12 months prior to the linking month. The window length becomes 25 months with the month in the middle being the linking month. More than two years of data are needed before indexes can be estimated, but this is hardly a drawback because data of a rather long period of time will be needed anyway for checking the plausibility of the estimated price (and quantity) index series to be published.

High frequency price indexes from scanner data are sometimes very volatile, and it can be argued that smoothing of the existing time series prior to splicing is called for. Compared to the choice of splicing method and window length, this issue is much less important.

6.2 Homogeneity, identification of items, and lack of matching

The CPI Manual (ILO et al., 2004) recommends the use of unit values as prices and unit value indexes as price indexes for homogeneous products. We could say that a product is homogeneous when all varieties are perfect substitutes and consumers are indifferent between them. Products often come in many broadly comparable varieties. Suppose that some varieties differ only in type of packaging or color of the item. From the observed expenditures on these varieties we cannot draw firm conclusions about substitutability because consumers are heterogeneous: some consumers have a preference for a certain type of packaging or item color whereas others are indifferent. Treating each variety as a different item and calculating unit values at this level seems warranted to avoid unit value bias.

Each product variety usually has a barcode (EAN/GTIN) and/or model number. Taking into account what has been mentioned above, barcodes or model numbers are natural keys to identify and distinguish items. Barcodes are always available in scanner data sets, and so observing unit values at the individual item level at a particular point in time is easy. For price comparisons over time, however, the use of barcodes to identify items may cause problems resulting from a lack of matching, in particular for the TPD approach. To get an idea of what can happen, let us look at traditional price collection. The first step followed by a statistical agency when an item is being replaced by a newly sampled item is to check whether or not the two items are comparable. If they are not fully comparable, then a quality adjustment should be made. If they are, the prices of the

two items are directly compared.[17] Direct comparison ensures that *hidden price changes* will be captured. Additional information and some guidelines are needed to determine whether items are comparable. Typically, the practitioner compares a set of important price determining characteristics of the items. If the quantities of those characteristics are the same, then the items are deemed fully comparable. This suggests that, when both items are available during a certain time period, it would be appropriate to calculate unit values across them.

In contrast, most empirical hedonic models are estimated on micro data, i.e. on the data of the various items as identified by barcode or model number. Exceptions are the studies on New Zealand scanner data for consumer electronics by Krsinich (2011) (2013) (2014) and de Haan and Krsinich (2014a) (2014b), who did not have access to such detailed items identifiers.[18] A potential problem in their data set was that unit value bias could not be completely ruled out because the data was aggregated across different types of outlet. Type of outlet is important for attaining homogeneity if the service that goes along with the purchase of a good differs across outlets.[19] More generally, ignoring important characteristics gives rise to unit value bias when items are identified by their characteristics and produces omitted variables bias in hedonic regressions.

If information on important characteristics is lacking, the only feasible solution is to rely on barcodes or article numbers to identify items (and use a TPD model rather than a hedonic model). But such identifiers may be too detailed for CPI purposes since different barcodes or article numbers can relate to items that are perfect substitutes from a consumer's perspective. This issue was mentioned already by Reinsdorf (1999) and de Haan (2002). Item churn will then be overestimated and matched model indexes will be based on fewer matches than desirable. Price changes of items whose barcodes or article numbers changed but otherwise remained unchanged are captured by hedonic methods, though the results will become increasingly model dependent. Hidden price changes are missed by matched model methods, including the TPD method.

---

[17] See for example Chapter 17 in the U.S. Bureau of Labor Statistics *Handbook of Methods*, available at www.bls.gov.

[18] The fit of hedonic regressions in terms of R squared will typically be much better for aggregated data than for the underlying micro data. However, aggregation involves some kind of weighting of the data, and so we should not directly compare the R squared values from the two types of regressions.

[19] An interesting paper on this topic is Ivancic and Fox (2013). Statistics Netherlands stratifies according to retail chain in order to reduce this problem.

6.3 Choice of method

The choice between the time dummy hedonic and the TPD method depends on the type of product and the availability of characteristics information. For example, for consumer electronics products, where quality change due to embodied technical progress plays an important role, hedonic regressions cannot be avoided. Running (time dummy) hedonic regressions on scanner data for consumer electronics is feasible in the Netherlands – the necessary information on characteristics is readily available from websites of retailers or manufacturers[20] or can alternatively be purchased from market research company GfK. Moreover, properly specifying hedonic models for consumer electronics should not be a problem, as previous research has shown.

