Testing unit value data price indices

Li-Chun Zhang[1,2], Ingvild Johansen[2], and Ragnhild Nygaard[2]

[1] *University of Southampton (email: L.Zhang@soton.ac.uk)*
[2] *Statistics Norway*

Abstract

Several index formulae have been proposed for scanner data in the recent literature. But there is currently a lack of consensus on how to evaluate them. We propose two overriding aspects of an index method, particularly important for the greater uptake of scanner data in practice: (a) it should accommodate available data of all items and not only the persistent ones over time, (b) it should as much as possible reduce the cost related to matching the actual items. Both point towards an index theory that extends beyond the traditional matched-model approach. In this paper, we present some work in two directions. First, we propose five formal tests explicitly formulated for a dynamic item universe, and compare theoretically several bilateral and multilateral indices in light of these tests, provided item-matching is unproblematic. Next, we outline an approach to segmented price indices, which can minimise the resource required for item-matching in practice, and illustrate the basic ideas using the Norwegian scanner data on food and non-alcoholic beverages. Future research should aim to develop shared explicit empirical criteria for well-behaving indices in practice.

1 Introduction

Scanner data was introduced in the Norwegian CPI in the late 90s. The methodology has changed overtime. The current method for the index of food and non-alcoholic beverages has been in use since 2013. There are three known shortcomings. First, it is based on persistent items (GTINs) over time, thus does not incorporate in-coming new items immediately and is not fully responsive to the constantly changing item universe. Second, the underlying matched-model approach, which requires the identification of persistent items over time despite changes in the GTIN, would have been resource consuming if it were pursued with full rigour. This is also a major challenge that has limited the uptake of scanner data in several other consumer groups of the CPI, where matching is more difficult and demanding. Third, Jevons index is currently applied at the elementary level, which does not make use of the available quantity information explicitly. At Statistics Norway, these issues have motivated research into more generic price index methodology, in order to make sound and effective use of Big-data sources that contain both unit value prices and transaction totals of the associated goods or goods-bundles.

Remark The term **matched-model** characterises an index theory approach, whereas we use the term *persistent item* to refer to an item that actually exists in the universe over time. Thus, e.g., the matched-model approach requires one to identify the persistent items over time.

Historic background Data for the Norwegian CPI has traditionally been collected through questionnaires, first through paper and then by web questionnaires where outlets report price and related information for a sample of products. Digitalisation has made it possible to collect full-scale price data. Scanner data in particular has been utilised in the Norwegian CPI for about two decades. In the beginning, only prices for a selection of items were used to replace price collection from questionnaires. The range of items included has been gradually expanded. The current approach is based on *cut-off thresholds*, which means that only the items with the highest expenditure shares are included. The cut-off limit is set empirically, in a way that more than 50 per cent of the items are excluded, but as much as 80 per cent of expenditure is included in the index. The unweighted Jevons index is calculated as a monthly chained matched-model index at an elementary level below COICOP6. Ideally, changes of consumption patterns should be reflected in a cost-of-living index, and products should be weighted according to their economic importance in a cost-of-goods index. However, the expenditure shares at the elementary level are not continuously updated in the current monthly chained index, due to the potential risk that the index might suffer from chain drift as a result.

Diversity of e-data index methods More than 20 per cent of the Norwegian CPI, measured in consumer-group weight shares of the total CPI, consists of scanner data. The main areas covered by scanner data are food and non-alcoholic beverages, pharmaceutical products and petrol. Together with other types of electronic data (*e-data*), including internet data, nearly half of the Norwegian CPI, in terms of the total weight share, is based on e-data. However, different index methods are currently implemented for different consumer groups. This has resulted in a complex production system, which is both resource demanding and potentially vulnerable for mistakes.

Practical challenge of item-matching Regardless of the choice of index formula, a major practical challenge with scanner data as well as other types of e-data is the need for item-matching over time. By ***item matching*** we mean the identification of the set of unique items over time, *no matter* how the items are coded or indexed in the datasets available. Discontinued item codes (e.g. GTIN) and insufficient metadata make it difficult to handle item-matching in a fully automatic manner, without which Big-data means Big-trouble literally speaking. For instance, as a result of imperfect item-matching, the price development related to relaunches may not be captured. This is one of the most important factors that have limited the uptake of scanner data in other consumer groups in the Norwegian CPI, like clothing and electronics, which have higher item churn rates than the grocery market. A potential theoretical implication is that, despite the increasing use of grocery market scanner data in official CPIs among NSOs, the underlying matched-model approach to price index used today may not be appropriate for other consumer groups.

Terms of reference We propose that two overriding aspects of an index method are particularly important for the greater uptake of scanner data in practice: (a) it should accommodate available data of all items and not only the persistent ones over time, (b) it should keep the cost related to matching the actual items at a sustainable level. Both point towards an index theory that extends beyond the traditional matched-model approach. Realistically speaking, ***we do not expect to be able to arrive at an ideal index formula, but aim at index methods that as much as possible fulfill these two overriding concerns***, which cover a number of often mentioned desirable features as special cases, including the following ones:

- incorporate quantity data explicitly to allow for a broader coverage of product offers,
- generic in the sense that the method can be used across different consumer groups,
- capture the dynamic product universe,
- handles substitution, which is only possible if in-coming items are included immediately,
- handles practical challenges in an effective manner – must avoid manual interference.

In addition, we maintain that both the cost-of-living and cost-of-goods perspectives are relevant, e.g. an index method should be

- harmonised with other NSOs and consistent with HICP principles and recommendations,
- transparent and easy to communicate to users.

Analytic approach of the paper Several index formulae have been proposed for scanner data in the recent literature. But there is currently a lack of consensus on how to evaluate them. The traditional axiomatic price index tests (Fisher, 1922) are all defined for a *fixed* item universe. Diewert (1999) and Balk (2001) outline tests for international comparisons. In Section 2.2, we propose five formal tests explicitly formulated for a dynamic item universe, and compare several bilateral and multilateral indices in light of these tests in Section 2.3, provided item-matching is unproblematic. These include in particular a modified Geary-Khamis index (MGK), which is closely related to the quality adjusted unit value index e.g. discussed by Chessa (2016), and an extension of it. In addition to the tests that specify certain necessary properties, we analyse in Section 2.4 the volatility of two aforementioned bilateral indices, by comparing them to their counterparts which are calculated only based on the persistent items.

Practical segmented price index In Section 3 we outline an approach to segmented price indices, which can minimise the resource required for item-matching in practice, and provide a theoretical interpretation of the ideal segment concept (Section 3.3) in terms of a cost-of-living index. In Section 4 we illustrate the basic ideas using the Norwegian scanner data on food and non-alcoholic beverages. While this demonstrates clearly that the segmented price index approach can produce comparable results to the existing methodology, with minimum effort for item-matching, it also raises several questions that require further study in order to better understand the properties of segmented price indices. In this respect, we strongly call for future research to develop shared explicit *empirical criteria* of well-behaving indices in practice.

2 On matched-model indices based on unit value data

We start by introducing formally the notation and terms that are necessary for the discussions to follow. Next, we propose 5 tests and motivate them both from the perspectives of ***cost-of-goods index (COGI)*** and ***cost-of-living index (COLI)***. We then examine the Geary-Khamis (GK, Geary, 1958) index using these tests, which is referred to as the Quality Adjusted Unit Value (QAUV) index by Chessa (2016). In light of the test results, we shall propose a modified GK (MGK) index and an extension of it. Considerations of a couple of other indices in the literature are given in the Appendices. At this stage ***we do not have an index that in general satisfies all the 5 tests***, which suggests that the set of proposed tests can be used as necessary conditions for any completely general index. We close the section with some additional evaluation of the theoretical properties of the MGK index and its extension.

2.1 Preliminary notation and terms

Consider a given ***item universe*** in period t, denoted by $U_t = \{1, 2, ..., N_t\}$, which constitutes a subset (or *sector*) of the entire CPI-universe. For instance, U_t may refer to all food and non-alcoholic beverages sold at supermarkets, or all personal computers sold at electro warehouses, etc. By ***unit value data*** we mean one has available the unit value price (p_i^t) and transaction total (v_i^t), for *any* $i \in U_t$, where $v_i^t = p_i^t q_i^t$ and q_i^t denote the corresponding quantity total. Denote by $q(U_t) = \{q_i^t; i \in U_t\}$ the set of item quantities and by $p(U_t) = \{p_i^t; i \in U_t\}$ that of unit value prices. Denote by $D_t = D(U_t) = \{p(U_t), q(U_t); U_t\}$ all the unit value data at t.

We observe the following. The unit value price refers to the average price of an item in the period t, even though the actual transaction price of the item may change many times over the same period. In general we allow the items to be specific for an outlet or chain. For example, iPhone 7 sold at two different outlets can be treated as two different items. Scanner data are typically available as unit value data. The term unit value data is more precise, and covers other situations where the term scanner data may appear less natural.

Denote by (U_0, U_t) the ***comparison universe***, for which we seek a price index from 0 to t, denoted by $P^{0,t}$. We refer to 0 as the *base* period and t the *statistical* period. Denote by $U_{R(0,t)} = \{U_r; r \in R\}$ the ***reference universe***, where R is the set of reference periods involved, and the notation $U_{R(0,t)}$ emphasises the dependence of R on the comparison universe – we may suppress $(0, t)$ whenever the context is clear. We consider price indices of the form

$$P^{0,t} = f(D_R)$$

i.e. based on the data $D_R = D(U_R)$. Two choices of R are most immediate, i.e.

$$R_B = \{0, t\} \quad \text{and} \quad U_{R_B} = U_0 \cup U_t$$
$$R_M = \{0, 1, ..., t\} \quad \text{and} \quad U_{R_M} = U_0 \cup U_1 \cup \cdots \cup U_t$$

where R_B implies direct comparison between 0 and t, e.g. December 2014 (0) and July 2015 (t), and R_M implies all the periods from 0 to t, e.g. December 2014 to July 2015. An index with reference universe U_{R_B} is referred to as a ***bilateral*** index, whereas it is ***multilateral*** with U_{R_M}. We notice that it is possible to use another reference universe for a multilateral index, such as a *rolling window* of fixed length counting backwards from t. However, the distinction is not important for the discussion in this paper, so we shall adopt the convention of multilateral reference universe U_{R_M} as specified above, unless it is explicitly stated otherwise.

2.2 Tests for dynamic item universe

Traditional tests are defined for a *fixed* item universe. Below we formulate 5 tests explicitly in the context of a *dynamic* item universe, and provide the motivation for each of them both from the perspectives of COGI and COLI. Afterwards a discussion is given for why we do not include here the traditional transitivity test.

Identity test (T1) If $U_0 = U_t$ and $p_i^0 \equiv p_i^t$ for any $i \in U_0$, then $P^{0,t} = 1$.

Since the item universe is the same at 0 and t, so must be the items eligible for a COGI, when the comparison universe is (U_0, U_t). Thus, despite changes of the item universe which are assumed to take place in general, i.e. $U_r \neq U_0$ for $r \neq 0$ and $r \in R_M$, the identity tests can be motivated for a COGI. Now that all the prices are the same at 0 and t, a fixed basket-of-goods index must necessarily be 1, regardless of how the reference quantities of the goods are calculated. So a COGI satisfies the identity test. Next, consider a COLI. Let $V^{0,t} = V^t/V^0 = \sum_{i \in U_t} q_i^t p_i^t / \sum_{i \in U_0} q_i^0 p_i^0$ be the ratio of total expenditures. Under the stipulated setting, it is obviously possible to maintain the same utility without changing the total expenditure. Thus, insofar as $V^{0,t} \neq 1$, all the change in expenditure must be attributed to the change in utility but *not* prices, under the assumption of rational consumer behaviour. A COLI should therefore be equal to 1.

Remark Provided a dynamic universe, where $U_r \neq U_0$ for $0 < r < t$, and $p_i^r \neq p_i^0$ for some $i \in U_0 \cap U_r$, an index that satisfies the identity test is said ***not to drift*** in this situation. We notice that chain drift is often contrasted with transitivity. However, as will be discussed later, we find it difficult to formulate a transitivity test for a dynamic universe. Thus the test T1 is the only test we can formulate at the moment with respect to chain drift.

Fixed basket test (T2) If $U_0 = U_t$ and $q_i^0 \equiv q_i^t$ for any $i \in U_0$, then $P^{0,t} = V^{0,t}$.