For products that do not exhibit substantial quality change, TPD indexes can be estimated and turned into quality-adjusted unit value indexes. An example is products sold in supermarkets. Statistics Netherlands already includes scanner data based indexes for supermarkets in the CPI and does not collect any prices at the stores anymore. At the lowest level of aggregation, monthly chained Jevons indexes are calculated; see van der Grient and de Haan (2010). A cut-off procedure, which effectively removes items with low expenditure shares within the product category, acts as a crude form of weighting. Nonetheless, the lack of explicit weighting at the item level is a weakness of the current method. Also, the method does not adjust for hidden price changes. This is becoming a serious problem: simultaneous changing of barcodes and prices (for fully comparable items) becomes more and more common. When applied to items defined by barcodes, the TPD method does not adjust for hidden price changes either. So an important issue is whether we will be able to synthetically match fully comparable items with different barcodes in the absence of detailed product descriptions.

For fashion products, such as clothing, synthetic or statistical matching over time may prove very difficult. This is partly because these products often exhibit a (strongly) seasonal pattern and are unavailable for a long time period before they re-appear. Also, "fashion" can make it hard to determine if two products are fully or nearly comparable. Without being able to match disappearing items and their successors, the TPD method should not be used: matched item indexes, including TPD indexes, will have significant downward bias due to continuous price declines typically observed for individual items

---

[20] Web scraping might be useful to collect data on product characteristics; see also the first topic of the research agenda in section 7.3.

during their availability.[21] So for clothing and other fashion/seasonal goods, alternative methods may have to be used. If not all price determining characteristics are available, time dummy hedonic indexes can still be estimated, but some bias is likely to occur. It remains to be seen whether this is acceptable.

7. Towards large scale use of scanner data

In this final section, we propose a framework for large scale use of scanner data in the Dutch CPI. For retailers who are unable or unwilling to deliver scanner data, Statistics Netherlands wants to use prices extracted from websites to reduce price collection costs. We explain how these online prices would fit into the proposed framework. A research agenda for the near future, explicitly aiming at implementation, is proposed also.

7.1 The proposed framework

Given the limited resources and the need to increase efficiency and transparency, only a small number of fully or semi automated methods can be implemented. Our framework supports this need by using the quality-adjusted unit value index as the sole aggregation formula at the product level and by allowing only time dummy hedonic or TPD indexes to enter the quality-adjusted unit value index formula. This means that, at this stage, we will ignore better methods that may be available. In the future, other methods could be considered, and we encourage further research aiming at the improvement of the use of scanner data in the CPI.

There are several other considerations:

- Product price indexes (quality-adjusted unit value indexes) are constructed at the retail chain level, not for individual stores. If a certain chain exists of store types with different service levels, a breakdown may be necessary.
- Purchases in physical shops and online purchases are both included in the CPI. It is useful to construct separate indexes for retail chains that sell online as well as "offline".

---

[21] De Haan and Hendriks (2013), using online price data, demonstrated that the (unweighted) TPD index for women's T-shirts was severely downward biased, as expected. Greenlees and McClelland (2010), using U.S. scanner data, observed the same phenomenon for rolling year GEKS apparel indexes.

- Different methods, i.e. time dummy hedonic or TPD methods, can be used for a single chain. Apart from data availability issues, the choice of method depends on the type of product and the market circumstances, not so much on the type of retail chain.

- Aggregation of the price indexes across products and across retail chains will be done using fixed expenditure weights which are updated annually. This is in line with current practices.

- There is no need for sampling of items according to expenditure or the use of cut-off procedures.

- Practitioners should not just focus on price changes. At the elementary level of aggregation, they should analyze (graphs of) quality-adjusted unit value indexes, unadjusted unit value indexes, quantity indexes and number of sales.

When scanner data for a retail chain becomes available, the following five-stage procedure could be considered.

1. *Data analysis and cleaning*. Analyzing the data is an important first step in order to answer the following questions. Does the data contain information to classify the items sold into product categories (according to COICOP)? Does it contain additional information on item characteristics? Can the data be merged with data from an external source to include more characteristics information? What is the rate of churn in terms of new and disappearing items? What do the distributions of item sales and expenditures look like? Some cleaning of the raw data may be required, but this should be done with care and not with the aim of smoothing sales or expenditures.

2. *Deciding on methods*. Is it possible to use a single method for the retail chain as input in the quality-adjusted unit value indexes or should a combination of time dummy hedonic and TPD be used? What is the appropriate aggregation level for running regressions? If not all the required characteristics variables are available for performing hedonic regressions, do we accept the resulting omitted variables bias?

3. *Preliminary calculations and plausibility checks*. Are the trends of the quality-adjusted unit value indexes for the various products plausible compared with the current CPI and the (unadjusted) unit value indexes? Does a comparison of the quantity indexes and the number of sales give any reasons for concern?

4. *Possible refinement of the method(s).* The above evaluation could lead to minor adjustments of the method(s) chosen.

5. *Implementation.* This includes the building of a (partially) automated system or the extension of an existing system. It also includes organizational matters, such as making arrangements about regular reviewing and updating of the regression models.

7.2 Online data

Statistics Netherlands has been experimenting with the collection of prices from the Internet through *web scraping*. Online prices could replace prices that are observed by price collectors for the compilation of the CPI.[22] The item samples have traditionally been quite small, particularly to keep things manageable and control costs. A large part of the costs associated with compiling a CPI stems from price collection in the stores. If web scraping turns out to be successful, CPI production costs can probably be reduced significantly (similar to scanner data), even when observing all items displayed on the websites rather than taking small samples.

Quantities and expenditures cannot be observed via the Internet. Weighted price indexes can therefore not be constructed, which is problematic. The lack of weights at the item level is not new to statistical agencies. Without having access to scanner data, the agencies are forced to construct unweighted indexes, for example Jevons indexes. For a particular product, the sample of narrowly defined items is typically kept fixed, at least for some time, and the index is based on a panel of (matched) items to compare "like with like". When new items are introduced into the sample to replace disappearing items, quality adjustments should be carried out.

Given the lack of quantities, it is not possible to construct quality-adjusted unit value indexes from online data. Depending on the available characteristics information, we can of course estimate *unweighted* time dummy hedonic or TPD indexes. If we want to combine scanner data based and online data based price indexes, this would have the

---

[22] Apart from efficiency considerations, web scraping has the advantage that prices can be monitored daily or weekly, allowing the estimation of high frequency price indexes. In the Billion Prices Project, a research initiative at MIT that uses online data to study high frequency price dynamics and inflation, daily price indexes have been calculated for a number of countries, including the Netherlands. The price indexes are currently being published by PriceStats, a private company; see www. PriceStats.com. For an example on Argentina data, see Cavallo (2012).

virtue of using (at some stage) similar methods for different sources. However, it is not a priori clear if, in the unweighted case, these regression based methods are better than traditional unweighted methods. For example, the monthly chained Jevons price indexes currently estimated in the Netherlands for supermarkets are not necessarily inferior to TPD indexes. Research is warranted, in particular to assess how well the implicit quality adjustment of the TPD method performs. A starting point could be the decomposition similar to equiation (43) for the unweighted case as described by de Haan and Hendriks (2013).

There are a number of issues that are specific to web scraping data. Online prices are sometimes above shelf prices due to delivery costs, and this difference may not be stable over time. In turn, shelf prices are likely to differ from average transaction prices (i.e. unit values) as a result of promotional sales and the like. Representativity of the online data is another issue. The range of products shown on websites need not be the same as offered in the corresponding physical outlets and tends to change frequently. Changes made to the website are a potential problem associated with web scraping as it could lead to missing price observations. Also, in particular for clothing and footwear, some online stores classify items that are on sale in a separate clearing sales category, and this category should not be overlooked.[23]

The last point raises an important issue. It is obvious that both regular and sales prices should be taken into account when measuring aggregate price change, but it is not clear how they should be treated. Suppose that prices are observed every day. The price change of a single item can be tracked as long as the item is available. This trajectory would show the true change in offer prices, but it does not necessarily reflect the correct trend for CPI purposes in case of promotional sales. Regular prices may stay the same over time while sales prices show an upward trend, for example. Since promotional sales occur infrequently relative to the number of days with regular prices, the measured overall trend will be flat. However, if consumers mainly buy the item at times of sales,[24] the change in sales prices would be a better indicator of the change in prices actually paid.

---

[23] For more details, see Griffioen, de Haan and Willenborg (2014).

[24] This seems to be the case for many items purchased in supermarkets that can be stored. An example for the most popular brand of detergents in the Netherlands can be found in de Haan (2008); quantities sold at the regular price were negligible but spiked enormously when the item was on sale.