T2 is obvious for a COGI, *provided* one restricts the reference quantities of the basket items to $q(U_0)$ and $q(U_t)$. Otherwise, in a dynamic universe, one may have $q_i^r \neq q_i^0$, for $0 < r < t$, so that the corresponding reference quantity can differ from $q_i^0 = q_i^t$, in which case a fixed basket-of-goods index may not be equal to $V^{0,t}$. It follows that, ***in order for a COGI to satisfy T2, one should adopt a bilateral index, and avoid the use of multilateral indices***. Paradoxically, the test seems more readily motivated for a COLI. Despite a quantity index is generally not the same as a utility index, there is practically no way to vary any measure of utility as long as the quantities remain the same. It follows that no utility adjustment of $V^{0,t}$ is practically feasible here, and any index targeted at a COLI would necessarily be equal to $V^{0,t}$.

Remark In the case of $t = 1$, both tests T1 and T2 are in fact tests for a fixed item universe. In this sense, we require that a dynamic-universe index should not have counter-intuitive properties in the special case of fixed universe.

Upper bound test (T3) If $U_0 \subseteq U_t$, and $p_i^t \leq p_i^0$ for all $i \in U_0$, then $P^{0,t} \leq 1$.

Let $U_{0t} = U_0 \cap U_t$ be the *persistent* item universe at 0 and t. That is, the item universe may be constant if $U_0 = U_t$ or strictly expanding if $U_0 \subset U_t$, and the price of each persistent item is the same or reduced, i.e. $p_i^t \leq p_i^0$ for all $i \in U_0 = U_{0t}$. To motivate the test, consider the following. Firstly, suppose substitution does not occur, in which case $q_i^t = q_i^0$ for all $i \in U_0$ and $q_i^t = 0$ for all $i \in U_{t\backslash 0}$, even if $U_{t\backslash 0}$ is nonempty. The actual comparison universe reduces then to U_{0t}, so that the test T2 applies, yielding $P^{0,t} = V^{0,t} \leq 1$ under the stipulated setting. Next, suppose substitution occurs only among the persistent items, i.e. $q_i^t = 0$ for $i \in U_{t\backslash 0}$ and $q_i^t \neq q_i^0$ for some $i \in U_{0t}$. Substitution can only be accounted for from the perspective of COLI. Now, given the actual $\{q_i^t; i \in U_{0t}\}$ and the corresponding utility at t, it cannot cost less for the same $\{q_i^t; i \in U_{0t}\}$ at 0 since $p_i^t \leq p_i^0$ for all $i \in U_{0t}$. It follows that a COLI must be less or equal to 1. Finally, suppose substitution also involves the items in $U_{t\backslash 0}$. Let $\{\tilde{q}_i^t; i \in U_{0t}\}$ be a hypothetical set of persistent items that would have yielded the same utility as the actual $\{q_i^t; i \in U_t\}$. Owing to rational behaviour, the expenditure of $\{\tilde{q}_i^t; i \in U_{0t}\}$ at t cannot be less than the actual expenditure of $\{q_i^t; i \in U_t\}$; whereas the expenditure of $\{\tilde{q}_i^t; i \in U_{0t}\}$ at 0 cannot be less than that at t. It follows again that a COLI must be less or equal to 1.

Lower bound test (T4) If $U_t \subseteq U_0$, and $p_i^t \geq p_i^0$ for all $i \in U_t$, then $P^{0,t} \geq 1$.

That is, the item universe may be constant or strictly shrinking, and the price of each persistent item is the same or increased. Firstly, suppose substitution does not occur, in which case $q_i^t = q_i^0$ for all $i \in U_{0t}$ and $q_i^t = 0$ for all $i \in U_{0\backslash t}$. Then, the comparison universe reduces to U_{0t}, and the test T2 applies, yielding $P^{0,t} = V^{0,t} \geq 1$ under the stipulated setting. Next, suppose substitution occurs only among the persistent items. Given any actual $\{q_i^t; i \in U_{0t}\}$ and the corresponding utility at t, it cannot cost more for the same $\{q_i^t; i \in U_{0t}\}$ at 0 since $p_i^0 \leq p_i^t$ for all $i \in U_{0t}$. It follows that a COLI must be greater or equal to 1. Finally, suppose substitution also involves the items in $U_{0\setminus t}$. Let $\{\tilde{q}_i^0; i \in U_{0t}\}$ be a hypothetical set of persistent units that would have yielded the same utility as the actual $\{q_i^0; i \in U_0\}$. The expenditure of $\{\tilde{q}_i^0; i \in U_{0t}\}$ at 0 cannot be less than the actual expenditure of $\{q_i^0; i \in U_0\}$; whereas the expenditure of $\{\tilde{q}_i^0; i \in U_{0t}\}$ at t cannot be less than that at 0. It follows again that a COLI must be greater or equal to 1.

Remark Under the setting of test T3, there exists a clear downwards trend of prices in the persistent universe. We should have $P^{0,t} \leq 1$ even if the price fall leads to an increase of expenditure, i.e. $V^{0,t} > 1$. Similarly, under the setting of test T4, there exists a clear upwards trend of prices. We should have $P^{0,t} \geq 1$ even if the price increase causes the expenditure to drop, i.e. $V^{0,t} < 1$.

Remark The following tests are somewhat sharper versions of tests T3 and T4. It is stated that $P^{0,t}$ can possibly deviate from 1 in a particular direction depending on whether the item universe is expanding or shrinking, *even if* all the prices of the persistent items remain the same. These are thus clearly the implications of the fact that the item universe is dynamic.

Test t3 If $U_0 \subset U_t$, i.e. $U_{t\setminus 0} \neq \emptyset$, and $p_i^0 = p_i^t$ for all $i \in U_0$, then $P^{0,t} \leq 1$.

Test t4 If $U_t \subset U_0$, i.e. $U_{0\setminus t} \neq \emptyset$, and $p_i^0 = p_i^t$ for all $i \in U_0$, then $P^{0,t} \geq 1$.

Responsiveness test (T5) For $U_0 \neq U_t$, $P^{0,t}$ should *not always* reduce to $f(D_{0t})$, where $D_{0t} = D(U_{0t})$ is based on the reference universe that consists only of the persistent items.

That is, one should not always be able to reduce a COGI to a fixed-basket index, where the basket items are only taken from U_{0t}. This is necessary for any COGI that in principle can be applied to a dynamic item universe. Whereas, since it must allow for substitution that involve items from $U_{t\setminus 0} = U_t \setminus U_{0t}$ and $U_{0\setminus t} = U_0 \setminus U_{0t}$, a COLI cannot be reduced to $f(D_{0t})$.

Result (R1) If $U_0 \neq U_t$, and $p_i^0 = p_i^t$ for all $i \in U_{0t}$, then a responsive index $P^{0,t}$ cannot always be equal to 1, regardless of $D(U_{t\setminus 0})$ and $D(U_{0\setminus t})$.

The result follows from test T1, t3, t4 and T5. According to T1, provided $p_i^0 = p_i^t$ for all $i \in U_{0t}$, the price index over the persistent universe U_{0t} must be 1. Any $P^{0,t}$ that is always equal to 1, regardless of $D(U_{t\setminus 0})$ or $D(U_{0\setminus t})$, is not responsive in this setting. As we will see later, result R1 is helpful for detecting whether an index is completely general in a dynamic universe.

Why not a transitivity test? We have not included some other tests, such as the time reversal test, because they are easily satisfied and it makes little difference in practice even if they are included. However, it is a different matter with transitivity. Loosely speaking, an index is transitive if $P^{0,t} = P^{0,r}P^{r,t}$ for any $r \neq 0, t$, provided all the three indices are *calculated in the same way.* Below we list some questions which, in our opinion, indicate why it is difficult to formulate an explicit transitivity test, especially in a dynamic universe.

(i) Consider a constant item universe over 3 periods, i.e. $U = U_0 = U_r = U_t$ where $0 < r < t$. Suppose $p_i^r < p_i^0 = p_i^t$ for all $i \in U$, i.e. after a general sales period at r all the prices return to the same at t. By test T1, we require $P^{0,t} = 1$, so that transitivity is the case if $P^{0,r} = 1/P^{r,t}$. Suppose the index is time reversible, so that $P^{t,r} = 1/P^{r,t} = P^{0,r}$. Then, the index needs to be invariant whether going from $q(U_0)$ to $q(U_r)$ or from $q(U_t)$ to $q(U_r)$, where $q(U_0) \neq q(U_t)$ in general. This can easily be achieved with a fixed-basket index. However, for a COLI to be invariant in these two situations, the utility difference between the three periods 0, r and t must necessarily be attributed to quantities alone. Is this acceptable according to the theory of utility?

(ii) When the item universe is constant over time, it seems intuitive that transitivity prevents chain drifting. In a dynamic universe, frequent chaining is used to avoid the difficulty that would have arisen when making direct price comparisons between U_0 and U_t. However, in order to check whether chain drifting is the case, one must compare the chained index between 0 and t to the direct index that could have been calculated between 0 and t. Thus, when trying to compare the chained and direct indices explicitly, one cannot avoid running into the same difficulty that has motivated the chaining in the first place. To push the logical difficulty to extreme, suppose $U_0 \cap U_t = \emptyset$, i.e. the item universe is completely renewed. While it is possible to devise a chained index between 0 and t, how can one state the conditions of non-drifting, *unless* one is able to define a theoretically correct direct index between 0 and t when $U_0 \cap U_t = \emptyset$?

(iii) The so-called GEKS index has been adapted for temporal price comparisons (Ivancic et al., 2011). However, the spatial extension is undirected and limited, whilst the temporal extension is directional and unlimited. For an index between 0 and t, it seems counter-intuitive to require $P^{0,r}P^{r,t} = P^{0,t}$, for an arbitrarily chosen period r where $r < 0$ or $r > t$. A practical consequence is that the GEKS index $P^{0,t}$ calculated at t will generally differ to $P^{0,t}$ calculated at $t+1$. It follows that, in reality, the disseminated GEKS indices over time will not be transitive – see Appendix A for details. Moreover, theoretically speaking, even when limited to R_M, it seems somewhat *ad hoc* to base the GEKS on all the *two-step* breakdowns, i.e. $P^{0,r}$ and $P^{r,t}$ for $0 \leq r \leq t$. For a truly transitive index, it should be possible to talk about, say, all the *three-step* breakdowns, i.e. $P^{0,r}$, $P^{r,s}$ and $P^{s,t}$ for $0 \leq r \neq s \leq t$. In other words, is there a unique expression for the multilateral index constructed in the spirit of the GEKS index?

2.3 Modification and extension of Geary-Khamis index

We test the GK index. The results show that it lacks responsiveness in certain situations. We propose a modified GK (MGK) index, as well as an extension of the MGK index. However, as we shall explain, the MGK is still not completely responsive, whereas its extension is responsive but fails other tests instead. An index that is somewhat similar to the GK index is the weighted geometric mean (WGM) index – see e.g. de Haan and Krsinich (2014) and Iklé (1972). The WGM index fails the responsiveness test T5 in the same situations as the GK index. In addition, it fails the fixed-basket T2 in general. Some properties of the WGM index are given in Appendix B.

2.3.1 A modified GK index

Deflating a *constant-value reference-price* quantity index yields the Geary-Khamis (GK) index:

$$P_{GK}^{0,t} = V^{0,t}/Q^{0,t} \tag{1}$$

$$Q^{0,t} = \sum_{i\in U_t} p_i q_i^t / \sum_{i\in U_0} p_i q_i^0 \qquad \text{and} \qquad p_i = \left(\sum_{r\in R_i} \frac{p_i^r}{P^{0,r}} q_i^r \right) / \left(\sum_{r\in R_i} q_i^r \right)$$

(Geary, 1958), where the observed price p_i^r is adjusted to a constant-value price by $P^{0,r}$. Notice that, in general, we allow for $U_0 \neq U_t$, and the GK index (1) is well-defined for a dynamic item universe. Notice also that, in the special case of fixed item universe, where $R_B = \{0, t\}$ and $U = U_0 = U_t$, the formula reveals the well-known relationship:

$$P_L^{0,t} Q_P^{0,t} = V^{0,t} = P_P^{0,t} Q_L^{0,t}$$

$$Q_P^{0,t} = (\sum_{i\in U} p_i^t q_i^t)/(\sum_{i\in U} p_i^t q_i^0) \qquad\qquad Q_L^{0,t} = (\sum_{i\in U} p_i^0 q_i^t)/(\sum_{i\in U} p_i^0 q_i^0)$$

Remark Chessa (2016) refers to (1) as a Quality Adjusted Unit Value (QAUV) index. In concept, the QAUV index is to be calculated for the so-called homogeneous products, which do not directly correspond to the observed items. The construction of artificial homogeneous products can be motivated for both theoretical and practical reasons. Theoretically, substitution should be meaningful among the items that belong to the same homogeneous product. Various hypotheses can be advanced as the motivation for this, including the theory of utility, in light of which the term Quality is unnecessarily limiting, especially when Quality has never been explicitly defined itself. The construction of homogeneous products provides also a means to overcome the difficulty associated with item-matching in practice. We shall discuss such methods in Section 3. However, it is then inappropriate to confuse a practical remedy, which unavoidably entails some bias in general, with the theoretical concept of exchangeable items (in utility, quality, etc.). Recall that in the discussions here in Section 2.3, item-matching is assumed to be unproblematic and ideal. For these reasons we prefer to retain the term GK index for (1).

A special case The index (1) fails the responsiveness test T5 for $R_B = \{0, t\}$ and $q_i^r = \delta_i^r$, where $\delta_i^r = 1$ if $i \in U_r$ and 0 otherwise, such as when the universe consists of rental objects. Let $V_0 = \sum_{U_0} p_i^0 = V_{0\cap t}^0 + V_{0\setminus t}^0$ over $U_0 \cap U_t$ and $U_0 \setminus U_t$, and $V_t = \sum_{U_t} p_i^t = V_{0\cap t}^t + V_{t\setminus 0}^t$. We have

$$P^{0,t} = \frac{V^{0,t}}{Q^{0,t}} = \left(\frac{V_{0\cap t}^t + V_{t\setminus 0}^t}{V_{0\cap t}^0 + V_{0\setminus t}^0} \right) / \left(\frac{\sum_{i\in U_t} p_i q_i^t}{\sum_{i\in U_0} p_i q_i^0} \right) = \left(\frac{V_{0\cap t}^t + V_{t\setminus 0}^t}{V_{0\cap t}^0 + V_{0\setminus t}^0} \right) / \left(\frac{\frac{V_{0\cap t}^0}{2} + \frac{V_{0\cap t}^t}{2P^{0,t}} + \frac{V_{t\setminus 0}^t}{P^{0,t}}}{\frac{V_{0\cap t}^t}{2P^{0,t}} + \frac{V_{0\cap t}^0}{2} + V_{0\setminus t}^0} \right)$$

where $p_i = (\delta_i^0 p_i^0 + \delta_i^t p_i^t / P^{0,t})/(\delta_i^0 + \delta_i^t)$. Successively rearranging the expression yields

$$P^{0,t}(V^0_{0\cap t} + V^0_{0\setminus t})(P^{0,t}V^0_{0\cap t} + V^t_{0\cap t} + 2V^t_{t\setminus 0}) = (V^t_{0\cap t} + V^t_{t\setminus 0})(V^t_{0\cap t} + P^{0,t}V^0_{0\cap t} + 2P^{0,t}V^0_{0\setminus t})$$
$$(P^{0,t}V^0_{0\cap t} - V^t_{0\cap t})(V^t_{0\cap t} + P^{0,t}V^0_{0\cap t} + P^{0,t}V^0_{0\setminus t} + V^t_{t\setminus 0}) = 0$$
$$P^{0,t}V^0_{0\cap t} - V^t_{0\cap t} = 0$$

i.e. $P^{0,t}$ depends only on the persistent item universe $U_0 \cap U_t$ in this situation.

Modification It is the constant-value price adjustment that causes the oddity above. Put

$$P^{0,t}_{MGK} = V^{0,t}/Q^{0,t} \tag{2}$$
$$Q^{0,t}_{RP} = \sum_{i\in U_t} p_i q_i^t / \sum_{i\in U_0} p_i q_i^0 \qquad \text{and} \qquad p_i = (\sum_{r\in R_i} p_i^r q_i^r)/(\sum_{r\in R_i} q_i^r)$$

where $Q^{0,t}_{RP}$ is a ***reference-price (RP)*** quantity index. We refer to the index (2) as the ***modified Geary-Khamis (MGK)*** index. When $R = \{0, t\}$ and $q_i^t = \delta_i^t$, we have

$$P^{0,t} = \left(\frac{V^0_{0\cap t} + V^t_{0\cap t} + 2V^t_{t\setminus 0}}{V^0_{0\cap t} + V^t_{0\cap t} + 2V^0_{0\setminus t}} \cdot V^{0,t} \cdot \frac{V^0_{0\cap t} + V^t_{0\cap t} + 2V^0_{0\setminus t}}{V^0_{0\cap t} + V^t_{0\cap t} + 2V^t_{t\setminus 0}}\right)^{\frac{1}{2}} = \sqrt{V^{0,t}}$$

which neither reduces to the matched universe $U_0 \cap U_t$ nor equals to the value index $V^{0,t}$.

Test results The MGK index (2) does not satisfy the identity test T1 except when $R(0, t) = R_B$:

$$P^{0,t} = \frac{V^{0,t}}{Q^{0,t}_{RP}} = \Big(\frac{\sum_{i\in U} p_i^0 q_i^t}{\sum_{i\in U} p_i^0 q_i^0}\Big) / \Big(\frac{\sum_{j\in U} p_j q_j^t}{\sum_{j\in U} p_j q_j^0}\Big) = 1 \qquad \text{where} \quad p_j = \frac{q_j^0 p_j^0 + q_j^t p_j^t}{q_j^0 + q_j^t} = p_j^0$$

Thus, ***in order for the (M)GK index to satisfy T1, one should adopt a bilateral index, and avoid the use of multilateral indices***. Next, it satisfies the fixed-basket test T2 since $Q^{0,t}_{RP} = 1$ then. Thirdly, it satisfies the upper bound test T3, provided the reference price of a persistent item is such that $p_i^t \leq p_i \leq p_i^0$ – which seems intuitive given R_B, since we have then

$$Q^{0,t}_{RP} = \frac{\sum_{j\in U_0} p_j q_j^t + \sum_{j\in U_{t\setminus 0}} p_j^t q_j^t}{\sum_{j\in U_0} p_j q_j^0} \geq \frac{\sum_{j\in U_0} p_j^t q_j^t + \sum_{j\in U_{t\setminus 0}} p_j^t q_j^t}{\sum_{j\in U_0} p_j^0 q_j^0} = V^{0,t}$$

Similarly, it satisfies the lower bound test T4, provided the reference price of a persistent item is such that $p_i^0 \leq p_i \leq p_i^t$ – which seems intuitive given R_B, since we have then $Q^{0,t}_{RP} \leq V^{0,t}$. However, it fails the responsiveness test T5 in the settings of tests t3 and t4. In the setting of test t3, i.e. $U_0 \subset U_t$ and $p_i^0 = p_i^t$ for all $i \in U_0$, the reference price under R_B must satisfy $p_i = p_i^0 = p_i^t$

for $i \in U_0$ and $p_i = p_i^t$ for $i \in U_{t\backslash 0}$, so that

$$Q_{RP}^{0,t} = \frac{\sum_{i \in U_t} q_i^t p_i}{\sum_{i \in U_0} q_i^0 p_i} = \frac{\sum_{i \in U_t} q_i^t p_i^t}{\sum_{i \in U_0} q_i^0 p_i^t} = V^{0,t} \qquad \Rightarrow \qquad P_{MGK}^{0,t} \equiv 1$$

regardless of $D(U_{t\backslash 0})$, i.e. $P_{MGK}^{0,t}$ is only a function of D_{0t} in this situation. Similarly, in the setting of test t4, i.e. $U_t \subset U_0$ and $p_i^0 = p_i^t$ for all $i \in U_t$, we obtain $Q_{RP}^{0,t} = V^{0,t}$, and $P_{MGK}^{0,t} \equiv 1$ regardless of $D(U_{0\backslash t})$ in this situation.

2.3.2 Reference-quantity-price (RQP) index

To enhance the responsiveness of the MGK index, consider the following. To obtain the MGK, the value index is deflated by a reference-price quantity index. One can equally obtain a quantity index by deflating the value index by a *reference-quantity* price index, which is given by

$$P_{RQ}^{0,t} = \sum_{i \in U_t} q_i p_i^t / \sum_{i \in U_0} q_i p_i^0$$

where q_i is the reference quantity of item i in $U_{R(0,t)}$.

Remark Traditional fixed-basket indices follow provided $R_B = \{0, t\}$ and $U = U_0 = U_t$:

$$\begin{cases} \text{Laspeyres: } P_L^{0,t} = \sum_{i \in U} q_i^0 p_i^t / \sum_{i \in U} q_i^0 p_i^0 & \text{where } q_i = q_i^0 \\ \text{Paasche: } P_P^{0,t} = \sum_{i \in U} q_i^t p_i^t / \sum_{i \in U} q_i^t p_i^0 & \text{where } q_i = q_i^t \\ \text{Marshall-Edgeworth (ME): } P_{ME}^{0,t} = \sum_{i \in U} q_i p_i^t / \sum_{i \in U} q_i p_i^0 & \text{where } q_i = (q_i^0 + q_i^t)/2 \\ \text{Walsh: } P_W^{0,t} = \sum_{i \in U} q_i p_i^t / \sum_{i \in U} q_i p_i^0 & \text{where } q_i = \sqrt{q_i^0 q_i^t} \end{cases}$$

In a dynamic universe, the last two each gives rise to a corresponding reference-quantity index:

$$\begin{cases} \text{RQ-ME: } P_{ME}^{0,t} & \text{if } q_i = \sum_{r \in R_i} q_i^r / |R_i| \\ \text{RQ-Walsh: } P_W^{0,t} & \text{if } q_i = (\prod_{r \in R_i} q_i^r)^{\frac{1}{|R_i|}} \end{cases}$$

where R_i is the set of periods in which item i is present, and $|R_i|$ its cardinality.

Test results Clearly, the index $P_{RQ}^{0,t}$ satisfies the identity test T1. It does *not* satisfy the fixed-basket test T2 except in the case of $R(0,t) = R_B$. It does not necessarily satisfy the upper bound test T3, unless we have $q_i^0 \leq q_i \leq q_i^t$ for all $i \in U_0 \subseteq U_t$, in which case we have

$$P_{RQ}^{0,t} = \frac{\sum_{i \in U_t} q_i p_i^t}{\sum_{i \in U_0} q_i p_i^0} \leq \frac{\sum_{i \in U_t} q_i^t p_i^t}{\sum_{i \in U_0} q_i p_i^0} \leq \frac{\sum_{i \in U_t} q_i p_i^t}{\sum_{i \in U_0} q_i^0 p_i^0} = V^{0,t}$$

In fact, it is straightforward to observe the failure of test t3, where $p_i^0 = p_i^t$ for all $i \in U_0$, but $P_{RQ}^{0,t} > 1$ as long as $q_i^t > 0$ for some $i \in U_{t\backslash 0}$. Similarly, $P_{RQ}^{0,t}$ does not satisfy the lower bound test t4, or T4 in general. However, it remains responsive in the settings of tests t3 and t4, since $P_{RQ}^{0,t}$ does not reduce to $f(D_{0t})$, as long as $q_i^t > 0$ for some $i \in U_{t\backslash 0}$ in the former case and $i \in U_{0\backslash t}$ in the latter, which is necessary in order for the test T5 to be applicable at all.