7.3 A research agenda

Finally, we propose a research agenda for the next couple of years. Given the focus of our paper, most of the topics are about scanner data, but we will address online data as well. Statistics Netherlands is keen to collaborate with other statistical agencies that are looking into the use of "big data" in CPI production with the aim of getting support for new developments. This does not mean, of course, that every agency has to agree with the views expressed in this paper, the methods proposed, or the research we intend to undertake.

Information on product characteristics

For classification purposes and for estimating hedonic regression models, information on characteristics is needed. In general, scanner data providers are unable or unwilling to provide this information. Two possibilities could be examined. The first possibility is retrieving item characteristics in the required format from product descriptions that are available in scanner data sets using some form of "text mining". This also applies to online data extracted from retailers' websites. The second possibility is retrieving item characteristics through web scraping. As mentioned earlier, for products like consumer electronics, a wealth of information on product characteristics can be found at websites of retailers, manufacturers or consumer organizations, and this information should be relatively easy to extract.

Matching of comparable items

When using the TPD method or traditional matched item methods, it is important to pick up hidden price changes that might occur when item identifiers such as EAN/GTIN codes are renewed while the items are essentially equivalent. In other words, we should match disappearing items and their (fully comparable) successors. The first step would be to compare the characteristics of disappearing and new items; see the previous point. There are two issues involved. First, it may take some time before natural successors are offered for sale; the extreme case is seasonal products which are unavailable for a long time period. This makes it difficult to match them with their predecessors. Second, in the absence of detailed product descriptions or characteristics, it will not be possible to check if some disappearing item and a new item, or a set of new items, differ in quality. A comparison of the prices and expenditures, coupled with the available information on

characteristics, may provide additional insight and reduce the probability of a mismatch. We suggest performing research into this type of statistical matching using scanner data for supermarket chains where Statistics Netherlands has gained experience in observing and adjusting for hidden price changes, or what CPI practitioners sometimes refer to as re-launches.

Splicing method and optimal window length

Pooled time dummy hedonic and TPD indexes are (repeatedly) estimated on data of a fixed time period, the estimation window. Previous research has shown that the results are dependent on the window length. The results also depend on the splicing method, i.e. on the way the indexes obtained from subsequent estimation windows are linked to obtain a non-revisable time series. Different splicing methods, including the standard method and variants of the window splice mentioned in section 6.1, should be compared and indicators be developed for choosing the optimal window length. We would expect the standard splicing to work satisfactorily for the time dummy hedonic approach. Thus, we suggest estimating TPD indexes and the resulting quality-adjusted unit value indexes from supermarket scanner data.

TPD method versus traditional methods (non-seasonal goods)

For non-seasonal goods, unweighted TPD seems to be a promising method for dealing with online data extracted from retailers' websites. However, at this moment there is little empirical evidence to support the view that this method is "better" than traditional methods. Also, as was mentioned in section 7.2, online data has a number of specific features that have to be taken into account, and the sets of observed prices can differ significantly from traditional small prices samples. Therefore, research into this topic exists of at least two components: choice of index number formula (TPD index versus e.g. chained Jevons index) and selection of items (all items displayed on the website, or a large sample thereof, versus a relatively small sample).

Consumer electronics: time dummy hedonic

A lot of theoretical and empirical research has been performed on New Zealand scanner data for consumer electronics products. Statistics New Zealand recently implemented the so-called ITRYGEKS method (see de Haan and Krsinich, 2014a). It will be useful

to explore how this method compares to the quality-adjusted unit value index method proposed by us. The two methods both use weighted time dummy hedonic price indexes as inputs; ITRYGEKS makes use of bilateral indexes (in a GEKS framework) whereas our method uses multilateral indexes. An advantage of our method is that it is easier to implement and explain to users. For research purposes, scanner data of a single retailer will suffice, possibly data from an online retailer for which Statistics Netherlands has been receiving scanner data for some time now. Information on item characteristics can be purchased from GfK. In addition, this information could be collected through web scraping (see the first point) and compared with the GfK data to find out how well web scraping works.

Combining characteristics information and the TPD method

As mentioned earlier, it is not possible to include both characteristics and item dummy variables in a time dummy model. Chessa (2014) estimated TPD models on relatively homogeneous product categories rather than micro data as given by EAN/GTIN codes. He defined the product categories by cross-classifying the (most important) categorical characteristics extracted from the product descriptions found in the scanner data set of a particular retailer. It will be useful to examine, both theoretically end empirically, how the results differ from the results obtained by directly estimating a time dummy hedonic model using the same set of characteristics. Chessa also used an alternative approach to splicing/chaining, which could be addressed in the topic above on the choice of splicing method and optimal window length.

Upper level aggregation

Current practice is to aggregate up the price indexes using (annually) fixed expenditure weights. However, if time dummy hedonic or TPD indexes are estimated from scanner data at levels below the lowest publication level, it would be better to aggregate them up to the publication level using a superlative index number formula, such as the Fisher formula, to take into account substitution effects. In general, we do not expect to find major differences compared with current fixed weight practices, but in some instances significant differences might occur. This piece of empirical research is straightforward and can be applied to any scanner data set.

Appendix: A comparison of standard splicing and window splicing

De Haan (2015) compares rolling year time dummy indexes based on standard splicing and window splicing. In this appendix, we summarize the most important findings for multilateral time dummy hedonic and time-product dummy indexes. Results from the estimation window starting in period x are indicated by (x). For example, $P_{TD}^{0,t}(0)$ is the expenditure share weighted time dummy hedonic index going from period 0 to period t, estimated on the data of the sample period $0,...,T$. After moving forward the estimation window by one period, the time dummy index between periods 1 and t is denoted by $P_{TD}^{1,t}(1)$.

The standard *movement splice* extends the existing time series $P_{TD}^{0,1}(0)....P_{TD}^{0,T}(0)$ by multiplying $P_{TD}^{0,T}(0)$ by the movement $P_{TD}^{1,T+1}(1)/P_{TD}^{1,T}(1)$. That is, the time dummy index with a movement splice (TDMS) for the "new" period $T+1$ and index reference period 0 is calculated as

$$P_{TDMS}^{0,T+1} = P_{TD}^{0,T}(0) \times \frac{P_{TD}^{1,T+1}(1)}{P_{TD}^{1,T}(1)} = P_{TD}^{0,T}(0) \times P_{TD}^{T,T+1}(1) = P_{TD}^{0,1}(0) \times P_{TD}^{1,T}(0) \times P_{TD}^{T,T+1}(1), \qquad \text{(A.1)}$$

using transitivity of the time dummy index. The *window splice* method extends the time series by multiplying the time dummy index for period 1, $P_{TD}^{0,1}(0)$, by the index going from period 1 to period $T+1$, $P_{TD}^{1,T+1}(1)$, based on the new estimation window. Thus, the time dummy index with a window splice (TDWS) for period $T+1$ with index reference period 0 is calculated as

$$P_{TDWS}^{0,T+1} = P_{TD}^{0,1}(0) \times P_{TD}^{1,T+1}(1) = P_{TD}^{0,1}(0) \times P_{TD}^{1,T}(1) \times P_{TD}^{T,T+1}(1). \qquad \text{(A.2)}$$

The ratio of $P_{TDWS}^{0,T+1}$ and $P_{TDMS}^{0,T+1}$ can be written as

$$\frac{P_{TDWS}^{0,T+1}}{P_{TDMS}^{0,T+1}} = \frac{P_{TD}^{1,T}(1)}{P_{TD}^{1,T}(0)} = \exp\left[\sum_{k=1}^{K}[\hat{\beta}_k(1) - \hat{\beta}_k(0)][\sum_{i \in U^1} s_i^1 z_{ik} - \sum_{i \in U^T} s_i^T z_{ik}]\right], \qquad \text{(A.3)}$$

where $\hat{\beta}_k(0)$ and $\hat{\beta}_k(1)$ are the parameter estimates from the two estimation windows. So if the parameter estimates from the two estimation windows are the same for all of the characteristics, then the window splice and the standard splice will produce identical results.

The last movement of the TDWS index, $P_{TDWS}^{0,T+1}/P_{TDWS}^{0,T}$, can be decomposed as

$$\frac{P_{TDWS}^{0,T+1}}{P_{TDWS}^{0,T}} = \frac{P_{TDWS}^{0,T+1}}{P_{TD}^{0,T}(0)} = \frac{P_{TD}^{0,1}(0) \times P_{TD}^{1,T+1}(1)}{P_{TD}^{0,1}(0) \times P_{TD}^{1,T}(0)} = \frac{P_{TD}^{1,T+1}(1)}{P_{TD}^{1,T}(0)} = \frac{P_{TD}^{1,T+1}(1)}{P_{TD}^{1,T}(1)} \times \frac{P_{TD}^{1,T}(1)}{P_{TD}^{1,T}(0)} . \qquad \text{(A.4)}$$

The first factor of (A.4) is identical to the index movement used in standard splicing; see (A.1). Because the month-on-month change from the standard splice depends on a single estimation window, it is easy to interpret. The month-on-month change from the window splice depends on two (adjacent) estimation windows.