Extension Consider an extension of the MGK index, which incorporates $P_{RQ}^{0,t}$ as follows:

$$P_{RQP}^{0,t} = \left(P_{RQ}^{0,t}\right)^{1-\alpha}\left(\frac{V^{0,t}}{Q_{RP}^{0,t}}\right)^{\alpha} \tag{3}$$

where α is a constant of choice, for $0 \leq \alpha \leq 1$. We refer to (3) as the ***reference-quantity-price (RQP)*** index, because it can make use of both reference-quantities and reference-prices. It is obtained from deflating the value index $V^{0,t}$ by the weighted geometric mean of two quantity indices $Q_{RP}^{0,t}$ and $V^{0,t}/P_{RQ}^{0,t}$, i.e.

$$\frac{V^{0,t}}{\left(Q_{RP}^{0,t}\right)^{\alpha}\left(V^{0,t}/P_{RQ}^{0,t}\right)^{1-a}} = \left(P_{RQ}^{0,t}\right)^{1-\alpha}\left(\frac{V^{0,t}}{Q_{RP}^{0,t}}\right)^{\alpha} = P_{RQP}^{0,t}$$

As special cases of the RQP index, we obtain the MGK index when $\alpha = 1$, and $P_{RQ}^{0,t}$ when $\alpha = 0$. In particular, at $\alpha = 0.5$, the RQP index can be considered to generalise the Fisher index, defined in the special case of $U = U_0 = U_t$ and $R(0,t) = R_B$, where

$$P_L^{0,t} P_P^{0,t} = P_L^{0,t} V^{0,t}/Q_L^{0,t} = P_P^{0,t} V^{0,t}/Q_P^{0,t}$$

Test results The test results of the RQP index can be deduced from those of $P_{MGK}^{0,t}$ and $P_{RQ}^{0,t}$. In particular, the *bilateral* RQP index satisfies the identity test T1 and the fixed-basket test T2; it does not satisfy the upper bound test T3 (or t3) unless $\alpha = 1$, nor the lower bound test T4 (or t4) unless $\alpha = 1$; it remains responsive in the settings of tests t3 and t4, provided $\alpha < 1$.

2.3.3 Summary of test results

Table 1 provides a summary of the test results above. We do not have an index that satisfies all the 5 tests proposed in this paper. Two observations are worth noting. First, only a bilateral but not multilateral index can satisfy the identity test T1 and the fixed-basket test T2. However, ***we do not therefore conclude that a bilateral index is generally preferable to a multilateral index***, since none is perfect and it is possible to compensate for a shortcoming in one respect with better properties in others. Second, since none of the indices considered above can satisfy the tests t3, t4 and T5 at the same time, ***one will have to compromise between the responsiveness test and the bound tests***, e.g. by using the RQP index (3), although it is unclear at this stage how one can set the constant α in practice.

Table 1: Test results for the MGK, RQ and RQP indices

	Identity	Fixed-basket	Upper-bound	Lower-bound	Responsiveness
$P^{0,t}_{MGK}$	Yes if R_B No if R_M	Yes	Yes	Yes	Not in the setting of t3 or t4
$P^{0,t}_{RQ}$	Yes	Yes if R_B No if R_M	Possibly for T3 No for t3	Possibly for T4 No for t4	Yes
$P^{0,t}_{RQP}$	Yes if R_B No, if R_M	Yes if R_B No if R_M	Possibly for T3 No for t3	Possibly for T4 No for t4	Yes

2.4 Volatility of bilateral index

Despite a bilateral index can satisfy the identity and fixed-basket tests in the special case of $U_0 = U_t$, there is also a need to exam the general case of $U_0 \neq U_t$. In particular, we compare a bilateral index defined for $U_0 \cup U_1$ to its counterpart defined for the persistent item universe $U_{01} = U_0 \cap U_1$, in order to understand whether the index becomes more volatile due to the inclusion of $U_{0\setminus 1} \cup U_{1\setminus 0}$. Notice that, since the reference universe is fixed at $R_B = \{0, 1\}$, we will drop the time denotation and simply write e.g. P_{MGK} instead of $P^{0,1}_{MGK}$ in Section 2.4. Moreover, we denote by P_M generically a bilateral index calculated over the persistent item universe U_{01}.

2.4.1 An artificial example

We start with an extremely simple example, and generalise the results in Section 2.4.2. Let there be an item A in both periods, with quantity and price (q^r_A, p^r_A) for $r = 0, 1$, respectively. Consider the following situations:

(I) $U_0 = \{A\}$, and $U_1 = \{A, B\}$ with (q^1_B, p^1_B), i.e. a new item B is introduced;

(II) $U_0 = \{A, B\}$ with (q^0_B, p^0_B), and $U_1 = \{A\}$, i.e. item B disappears;

(III) $U_0 = \{A, B\}$ with (q^0_B, p^0_B), and $U_1 = \{A, C\}$ with (q^0_C, p^0_C), i.e. item B disappears whilst item C is introduced.

Below we compare the RQ and MGK index to their persistent-item counterpart $P_M = p^1_A/p^0_A$.

Situation (I) Assume any method that yields reference price $p_B = p^1_B$, we have

$$P_{RQ} = \frac{q_A p^1_A + q_B p^1_B}{q_A p^0_A} > P_M = \frac{p^1_A}{p^0_A}$$

$$P_{MGK} = \Big(\frac{q^1_A p^1_A + q^1_B p^1_B}{q^0_A p^0_A}\Big) \cdot \Big(\frac{q^0_A p_A}{q^1_A p_A + q^1_B p_B}\Big) \Rightarrow \begin{cases} P_{MGK} = 1 = P_M & \text{if } p^0_A = p_A = p^1_A \\ P_M < P_{MGK} < \frac{p_A}{p^0_A} < 1 & \text{if } p^0_A > p_A > p^1_A \\ P_M > P_{MGK} > \frac{p_A}{p^0_A} > 1 & \text{if } p^0_A < p_A < p^1_A \end{cases}$$

Thus, P_{MGK} is *less volatile* than its persistent-item counterpart P_M, since P_{MGK} lies between 1 and P_M, no matter how the price of A changes. Meanwhile, P_{RQ} is *always* larger than P_M.

Situation (II) Assume any method that yields $p_B = p_B^0$, we have

$$P_{RQ} = \frac{q_A p_A^1}{q_A p_A^0 + q_B p_B^0} < P_M = \frac{p_A^1}{p_A^0}$$

$$P_{MGK} = \Big(\frac{q_A^1 p_A^1}{q_A^0 p_A^0 + q_B^0 p_B^0}\Big) \cdot \Big(\frac{q_A^0 p_A + q_B^0 p_B}{q_A^1 p_A}\Big) \Rightarrow \begin{cases} P_{MGK} = 1 = P_M & \text{if } p_A^0 = p_A = p_A^1 \\ P_M < P_{MGK} < \frac{p_A^1}{p_A} < 1 & \text{if } p_A^0 > p_A > p_A^1 \\ P_M > P_{MGK} > \frac{p_A^1}{p_A} > 1 & \text{if } p_A^0 < p_A < p_A^1 \end{cases}$$

Again, P_{MGK} is *less volatile* than P_M, whereas P_{RQ} is *always* smaller than P_M.

Remark In light of both situations (I) and (II), P_{RQ} can be more volatile than P_{MGK} or P_M, when the item universe strictly expands or shrinks. This is not surprising because P_{RQ} does not have any inherent adjustment for the different sizes of U_0 and U_t.

Situation (III) Assume any method that yields $p_B = p_B^0$ and $p_C = p_C^1$, we have

$$P_{RQ} = \frac{q_A p_A^1 + q_C p_C^1}{q_A p_A^0 + q_B p_B^0} \Rightarrow \begin{cases} P_{MGK} = 1 = P_M & \text{if } q_B p_B^0 = q_C p_C^1 \\ P_M < P_{RQ} & \text{if } q_B p_B^0 < q_C p_C^1 \\ P_M > P_{RQ} & \text{if } q_B p_B^0 > q_C p_C^1 \end{cases}$$

$$P_{MGK} = \Big(\frac{q_A^1 p_A^1 + q_C^1 p_C^1}{q_A^1 p_A + q_C^1 p_C}\Big) \cdot \Big(\frac{q_A^0 p_A + q_B^0 p_B}{q_A^0 p_A^0 + q_B^0 p_B^0}\Big) \Rightarrow \begin{cases} P_{MGK} = 1 = P_M & \text{if } p_A^0 = p_A = p_A^1 \\ P_M < P_{MGK} < 1 & \text{if } p_A^0 > p_A > p_A^1 \\ P_M > P_{MGK} > 1 & \text{if } p_A^0 < p_A < p_A^1 \end{cases}$$

Once again, P_{MGK} is *less volatile* than P_M as the item universe changes, whereas P_{RQ} behaves differently and depends more on the expenditures of items B and C.

2.4.2 A generalisation

Let U_{01} consist of the persistent items in both periods, and U_{0c} the items that are present at 0 but not at 1, and U_{1c} those that are present at 1 but not at 0. Denote by $(\mathbf{q}_{01}^r, \mathbf{p}_{01}^r)$ the quantities and prices of all the items in U_{01}, at $r = 0, 1$, and by $(\mathbf{q}_{0c}^0, \mathbf{p}_{0c}^0)$ those of all the items in U_{0c}, and by $(\mathbf{q}_{1c}^1, \mathbf{p}_{1c}^1)$ those of all the items in U_{1c}. Let $\mathbf{p}_{01}$ be the reference prices of all the items in U_{01}, and $\mathbf{p}_{0c}$ those in U_{0c}, and $\mathbf{p}_{1c}$ those in U_{1c}. Under the assumption that $\mathbf{p}_{0c} = \mathbf{p}_{0c}^0$ for items in U_{0c}, and $\mathbf{p}_{1c} = \mathbf{p}_{1c}^1$ for items in U_{1c}, we have

$$P_{MGK} = \Big(\frac{V(\mathbf{q}_{01}^1, \mathbf{p}_{01}^1) + V(\mathbf{q}_{1c}^1, \mathbf{p}_{1c}^1)}{V(\mathbf{q}_{01}^1, \mathbf{p}_{01}) + V(\mathbf{q}_{1c}^1, \mathbf{p}_{1c})}\Big)\Big(\frac{V(\mathbf{q}_{01}^0, \mathbf{p}_{01}) + V(\mathbf{q}_{0c}^0, \mathbf{p}_{0c})}{V(\mathbf{q}_{01}^0, \mathbf{p}_{01}^0) + V(\mathbf{q}_{0c}^0, \mathbf{p}_{0c}^0)}\Big)$$
$$= \Big(\frac{V(\mathbf{q}_{01}^1, \mathbf{p}_{01}^1) + V_{1c}}{V(\mathbf{q}_{01}^1, \mathbf{p}_{01}) + V_{1c}}\Big)\Big(\frac{V(\mathbf{q}_{01}^0, \mathbf{p}_{01}) + V_{0c}}{V(\mathbf{q}_{01}^0, \mathbf{p}_{01}^0) + V_{0c}}\Big)$$

where the various expenditures are $V(\mathbf{q}_{01}^r, \mathbf{p}_{01}^r) = \sum_{i \in U_{01}} q_i^r p_i^r$ and $V(\mathbf{q}_{01}^r, \mathbf{p}_{01}) = \sum_{i \in U_{01}} q_i^r p_i$, for $r = 0, 1$, and $V_{0c} = V(\mathbf{q}_{0c}^0, \mathbf{p}_{0c}^0) = \sum_{i \in U_{0c}} q_i^0 p_i^0 = \sum_{i \in U_{0c}} q_i^0 p_i = V(\mathbf{q}_{0c}^0, \mathbf{p}_{0c})$ because $\mathbf{p}_{0c} = \mathbf{p}_{0c}^0$ for U_{0c}, and $V_{1c} = V(\mathbf{q}_{1c}^1, \mathbf{p}_{1c}^1) = \sum_{i \in U_{1c}} q_i^1 p_i^1 = \sum_{i \in U_{1c}} q_i^1 p_i = V(\mathbf{q}_{1c}^1, \mathbf{p}_{1c})$ because $\mathbf{p}_{1c} = \mathbf{p}_{1c}^1$ for U_{1c}. Thus, in a situation where prices *overall* move downwards, such that

$$V(\mathbf{q}_{01}^1, \mathbf{p}_{01}^1) < V(\mathbf{q}_{01}^1, \mathbf{p}_{01}) \qquad \text{and} \qquad V(\mathbf{q}_{01}^0, \mathbf{p}_{01}) < V(\mathbf{q}_{01}^0, \mathbf{p}_{01}^0)$$

we have

$$P_M = \Big(\frac{V(\mathbf{q}_{01}^1, \mathbf{p}_{01}^1)}{V(\mathbf{q}_{01}^1, \mathbf{p}_{01})}\Big)\Big(\frac{V(\mathbf{q}_{01}^0, \mathbf{p}_{01})}{V(\mathbf{q}_{01}^0, \mathbf{p}_{01}^0)}\Big) < P_{MGK} < 1$$

where P_M be the MGK index defined for U_{01}. Similarly, in a situation where prices *overall* move upwards, such that

$$V(\mathbf{q}_{01}^1, \mathbf{p}_{01}^1) > V(\mathbf{q}_{01}^1, \mathbf{p}_{01}) \qquad \text{and} \qquad V(\mathbf{q}_{01}^0, \mathbf{p}_{01}) > V(\mathbf{q}_{01}^0, \mathbf{p}_{01}^0)$$

we have

$$P_M = \Big(\frac{V(\mathbf{q}_{01}^1, \mathbf{p}_{01}^1)}{V(\mathbf{q}_{01}^1, \mathbf{p}_{01})}\Big)\Big(\frac{V(\mathbf{q}_{01}^0, \mathbf{p}_{01})}{V(\mathbf{q}_{01}^0, \mathbf{p}_{01}^0)}\Big) > P_{MGK} > 1$$

In other words, ***one can expect the MGK index defined for*** $U_0 \cup U_1$ ***to be <u>less</u> volatile than if one had restricted the index to the persistent universe only***. Notice that, in the theoretical discussion here, it is assumed that item-matching is unproblematic. The property of the MGK index in the presence of *practical mismatching* will be considered in Section 3.