The time-product dummy or fixed effects (FE) indexes with a movement splice (FEMS) and a window splice (FEWS), respectively, for period $T+1$ and with reference period 0 are calculated as

$$P_{FEMS}^{0,T+1} = P_{FE}^{0,T}(0) \times \frac{P_{FE}^{1,T+1}(1)}{P_{FE}^{1,T}(1)} = P_{FE}^{0,T}(0) \times P_{FE}^{T,T+1}(1) = P_{FE}^{0,1}(0) \times P_{FE}^{1,T}(0) \times P_{FE}^{T,T+1}(1); \qquad \text{(A.5)}$$

$$P_{FEWS}^{0,T+1} = P_{FE}^{0,1}(0) \times P_{FE}^{1,T+1}(1) = P_{FE}^{0,1}(0) \times P_{FE}^{1,T}(1) \times P_{FE}^{T,T+1}(1) . \qquad \text{(A.6)}$$

The only difference between $P_{FEMS}^{0,T+1}$ and $P_{FEWS}^{0,T+1}$ is the use of $P_{FE}^{1,T}(0)$ rather than $P_{FE}^{1,T}(1)$ in the above decompositions, similar to the hedonic counterparts.

To evaluate the effect of new items on the FEMS and FEWS indexes, suppose first that a new item was introduced in period $T+1$. This item affects neither $P_{FEMS}^{0,T+1}$ nor $P_{FEWS}^{0,T+1}$ because it is observed only once in the estimation window (1), hence zeroed out, and unobserved in the estimation window (0). Suppose next that a new item was being introduced in the previous period T. This item will usually be purchased in period $T+1$ as well; its price change between and T to $T+1$ affects $P_{FE}^{T,T+1}(1)$ in (A.5) and (A.6) and therefore impacts on both $P_{FEMS}^{0,T+1}$ and $P_{FEWS}^{0,T+1}$. In addition, the FEWS method incorporates the effect of this item into the price movement for back periods through $P_{FE}^{1,T}(1)$ whereas the FEMS method does not "revise" this longer term price movement because $P_{FE}^{1,T}(0)$ is based on the previous estimation window. This form of implicit revision is a strong point of Krsinich's (2014) FEWS method.

The FEWS index for period $T+1$ can be written as

$$P_{FEWS}^{0,T+1} = \frac{P_{FE}^{1,T}(1)}{P_{FE}^{1,T}(0)} \times P_{FEMS}^{0,T+1}, \qquad \text{(A.7)}$$

where the ratio of $P_{FE}^{1,T}(0)$ and $P_{FE}^{1,T}(1)$ equals

$$\frac{P_{FE}^{1,T}(1)}{P_{FE}^{1,T}(0)} = \exp\left[\sum_{i \in U^1} s_i^1 [\hat{\gamma}_i(1) - \hat{\gamma}_i(0)] - \sum_{i \in U^T} s_i^T [\hat{\gamma}_i(1) - \hat{\gamma}_i(0)]\right]. \qquad \text{(A.8)}$$

The predicted price for the item being introduced in period T from the regression ran on the window (0) with or without this item equals $\hat{p}_i^0(0) = p_i^0 = \exp(\hat{\delta}^0(0))\exp(\hat{\gamma}_i(0))$. In other words, the new item's fixed effect is trivially estimated by $\hat{\gamma}_i(0) = \ln(p_i^0) - \hat{\delta}^0(0)$, where $\hat{\delta}^0(0)$ is the intercept. The FEWS method updates the trivial estimate $\hat{\gamma}_i(0)$ for the new item, which belongs to U^T but not U^1 in (A.8), by the realistic estimate $\hat{\gamma}_i(1)$. It also updates the fixed effects estimates for the other items while the FEMS method is based on the previous fixed effects estimates.

References

Aizcorbe, A., C. Corrado and M. Doms (2003), "When Do Matched-Model and Hedonic Techniques Yield Similar Price Measures?", Working Paper no. 2003-14, Federal Reserve Bank of San Francisco.

von Auer, L. (2014), "The Generalized Unit Value Index Family", *Review of Income and Wealth* 60, 843-861.

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

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

Bradley, R. (2005), "Pitfalls of Using Unit Values as a Price Measure or Price Index", *Journal of Economic and Social Measurement* 30, 39-61.

Cavallo, A. (2012), "Online and Official Price Indexes: Measuring Argentina's Inflation", *Journal of Monetary Economics*, online version, 25 October 2012.

Chessa, A. (2014), "Een prijsindexmethode voor HEMA-scannerdata", Research paper [in Dutch], Statistics Netherlands, The Hague, The Netherlands.

Bradley, R., B. Cook, S.G. Leaver and B. Moulton (1997), "An Overview of Research on Potential Uses of Scanner Data in the U.S. CPI", Paper presented at the third meeting of the Ottawa Group, 16-18 April 1997, Voorburg, The Netherlands.

Dalén, J. (1997), "Experiments with Swedish Scanner Data", Paper presented at the third meeting of the Ottawa Group, 16-18 April 1997, Voorburg, The Netherlands.

Dalén, J. (1998), "On the Statistical Objective of a Laspeyres' Price Index", Paper presented at the fourth meeting of the Ottawa Group, 22-24 April 1998, Washington D.C., USA.

Dalén, J. (2001), "Statistical Targets for Price Indexes in Dynamic Universes", Paper presented at the sixth meeting of the Ottawa Group, 2-6 April 2001, Canberra, Australia.

Dalén, J. (2014), "The Use of Unit Values in Scanner Data", Paper presented at the Scanner Data Workshop, 2-3 October 2014, Vienna, Austria.

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*, Studies in Income and Wealth, Vol. 61. Chicago: University of Chicago Press.

Diewert, W.E. (2004), "On the Stochastic Approach to Linking the Regions in the ICP", Discussion Paper no. 04-16, Department of Economics, The 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*, Studies in Income and Wealth, Vol. 70. University of Chicago Press, Chicago.

van Garderen, K.J. and C. Shah (2002), "Exact Interpretation of Dummy Variables in Semilogarithmic Equations", *Econometrics Journal* 5, 149-159.

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

van der Grient, H. and J. de Haan (2010), "The Use of Supermarket Scanner Data Price in the Dutch CPI", Paper presented at the meeting of the group of experts on consumer price indices, 10-12 May 2010, Geneva, Switzerland.

van der Grient, H. and J. de Haan (2011), "Scanner Data Price Indexes: The "Dutch" Method versus Rolling Year GEKS", Paper presented at the twelth meeting of the Ottawa Group, 4-6 May 2011, Wellington, New Zealand.

Griffioen, R., J. de Haan and L.Willenborg (2014), "Collecting Cloting Data from the Internet", Paper presented at the meeting of the group of experts on consumer price indices, 26-28 May 2014, Geneva, Switzerland.

de Haan, J. (2002), "Generalised Fisher Price Indexes and the Use of Scanner Data in the Consumer Price Index", *Journal of Official Statistics* 1, 61-85.

de Haan, J. (2004a), "Estimating Quality-Adjusted Unit Value Indexes: Evidence from Scanner Data", Paper presented at the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, 30 June – 3 July 2004, Vancouver, Canada.

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

de Haan, J. (2008), "Reducing Drift in Chained Superlative Price Indexes for Highly Disaggregated Data", Paper presented at the Economic Measurement Group Workshop, 10-12 December 2008, Sydney, Australia.

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

de Haan, J. (2015), "Rolling Year Time Dummy Indexes and the Choice of Splicing Method", unpublished paper, Statistics Netherlands.

de Haan, J. and E. Opperdoes (1997a), "Estimation of the Coffee Price Index Using Scanner Data: Simulation of Official Practices", Paper presented at the third meeting of the Ottawa Group, 16-18 April 1997, Voorburg, The Netherlands.

de Haan, J. and E. Opperdoes (1997b), "Estimation of the Coffee Price Index Using Scanner Data: The Choice of the Micro Index", Paper presented at the third meeting of the Ottawa Group, 16-18 April 1997, Voorburg, The Netherlands.

de Haan, J., E. Opperdoes and C.M. Schut (1999), "Item Selection in the Consumer Price Index: Cut-off versus Probability Sampling", *Survey Methodology* 32, 197-216.

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 R. Hendriks (2013), "Online Data, Fixed Effects and the Construction of High-Frequency Price Indexes", Paper presented at the Economic Measurement Group Workshop, 28-29 November 2013, Sydney, Australia.

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

de Haan, J. and F. Krsinich (2014b), "Time Dummy Hedonic and Quality-Adjusted Unit Value Indexes: Do They Really Differ?", Paper presented at the Society for Economic Measurement Conference, 18-20 August 2014, Chicago, U.S.