2.4.3 P_{RQ} vs. P_{MGK} over persistent items

In the above it is seen that the RQ index behaves quite differently to the MGK index when the item universe is dynamic. We now compare the two over the persistent item universe U_{01}. If they are close to each other over U_{01}, then their difference in a dynamic universe is largely due to how they handle the items in $U_{0\setminus 1}$ and $U_{1\setminus 0}$. To facilitate the comparison, we postulate the existence of an unambiguous price trend, where $p_i^1 \equiv \theta p_i^0$ for any $i \in U$, where U is a shorthand of U_{01} for simplicity. Notice that the results below can equally be established if the proportional assumption holds in expectation instead of mathematically, i.e. by assuming $p_i^1 = \theta p_i^0 + \epsilon_i$, with noise ϵ_i having zero expectation $E(\epsilon_i) = 0$ and individual but bounded variance $E(\epsilon_i) = \sigma_i^2$.

Free varying quantities Suppose that (q_i^0, q_i^1) can vary freely. We have

$$P_{RQ} = \frac{\sum_{i\in U} q_i p_i^1}{\sum_{i\in U} q_i p_i^0} = \theta$$

$$P_{MGK} = \frac{(\sum_{i\in U} q_i^1 \theta p_i^0)/(\sum_{i\in U} q_i^0 p_i^0)}{(\sum_{i\in U} q_i^1 p_i)/(\sum_{i\in U} q_i^0 p_i)} = \theta \frac{(\sum_{i\in U} q_i^1 p_i^0)/(\sum_{i\in U} q_i^0 p_i^0)}{(\sum_{i\in U} q_i^1 p_i^0 \kappa_i)/(\sum_{i\in U} q_i^0 p_i^0 \kappa_i)}$$

where the reference price is given as $p_i = p_i^0 \kappa_i$, and

$$\kappa_i = \frac{\theta q_i^1 + q_i^0}{q_i^1 + q_i^0} = 1 + \delta c_i \approx 1 + \delta \bar{c} + \delta (c_i - \bar{c}) \qquad \text{and} \qquad c_i = \frac{q_i^1}{q_i^1 + q_i^0}$$

and $\bar{c}$ is the mean of c_i over U, and $\delta = \theta - 1$ is a small value close to 0. It follows that

$$\frac{\sum_{i\in U} q_i^1 p_i^0 \kappa_i}{\sum_{i\in U} q_i^0 p_i^0 \kappa_i} = \frac{(1+\delta\bar{c}) \sum_{i\in U} q_i^1 p_i^0 + \delta \sum_{i\in U} q_i^1 p_i^0 (c_i - \bar{c})}{(1+\delta\bar{c}) \sum_{i\in U} q_i^0 p_i^0 + \delta \sum_{i\in U} q_i^0 p_i^0 (c_i - \bar{c})} \approx \frac{\sum_{i\in U} q_i^1 p_i^0}{\sum_{i\in U} q_i^0 p_i^0}$$

since $\delta |\sum_{i\in U} q_i^1 p_i^0 (c_i - \bar{c})| < \delta \sum_{i\in U} q_i^1 p_i^0$ in the numerator and $\delta |\sum_{i\in U} q_i^0 p_i^0 (c_i - \bar{c})| < \delta \sum_{i\in U} q_i^0 p_i^0$ in the denominator, i.e. each much smaller than the respective term in front. One may therefore expect the RQ and MGK indices to be close to each other, i.e. $P_{RQ} = \theta \approx P_{MGK}$.

Substitution Substitution may take place as the prices change, in which case (q_i^0, q_i^1) no longer vary freely. To stipulate the consequences, suppose substitution can be described as a constrained optimisation process, where the q_i^1's are changed minimally compared to the q_i^0's, subjected to constant expenditure at both 0 and 1. Put the Lagrangian

$$L = \frac{1}{2} \sum_{i\in U} (q_i^1 - q_i^0)^2 - \lambda (\sum_{i\in U} q_i^1 p_i^1 - \sum_{i\in U} q_i^0 p_i^0)$$

which yields the following solution

$$q_i^1 = q_i^0 + \lambda p_i^1 = q_i^0 + \lambda \theta p_i^0$$

$$\lambda = \frac{\sum_{i\in U} q_i^0 p_i^0 - \sum_{i\in U} q_i^0 p_i^1}{\sum_{i\in U} (p_i^1)^2} = \frac{(1-\theta) \sum_{i\in U} q_i^0 p_i^0}{\theta^2 \sum_{i\in U} d_i} \qquad \text{where} \quad \sqrt{d_i} = p_i^0$$

We have, for $p_i = p_i^0 \kappa_i$,

$$P_{RQ} = \frac{\sum_{i\in U} q_i p_i^1}{\sum_{i\in U} q_i p_i^0} = \theta$$

$$P_{MGK} = \frac{1}{Q_{RP}} = \frac{\sum_{i\in U} q_i^0 p_i}{\sum_{i\in U} q_i^1 p_i} = \frac{\sum_{i\in U} q_i^0 p_i^0 \kappa_i}{\sum_{i\in U} q_i^1 p_i^0 \kappa_i} = \frac{\sum_{i\in U} q_i^0 p_i^0 \kappa_i}{\sum_{i\in U} q_i^0 p_i^0 \kappa_i + \lambda\theta \sum_{i\in U} d_i \kappa_i}$$

$$= \frac{\sum_{i\in U} q_i^0 p_i^0 \kappa_i}{\sum_{i\in U} q_i^0 p_i^0 \kappa_i + \frac{1-\theta}{\theta} \sum_{i\in U} q_i^0 p_i^0 \frac{\sum_{i\in U} d_i \kappa_i}{\sum_{i\in U} d_i}} \approx \frac{1}{1 + \frac{1-\theta}{\theta}} = \theta$$

where we make use of $\sum_{i\in U} q_i^0 p_i^0 \kappa_i \approx (1+\delta\bar{c}) \sum_{i\in U} q_i^0 p_i^0$ and $\sum_{i\in U} d_i \kappa_i \approx (1+\delta\bar{c}) \sum_{i\in U} d_i$ by the same argument as above. Again, the RQ and MGK indices may be close to each other, when substitution takes place as stipulated in the presence of unambiguous trend in prices.

2.4.4 Summary

First, in the presence of clear price trend in a dynamic universe (Section 2.4.2), the bilateral MGK index (2) can be expected to be less volatile than if the index had been restricted to the persistent item universe. It seems that one cannot motivate the use of multilateral indices and more frequent chaining, by arguing that the MGK index could be more volatile than its persistent-universe counterpart in general. Second, one can expect the RQP index (3) to be close to the MGK index over the persistent item universe. The choice between the two is unlikely to be of critical importance, provided the comparison universe is dominated by the persistent items. When this is not the case, the RQP index can become increasingly more responsive and potentially more volatile than the MGK index, as the constant α in (3) moves from 1 to 0.

3 Practical segmented price index

In the above we have considered price index under the matched-model approach. For unit value data, however, item matching for the whole comparison or reference universe can be costly, if the input data does not have precise metadata that can support automatic processing. In some parts of the CPI-universe, such as clothes, automatic item matching has been particularly challenging. Below we develop an approach of ***segmented index***, as a practical means for retaining responsiveness at an affordable cost, which generates a family of indices depending on the construction. To focus on the central idea we describe only *bilateral* indices below. We discuss the interpretation of segmented items, and provide some empirical results based on the Norwegian scanner data. The work is of still going on, and we hope that refinement of the segmentation methodology will lead to generally viable approaches and increase the up-take of unit value data in practice.

3.1 Automatic matched-model indices

Usually some metadata exist with the unit value data, so that ***automatic item matching (AIM)*** is feasible for a subset of U_{0t}, denoted by $U_{0t}^A \subseteq U_{0t}$. For example, one may simply use (outlet, GTIN) as the ***match key***. It is then possible to calculate a matched-model index, *as if* $U_{0t}^A = U_{0t}$. Consider below an automatic MGK index and its RQP extension.

Automatic MGK (AMGK) index Clearly, the value index $V^{0,t}$ is unaffected whether or not $U_{0t}^A = U_{0t}$, but the quantum index is. Let the *automatic* MGK index be given by

$$\widehat{P}_{MGK}^{0,t} = \frac{V^{0,t}}{\widehat{Q}^{0,t}} \quad \text{and} \quad \widehat{Q}^{0,t} = \frac{\sum_{i\in U_{0t}^A} q_i^t p_i + \sum_{i\in U_{t\setminus 0}^A} q_i^t p_i^t}{\sum_{i\in U_{0t}^A} q_i^0 p_i + \sum_{i\in U_{0\setminus t}^A} q_i^0 p_i^0} \tag{4}$$

where $U^A_{t\backslash 0} = U_t \setminus U^A_{0t}$ and $U^A_{0\backslash t} = U_0 \setminus U^A_{0t}$, and we assume that $p_i = p^t_i$ for $i \in U^A_{t\backslash 0}$ and $p_i = p^0_i$ for $i \in U^A_{0\backslash t}$. When $U^A_{0t} = U_{0t}$ happens to be the case, we have $\widehat{P}^{0,t}_{MGK} = P^{0,t}_{MGK}$. However, when none of the persistent items is matched automatically, i.e. $U^A_{0t} = \emptyset$, we would have

$$\widehat{Q}^{0,t} = \frac{\sum_{i\in U_t} q^t_i p^t_i}{\sum_{i\in U_0} q^0_i p^0_i} = V^{0,t} \qquad \Rightarrow \qquad \widehat{P}^{0,t}_{MGK} = 1$$

Thus, ***missing persistent items lead to bias of the AMGK index***. Observe the following.

- If $p^t_i < p_i < p^0_i$ for $i \in U_{0t} \setminus U^A_{0t}$, then $\widehat{Q}^{0,t} < Q^{0,t}_{RP}$ so that $\widehat{P}^{0,t}_{MGK} > P^{0,t}_{MGK}$. Thus, the bias of $\widehat{P}^{0,t}_{MGK}$ can possibly cause it to fail the upper-bound test T4. However, the AMGK still satisfies the test t3, where $p_i = p^0_i = p^t_i$ for all $i \in U_{0t}$, whether or not the item belongs to U^A_{0t}, so that $\widehat{P}^{0,t}_{MGK} = P^{0,t}_{MGK} \leq 1$ in this case.
- If $p^t_i > p_i > p^0_i$ for $i \in U_{0t} \setminus U^A_{0t}$, then $\widehat{Q}^{0,t} > Q^{0,t}$ so that $\widehat{P}^{0,t}_{MGK} < P^{0,t}_{MGK}$. Thus, the bias of $\widehat{P}^{0,t}_{MGK}$ can possibly lead to violation of the lower bound test T4, but not t3.
- The AMGK index (4) satisfies the identity test T1, where $\widehat{Q}^{0,t} = Q^{0,t}_{RP} = V^{0,t}$ since $p_i = p^0_i = p^t_i$ whether or not $i \in U^A_{0t}$. It satisfies the fixed-basket test T2, since $q_i = q^0_i = q^t_i$ whether or not $i \in U^A_{0t}$. It will most likely remain responsive where the MGK index is.