Hawkes, W.J. (1997), "Reconciliation of Consumer Price Index Trends with Corresponding Trends in Average Prices for Quasi-Homogeneous Goods Using Scanning Data", Paper presented at the third meeting of the Ottawa Group, 16-18 April 1997, Voorburg, The Netherlands.

Hawkes, W.J. (1998), "The Use of Scanner Data in Reconciling Time-Series (Consumer Price Index) and Geographic (Place-to- Place) Comparisons", Paper presented at

the fourth meeting of the Ottawa Group, 22-24 April 1998, Washington D.C., USA.

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

Ioannides, C. and M. Silver (1997), "Chained, Exact and Superlative Hedonic Price Changes: Estimates from Micro Data", Paper presented at the third meeting of the Ottawa Group, 16-18 April 1997, Voorburg, The Netherlands.

Ivancic, L. (2007), *Scanner Data and the Construction of Price Indices*, Ph.D. thesis, University of New South Wales, Sydney, Australia.

Ivancic, L., W.E. Diewert and K.J. Fox (2009), "Scanner Data, Time Aggregation and the Construction of Price Indexes", Discussion paper 09-09, Department of Economics, The University of British Columbia, Vancouver, Canada.

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.

Jain, M. and J. Caddy (2001), "Using Scanner Data to Explore Unit Value Indexes", Paper presented at the sixth meeting of the Ottawa Group, 2-6 April 2001, Canberra, Australia.

Jain, M. and R. Abello (2001), "Construction of Price Indexes and Exploration of Biases Using Scanner Data", Paper presented at the sixth meeting of the Ottawa Group, 2-6 April 2001, Canberra, Australia.

Kennedy, P.E. (1981), "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations", *American Economic Review* 71, 801.

Krsinich, F. (2011), "Price Indexes from Scanner Data: A Comparison of Different Methods", Paper presented at the twelfth meeting of the Ottawa Group, 4-6 May 2011, Wellington, New Zealand.

Krsinich, F. (2013), "Using the Rolling Year Time Product Dummy Method for Quality Adjustment in the Case of Unobserved Characteristics", Paper presented at the thirteenth meeting of the Ottawa Group, 1-3 May 2013, Copenhagen, Denmark.

Krsinich, F. (2014), "The FEWS Index: Fixed Effects with a Window Splice – Non-Revisable Quality-Adjusted Price Indexes with No Characteristic Information", Paper presented at the meeting of the group of experts on consumer price indices, 26-28 May 2014, Geneva, Switzerland.

Melser, D. (2011), "Constructing High Frequency Price Indexes Using Scanner Data", Paper presented at the twelfth meeting of the Ottawa Group, 4-6 May 2011, Wellington, New Zealand.

Okamoto, M. and T. Salou (2001), "Comparisons of Hedonic Method and Matched Models Using Scanner Data: The Case of PCs, TVs, and Digital Cameras, Paper presented at the sixth meeting of the Ottawa Group, 2-6 April 2001, Canberra, Australia.

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

Richardson, D. (2001), "Scanner Indexes for the CPI", Paper presented at the sixth meeting of the Ottawa Group, 2-6 April 2001, Canberra, Australia.

Saglio, A. (1994), "Comparative Changes in Average Price and Price Index: Two Case Studies", Paper presented at the first meeting of the Ottawa Group, 31 October -4 November 1994, Ottawa, Canada.

Silver, M. (2008), "Do Unit Value Export, Import and Terms of Trade Indexes Represent or Misrepresent Price Indexes?", *IMF Staff Papers* 55 (September).

Silver, M. (2009), "An Index Number Formula Problem: The Aggregation of Broadly Comparable Items", IMF Working Paper no. 09/19, International Monetary Fund, Washington D.C., U.S.A.

Silver, M., C. Ioannides and M. Haworth (1997), "Hedonic Quality Adjustments for Non-Comparable Items for Consumer Price Indices", Paper presented at the third meeting of the Ottawa Group, 16-18 April 1997, Voorburg, The Netherlands.

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, 269-281.

Silver, M. and S. Heravi (2007a), "The Difference Between Hedonic Imputation Indexes and Time Dummy Hedonic Indexes", *Journal of Business & Economic Statistics* 25, 239-246.

Silver, M. and S. Heravi (2007b), "Why Elementary Price Index Number Formulas Differ: Evidence on Price Dispersion", *Journal of Econometrics* 140, 874-883.

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.