In short, ***the AMGK index can potentially become more volatile than the MGK index due to the missing matched items***, without violating the other tests T1, T2, t3, t4 and T5.

Remark The automatic match key can be broken for a persistent item in two situations:

- The assignment of new GTIN code is associated with ***repricing***, i.e. $p^0_i \neq p^t_i$.
- The assignment of new GTIN code is only associated with ***repackaging***, i.e. $p^0_i = p^t_i$.

The potential bias of the AMGK index due to broken match-keys is only caused by repricing.

Automatic RQP (ARQP) index Let the *automatic* RQ index be given by

$$\widehat{P}^{0,t}_{RQ} = \frac{\sum_{i\in U^A_{0t}} q_i p^t_i + \sum_{i\in U^A_{t\backslash 0}} q^t_i p^t_i}{\sum_{i\in U^A_{0t}} q_i p^t_i + \sum_{i\in U^A_{0\backslash t}} q^0_i p^0_i}$$

If $U^A_{0t} = U_{0t}$ happens to be the case, we have $\widehat{P}^{0,t}_{RQ} = P^{0,t}_{RQ}$. Whereas, if $U^A_{0t} = \emptyset$, we would have $\widehat{P}^{0,t}_{RQ} = V^{0,t}$. Thus, ***missing persistent items lead to bias of*** $\widehat{P}^{0,t}_{RQ}$. The more missing items, the closer $\widehat{P}^{0,t}_{RQ}$ gets to $V^{0,t}$. For the resulting *automatic RQP* index, we have

$$\widehat{P}^{0,t}_{RQP} = \left(\widehat{P}^{0,t}_{RQ}\right)^{1-\alpha} \left(\frac{V^{0,t}}{\widehat{P}^{0,t}_{RQ}}\right)^{\alpha} \quad \xrightarrow{U^A_{0t}\to\emptyset} \quad \left(V^{0,t}\right)^{1-\alpha}$$

3.2 Segmented price indices

The AMGK and ARQP indices above treat $U^A_{t\backslash 0}$ *as if* it were completely new, and $U^A_{0\backslash t}$ as *as if* it had completely disappeared. A different approach is to divide the item universe into ***segments***, and to treat the items in each segment *as if* they were *exchangeable*, so that *each segment at 0 is matched to a corresponding segment at t*. Varying the construction, one can easily obtain different types of segmented price indices from 0 to t. Below we outline two possibilities.

3.2.1 Segment unit-value MGK and RQP indices

Consider a partition of the comparison universe into segments, denoted by $U_0 = \cup_{g=1}^G U_{0g}$ and $U_t = \cup_{g=1}^G U_{tg}$. For each segment, let $\bar{p}_g^t$ be the corresponding unit-value price at t, where

$$\bar{p}_g^t = V_g^t / Q_g^t \qquad \text{and} \qquad V_g^t = \sum_{i\in U_{tg}} q_i^t p_i^t \qquad \text{and} \qquad Q_g^t = \sum_{i\in U_{tg}} q_i^t$$

Similarly for $(\bar{p}^0, Q_g^0, V_g^0)$. Without the need for item-matching at all, a ***segment unit-value (SUV)*** price index can be given by any standard index formula, where *each segment is treated as if it were a genuine item.*

SUV-MGK index Let $\bar{p}_g = (V_g^0 + V_g^t)/(Q_g^0 + Q_g^t)$ be the *reference SUV-price* given $R_B = \{0, 1\}$. The ***SUV-MGK*** index is given by

$$\begin{aligned} \bar{P}^{0,t}_{sMGK} &= V^{0,t}/Q_S^{0,t} \\ Q_S^{0,t} &= \sum_{g=1}^G Q_g^t \bar{p}_g / \sum_{g=1}^G Q_g^0 \bar{p}_g = \sum_{g=1}^G V_g^t(\bar{p}) / \sum_{g=1}^G V_g^t(\bar{p}) = V^t(\bar{p})/V^0(\bar{p}) \end{aligned} \tag{5}$$

where $Q_S^{0,t}$ is the SUV-reference-price quantity index. We have

$$\begin{aligned} \bar{P}^{0,t}_{sMGK} &= \frac{V^{0,t}}{Q_S^{0,t}} = \frac{1}{Q_S^{0,t}} \sum_{g=1}^G \frac{V_g^t}{V^0} \cdot \frac{Q_g^t}{Q_g^0} \cdot \frac{Q_g^0}{Q_g^t} \cdot \frac{V_g^0}{V_g^0} = \sum_{g=1}^G \frac{Q_g^t/Q_g^0}{Q_S^{0,t}} \cdot \frac{V_g^0}{V^0} \cdot \Big(\frac{V_g^t}{V_g^0} \cdot \frac{Q_g^0}{Q_g^t} \Big) \\ &= \sum_g \frac{Q_g^{0,t}}{Q_S^{0,t}} \cdot \frac{V_g^0}{V^0} \cdot \bar{P}_g^{0,t} \end{aligned}$$

where $\bar{P}_g^{0,t} = V_g^{0,t}/Q_g^{0,t}$ is the segment UV index.

SUV-RQP index Let $\bar{q}_g = (Q_g^0 + Q_g^t)/2$ be the *reference segment quantity* given $R_B = \{0,1\}$. The ***SUV-RQP*** index is given by

$$P_{sRQP}^{0,t} = \left(P_{sRQ}^{0,t}\right)^{1-\alpha}\left(\frac{V^{0,t}}{Q_S^{0,t}}\right)^{\alpha}$$

$$P_{sRQ}^{0,t} = \frac{\sum_{g=1}^{G}\bar{q}_g\bar{p}_g^t}{\sum_{g=1}^{G}\bar{q}_g\bar{p}_g^0} = \sum_g \frac{\bar{q}_g\bar{p}_g^0}{V^0(\bar{q})}\cdot\frac{\bar{p}_g^t}{\bar{p}_g^0} = \sum_g \frac{V_g^0(\bar{q})}{V^0(\bar{q})}\bar{P}_g^{0,t}$$

where it seems possible to refer to $P_{sRQ}^{0,t}$ as a segment ME index.

3.2.2 Hybrid SUV-MGK and -RQP indices

In the above, segmentation is applied to the whole item universe. An apparent consequence is that one cannot ensure that a persistent item is always allocated to the same segment over time, no matter how the segments are formed in practice. But one can at least avoid this for the persistent items that can be automatically matched. In other words, one may apply segmentation *only* to $U_{t\backslash 0}^A$ and $U_{0\backslash t}^A$, i.e. put $U_{t\backslash 0}^A = \cup_{g=1}^G U_{t\backslash 0g}^A$ and $U_{0\backslash t}^A = \cup_{g=1}^G U_{0\backslash tg}^A$. The ***hybrid SUV-MGK*** and ***hybrid SUV-RQP*** indices are given below.

Hybrid SUV-MGK index Let $\bar{p}_g^A = (V_{0\backslash tg}^A + V_{t\backslash 0g}^A)/(Q_{0\backslash tg}^A + Q_{t\backslash 0g}^A)$ be the *reference SUV-price* for the g-th segment of $(U_{0\backslash t}^A, U_{t\backslash 0}^A)$. The ***hybrid SUV-MGK*** index is given by

$$\bar{P}_{hsMGK}^{0,t} = V^{0,t}/Q_{HS}^{0,t} \tag{6}$$

$$Q_{HS}^{0,t} = \frac{\sum_{i\in U_{0t}^A} q_i^t p_i + \sum_{g=1}^G Q_{t\backslash 0g}^A \bar{p}_g^A}{\sum_{i\in U_{0t}^A} q_i^0 p_i + \sum_{g=1}^G Q_{0\backslash tg}^A \bar{p}_g^A}$$

where $Q_{HS}^{0,t}$ is the hybrid reference-price quantity index.

Hybrid SUV-RQP index Let $\bar{q}_g^A = (Q_{0\backslash tg}^A + Q_{t\backslash 0g}^A)/2$ be the *reference segment quantity* for the g-th segment of $(U_{0\backslash t}^A, U_{t\backslash 0}^A)$. The ***hybrid SUV-RQP*** index is given by

$$P_{hsRQP}^{0,t} = \left(P_{hsRQ}^{0,t}\right)^{1-\alpha}\left(\frac{V^{0,t}}{Q_{HS}^{0,t}}\right)^{\alpha}$$

$$P_{hsRQ}^{0,t} = \frac{\sum_{i\in U_{0t}^A} q_i p_i^t + \sum_{g=1}^G \bar{q}_g^A \bar{p}_{t\backslash 0g}^A}{\sum_{i\in U_{0t}^A} q_i p_i^0 + \sum_{g=1}^G \bar{q}_g^A \bar{p}_{0\backslash tg}^A}$$

where one may refer to $P_{hsRQ}^{0,t}$ as a hybrid ME index, provided $q_i = (q_i^0 + q_i^t)/2$ for $i \in U_{0t}^A$.

3.3 Interpretation

Necessary properties of exchangeability The following properties of exchangeability seem necessary for motivating segmentation based on the unit-value data available.

1. Exchangeability is a local property. To allow for substitution among different items, they must be ***exchangeable*** in some meaningful sense. For instance, imagine if one starts buying less fruit (and food items) in order to pay for mortgage on house. One would refer to it as saving rather than substitution. In reality, substitution as a reaction to price differences (or changes) can only take place 'locally', i.e. among a limited group of items, instead of 'globally' across the whole range of items. For instance, not only is substitution meaningless between house and food, it may be difficult to insist substitution can be meaningful between an iPhone and a chocolate-bar mobile phone that can only be used to make calls and send text messages.

2. Exchangeability is more fundamental than observable traits. Assume a dynamic universe, where $U_0 \neq U_t$. Unless exchangeability between items is possible, there would be no theoretical justification for bringing into comparison the items in $U_{t\setminus 0}$ and $U_{0\setminus t}$. Moreover, in order to avoid spurious missing item-matching, e.g. when an item is assigned different GTIN codes before and after repackaging, item matching can only be based on exchangeability, but not any tangible or directly observable characteristics. The matched-model approach and the hedonic approach based on price-determining factors are best understood as practical means, either for identifying exchangeable items directly or for achieving exchangeability indirectly.

3. Exchangeability is discrete. What makes the different items exchangeable can be conceptualised, variously in terms of utility, quality, satisfaction, etc. To differentiate exchangeability at all, it seems *necessary and sufficient* to assume the existence of *package-exchangeability* across the item universe, i.e. when goods or services are compared in their available packaging, or otherwise delineated in one way or another as-is, such as a month's subscription of broadband or a litre-pack of milk. Exchangeability is thus discrete, in the sense that it is neither necessary nor helpful to treat whatever underlies exchangeability (e.g. utility) over a continuum.

Assumption of exchangeability as a function of unit-value price Without losing generality, let utility be that which makes items exchangeable. It follows from the necessary properties of exchangeability above that, given any item i, it should be possible to define its utility, denoted by u_i, as a function of $q_i = 1$, i.e. in its available packaging. It becomes then equally possible to express $u_i = f(p_i)$ as a function of its *unit-value price*. We make the following assumption. ***Among a suitably chosen set of items, utility is a discrete, positive function of the unit-value price, which is increasing in the latter in segments.*** A function $u_i = f(p_i)$ is increasing in segments, provided for any $p > 0$, there exists an interval $[p_L, p_U]$ around it, i.e. $p \in [p_L, p_U]$, such that $u(p') < u(p)$ for any $p' < p_L$ and $u(p') > u(p)$ for any $p' > p_U$.

Ideal segmentation Denote by $u_1, .., u_G$ the set of discrete unit-value utilities that exist for the comparison universe (U_0, U_t). From both the COLI and COGI perspectives, an *ideal segmentation* method is such that any two items i and j, where $i \in U_0$ and $j \in U_t$, are assigned to the same segment g, for $g = 1, ..., G$, whenever $f_0(p_i^0) = f_t(p_j^t) = u_g$. Notice that, for bilateral indices, one needs to use p_i^0 at 0 and p_j^t at t, and the functions f_0 and f_t refer to 0 and t, respectively. In the special case of $U_0 = U_t$, correct matching of the persistent items over time can yield an ideal segmentation method. However, the approach is inadequate in principle for a dynamic universe, due to the existence of $U_{t\setminus 0}$ and $U_{0\setminus t}$. It follows that ***other segmentation methods are necessary in a dynamic universe***. One can envisage ideal segmentation as the process of assigning each item in U_0 and U_t, respectively, to the unknown *latent* segment it belongs to. Moreover, via the g-th segment, those items in U_0 for which $u_g = f_0(p_i^0)$ are indirectly "matched" to those items in U_t for which $u_g = f_t(p_i^t)$.

Bias of segmented index The UV index over an ideal segment of items, i.e. $\bar{P}_g^{0,t} = \bar{p}_g^t / \bar{p}_g^0 = V^{0,t}/Q^{0,t}$, is a true price index with respect to whatever that makes the items exchangeable to start with. Nevertheless, additional assumptions are needed in order to motivate a segmented index, which combines all the segment-specific UV indices in one way or another. It follows that there are *two sources of bias* for a segmented index.

- ***Misallocation of segment*** This happens when an item is assigned to a wrong segment.
- ***Misspecification of index formula*** One can envisage a segmented index as a function of the segment UV indices $P_S^{0,t} = f(\bar{P}_1^{0,t}, ..., \bar{P}_G^{0,t})$. Misspecification seems unavoidable.

Insofar as the two effects are unclear, it remains impossible e.g. to determine in theory whether the segmented MGK index (5) is more or less biased than its hybrid counterpart (6).

Remark Excessive volatility of any segment UV price index $\bar{P}_g^{0,t}$ can possibly indicate unequal exchangeability in the constructed segment, i.e. misallocation of segment.

Remark Missing match in practice due to broken match-keys of a persistent item does not necessarily result in bias, unless it leads to misallocation of segment in addition.

An indicator of misallocation error It is a clear indication of misallocation error when a known persistent item is assigned to different segments. Provided automatic matching results in a non-empty set U_{0t}^A, which is of the size N_{0t}^A, let M_{0t}^A be the number of items in U_{0t}^A that are assigned to the *same* segment. An indicator of the accuracy of segmentation can be given by

$$\gamma_S = M_{0t}^A / N_{0t}^A \qquad \text{where} \quad 0 \le \gamma_S \le 1.$$

3.4 Formation

Empirically, to form the segments, one needs to address two questions: (i) how many segments? and (ii) where are the segment boundaries? On the one hand, too few segments (e.g. $G = 1$) is bound to lead to unequal exchangeability within each segment; on the other hand, too many segments is likely to increase the misallocation error. Below we outline briefly two methods, taken from the field of survey sampling. The study of segmentation methods is still going on. Notice that the method is applied for each time point separately. For simplicity we have suppressed the time denotation in the description below.

ANOVA The idea is to minimise the within-segment price variance and maximise the between-segment price variance at the same time. Let the segment boundaries be given by $(\mu_0, \mu_1, ..., \mu_G)$, where $\mu_0 < \bar{p}_1 < \mu_1 < \bar{p}_2 < \cdots < \bar{p}_g < \mu_g < \bar{p}_{g+1} < \cdots < \bar{p}_G < \mu_G$. These should satisfy

$$\frac{(\mu_g - \bar{p}_g)^2 + S_g^2}{S_g} = \frac{(\mu_g - \bar{p}_{g+1})^2 + S_{g+1}^2}{S_{g+1}}$$

where S_g^2 is the empirical variance of all the p_i's in g-th segment. The ANOVA approach is similar to optimal design for stratified simple random sampling. It seems especially intuitive when the observed unit-value prices are naturally clustered, in which case iterations with different G's will be able to identify the natural clusters of price as the segments.

Equal weight share (EWS) Let w_i be the *weight* of item i, and W_g the sum of w_i over the g-th segment, and $W = \sum_{g=1}^{G} W_g$. By the EWS method, the segment boundaries $(\mu_0, \mu_1, ..., \mu_G)$, where $\mu_0 < \bar{p}_1 < \mu_1 < \bar{p}_2 < \cdots < \bar{p}_g < \mu_g < \bar{p}_{g+1} < \cdots < \bar{p}_G < \mu_G$, will now satisfy

$$W_1 = \cdots = W_G = W/G$$

In particular, setting $w_i = p_i q_i$ implies that the segments will have equal expenditure shares, whereas setting $w_i \equiv 1$ implies that the segments will have equal number of items.

Two variants of equal size segmentation (ESS) Both will be illustrated in Section 4. Let p_{is} be the 'normal price' of an item in U_0, e.g. calculated as the average price of the item in the previous 12 months including period 0. Let U_{0g} be the g-th ESS-segment of item universe U_0, which contains N_0/G items and is such that $p_{is} \leq p_{js}$ for any $i \in U_{0g}$, and $j \in U_{0h}$ if $g < h$, and $p_{js} \leq p_{is}$ for any $j \in U_{0h}$ if $h < g$. Consider next the item universe U_t. For any item $i \in U_{0t}^A$, i.e. an automatically matched persistent item, let p_{is} be the same 'normal price' as in period 0. For any item $i \in U_t \setminus U_{01}^A$, i.e. an apparent new item at period t, let $p_{is} = p_i^t$. We have experimented with the following two alternative ways of forming the segments of U_t.

- *(Dynamic) segments*: form the G segments based on $\{p_{is}; i \in U_t\}$, regardless of which segment of U_0 an item in U_{01}^A belongs to. Thus, an item in U_{01}^A can be placed in U_{0g} and

U_{th}, where $g \neq h$, i.e. different segments. The resulting segments are of equal size *both* of U_0 and U_t, although the segment size may differ between the two time periods.

- *Fixed segments*: place each item $i \in U_{01}^A$ in the segment of U_0 it belongs to, so that no automatically matched item changes segment from 0 to t; assign the segment for an item in $U_t \setminus U_{01}^A$ according to the *fixed* segment boundaries of U_{0g}'s. As a result, the segments of U_t may have different sizes.

Metadata It may be possible to form the segments based on the available metadata. For instance, text mining techinuqes may be used in combination with (outlet, GTIN). A simple experiment of this method will be presented in Section 4.

4 Illustrations

We illustrate the following indices using scanner data on food and non-alcoholic beverages from the Norwegian grocery market over a two years period (2014 - 2015):

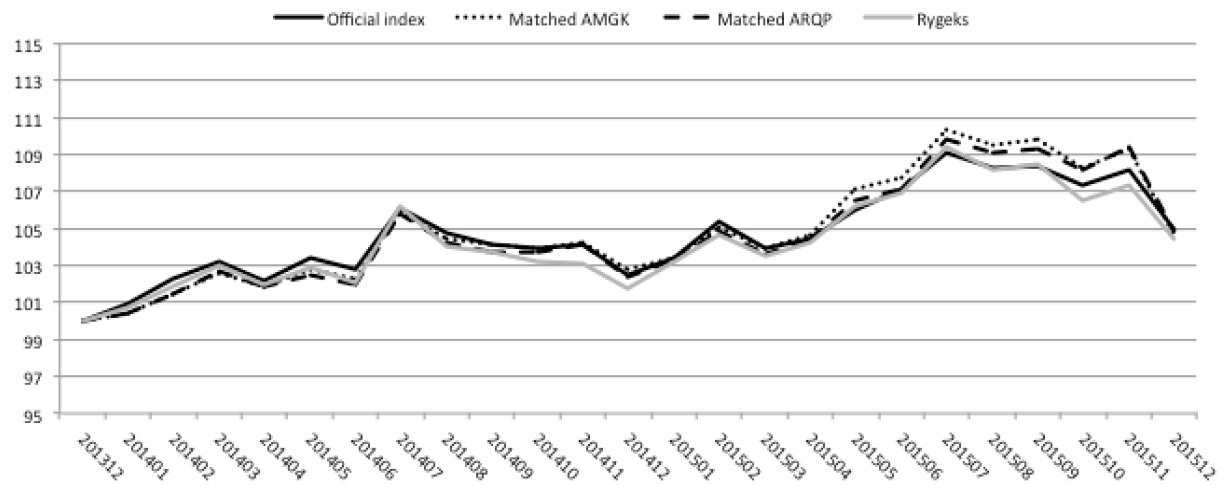
- the official published index, monthly chained with cut-off item universe (Section 1)
- RYGEKS (de Haan and Krsinich, 2014), calculated as a monthly-chained index based on relevant bilateral (weighted) Törnquist indices
- the 6 indices corresponding to crossings (MGK, RQP) $\times$ (automatic, segmented, hybrid), where $\alpha = 0.5$ for the RQP index, and segmentation is either dynamic or fixed
- the bilateral version of the hybrid segmented index, with the reference universe R_B

All the indices are multilateral, i.e based on reference universe $R_M = \{0, 1, ..., t\}$, except the one bilateral hybrid segmented index. In all the cases the base-0 period December is updated once a year; see also relevant discussion of how to build an index over time in Section 5. Below we compare the indices in groups, draw some tentative conclusions and point out several issues raised. Two types of comparisons are of particular interest here: among the different matched-model indices themselves, and against the practical segmented indices beyond the matched-model approach.

4.1 Automatic vs. official index

Consider first the automatic MGK and RQP indices based on the automatically matched items only, where matching based on (outlet, GTIN) is fully automatic, i.e. with minimum resource required. The two indices are close to each other, as suggested by the analysis in Section 2.4.3. Also given is the RYGEKS index based on automatic matching. Meanwhile, the official index is based on a cut-off item universe, and unweighted Jevons index at the elementary level. However, judging from Figure 1, there is little evidence of increasing volatility of the automatic indices at this COICOP2 level. It seems that including expenditure shares is able to sufficiently down-weigh those cut-off items in the current practice. Despite there must exist some missed matches, the automatic indices are less *ad hoc* in construction and give comparable results to the existing method. However, since all these indices are only based on the persistent items, they are not responsive in a dynamic item universe.

Figure 1: Automatic MGK and RQP, RYGEKS vs. Official index. (December 2013 = 100)



Remark It is known that the automatic indices will encounter difficulties in other consumer groups, where one can expect many more missing automatic matches, e.g. due to relauches.

4.2 SUV indices

Figure 2 and 3 show the SUV-MGK and SUV-RQP indices, respectively. Segmentation requires a certain amount of effort in practice, compared to the fully automatic indices above. Nevertheless, ***the approach keeps the demand of resource at a minimum, in order to be responsive to the dynamic item universe***, allowing all the new items and their quantity data to be used as soon as they enter the item universe U_t. By fixed segmentation, one ensures that an automatically matched item is kept in the same segment. This makes use of the match keys available instead of ignoring them, which can be helpful for reducing mismatched price comparisons. Judging from these results dynamic segmentation can yield substantial difference to the matched-model approach, though not necessarily much more volatile. Part of the difference can certainly be attributed to a bias due to departures from ideal segmentation. Compared to the official index, whereas there are some differences in the short-term price movements, the trend of the fixed-segment SUV-indices is clearly consistent with the official indices.

Figure 2: SUV-MGK (dynamic & fixed) vs. Official index. (December 2013 = 100)

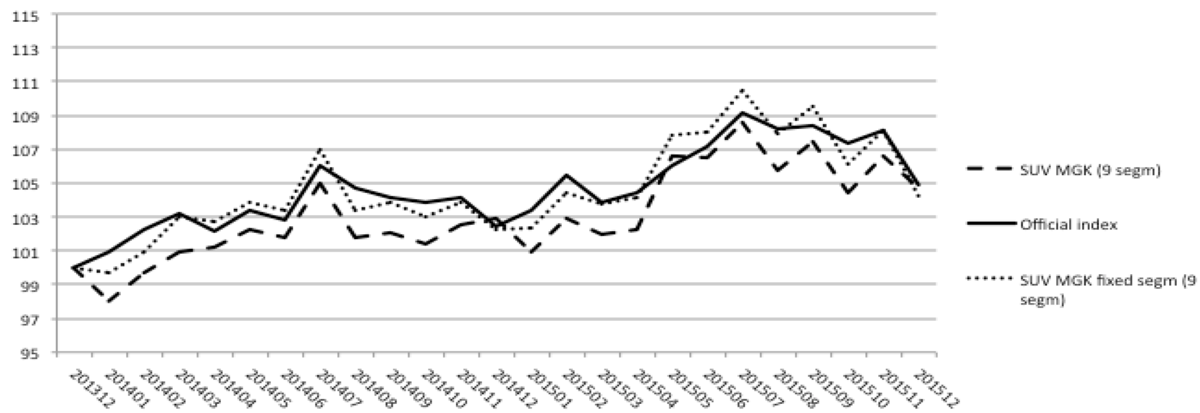
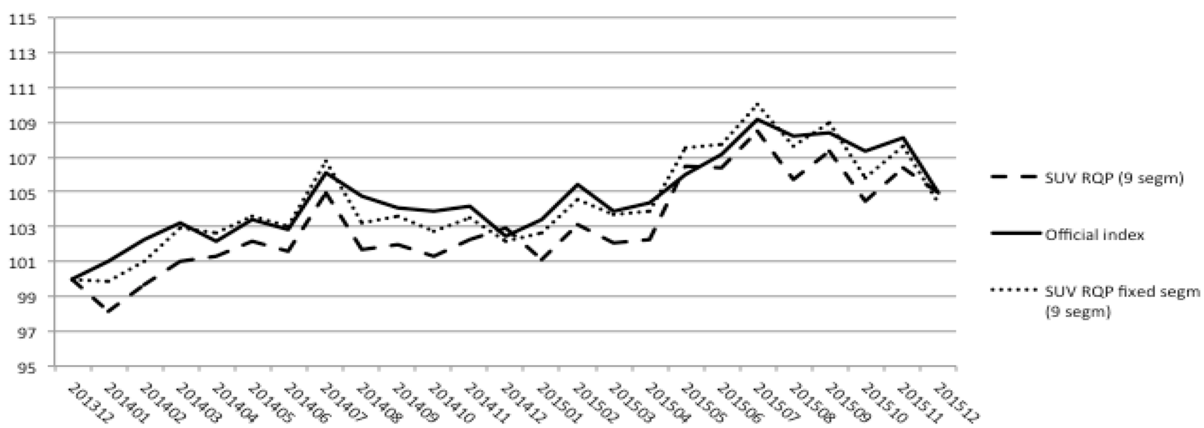


Figure 3: SUV-RQP (dynamic & fixed) vs. Official index. (December 2013 = 100)



Remark There is obviously a connection between segment and "homogenous product group" (Chessa, 2015). However, notice that we have only used 9 segments here for every COICOP6-domain, which is much less than the number of "homogenous product group", as it is currently envisaged and being tested elsewhere. Moreover, the segmentation methods that we have used here do not require any extra effort for collecting and making use of additional metadata. In a way, one may consider these segmentation methods to present an extreme practical means, which neither requires any additional metadata nor the effort to deal with them.

Remark On the one hand, departures from the ideal segmentation can cause bias and potentially increase the volatility of an segmented index. On the other hand, one might also expect an index that is responsive to the dynamic item universe to be somewhat more volatile. However, it is unclear at this stage ***how to assess the noise arising from lack of ideal segmentation against the responsiveness due to incorporation of non-persistent items***.

4.3 Hybrid indices

Figure 4: Hybrid-MGK vs. SUV-MGK and Official index. (December 2013 = 100)

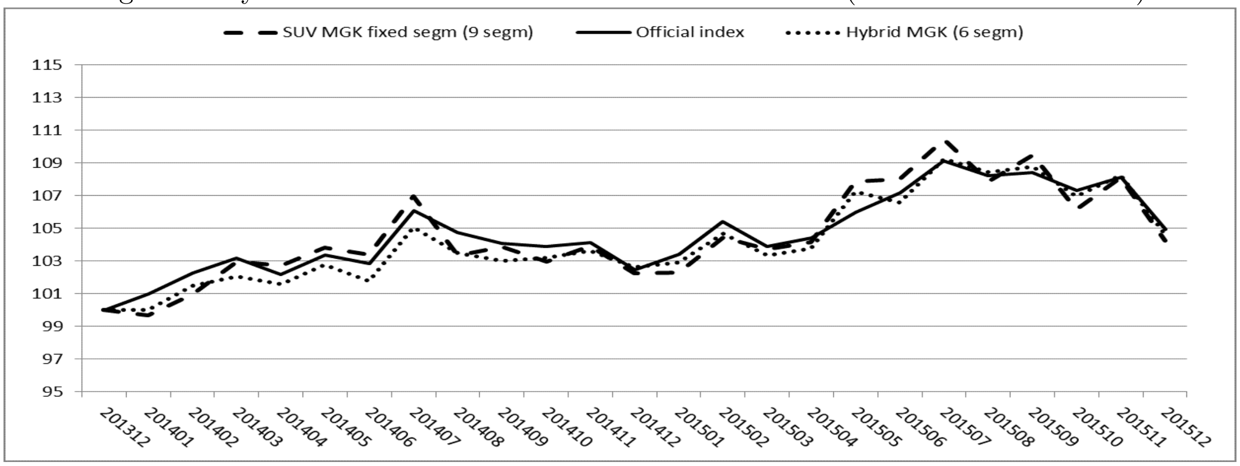


Figure 5: Hybrid-RQP vs. SUV-RQP and Official index. (December 2013 = 100)

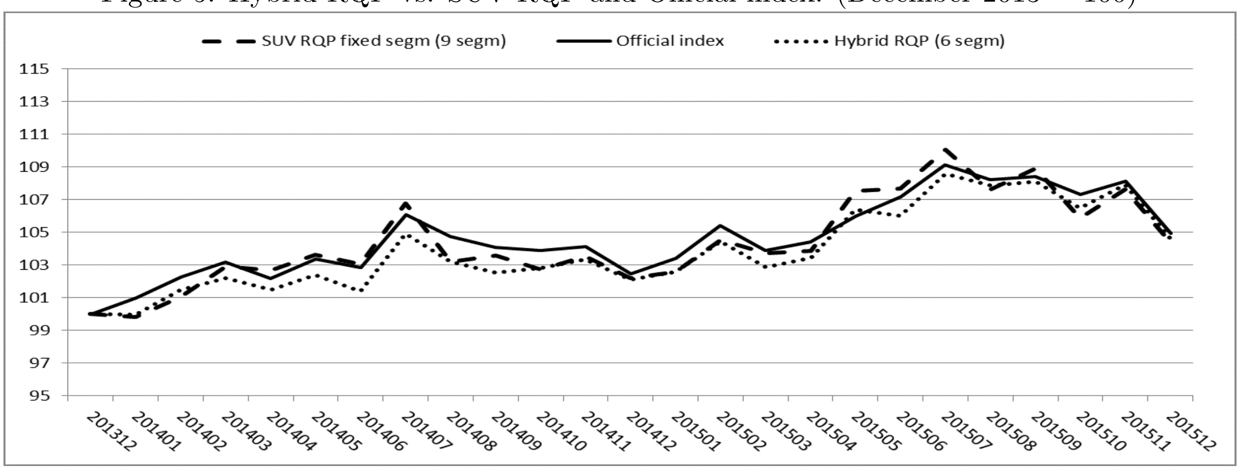


Figure 4 and 5 show the hybrid-MGK and hybrid-RQP indices, respectively. The differences to the SUV counterparts are small. Indeed, the short-term movements of the hybrid indices seem to agree somewhat better with the official index than the SUV-indices. However, there is no reason to expect this generally, as explained below. Both types of index avoid separating an automatically matched item into different segments, but differ in the way the rest items enter the index. In an SUV index with fixed segments, the unmatched items, i.e. in $U_0 \setminus U_{0t}^A$ and $U_t \setminus U_{0t}^A$, are mixed *with* the matched items to form 'large' segments for comparison. In a hybrid index, the unmatched items form 'small' segments *apart from* the matched items – since there are fewer unmatched items here, the hybrid indices are based on 6 instead of 9 segments, to reduce the misallocation of unmatched items. In situations where many of the unmatched items are actually non-persistent items, it is possible that the in-coming and out-going items are not directly comparable. The hybrid index may then be less plausible than the SUV index with fixed segments.

Remark One can manually check an audit item sample from $U_0 \setminus U_{0t}^A$ and $U_t \setminus U_{0t}^A$, to control for the comparability between them, and to understand which type of index is more plausible.

Hybrid approach by weighted average It seems natural in future to investigate another type of hybrid index, given as a suitable weighted average of an automatic and a SUV index.

4.4 Bilateral indices

In Chapter 2 both bilateral and multilateral indices are tested formally. Here, in Figure 6 and 7 we compare bilateral and multilateral hybrid indices. The two types of index differ little in terms of volatility for these data. The MGK indices differ somewhat in trend, where the bilateral index is mostly lower than the multilateral index. The same is not the case with the RQP indices.

Figure 6: Bilateral vs. multilateral hybrid-MGK and Official index. (December 2013 = 100)

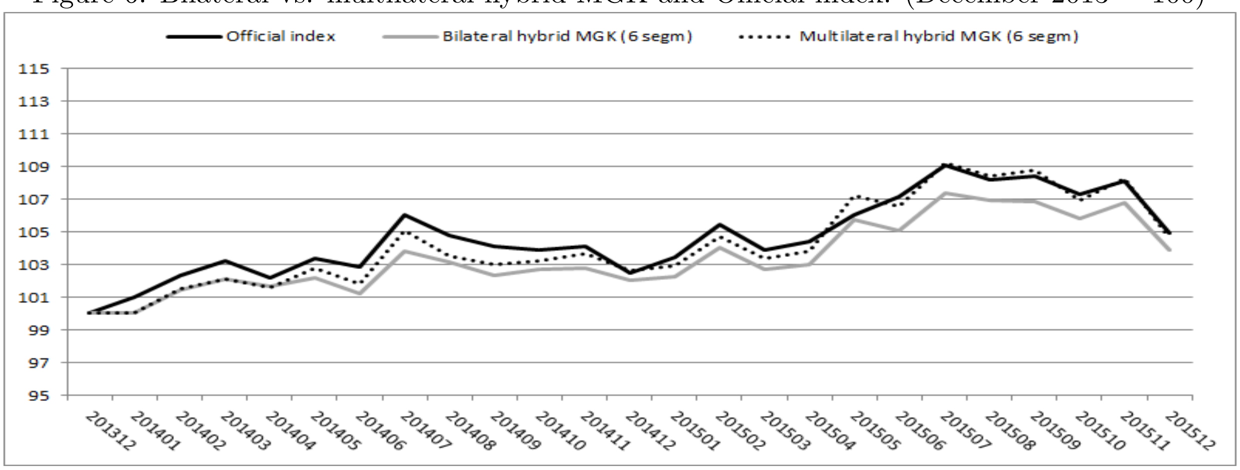
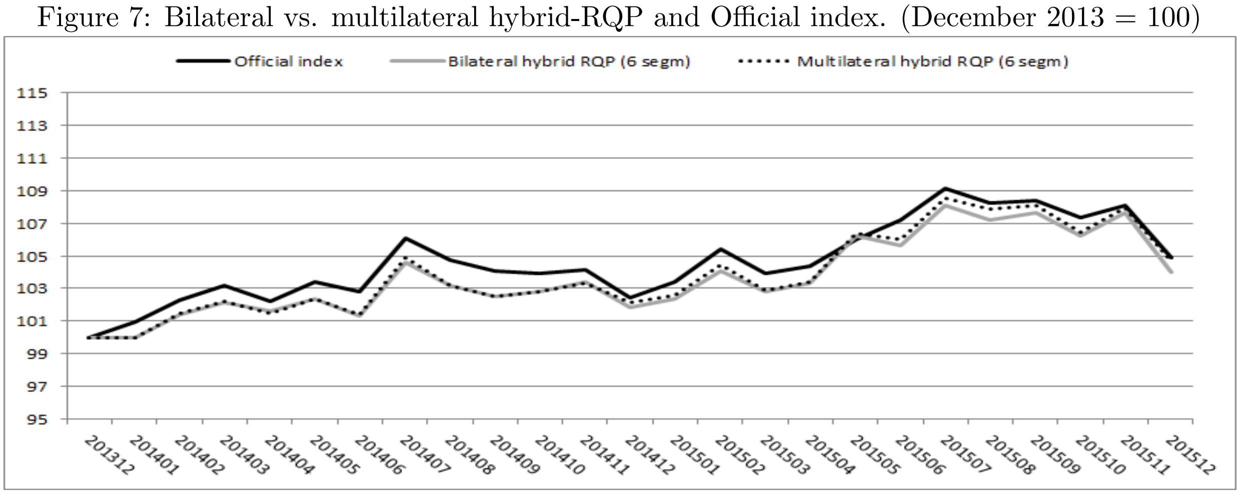


Figure 7: Bilateral vs. multilateral hybrid-RQP and Official index. (December 2013 = 100)

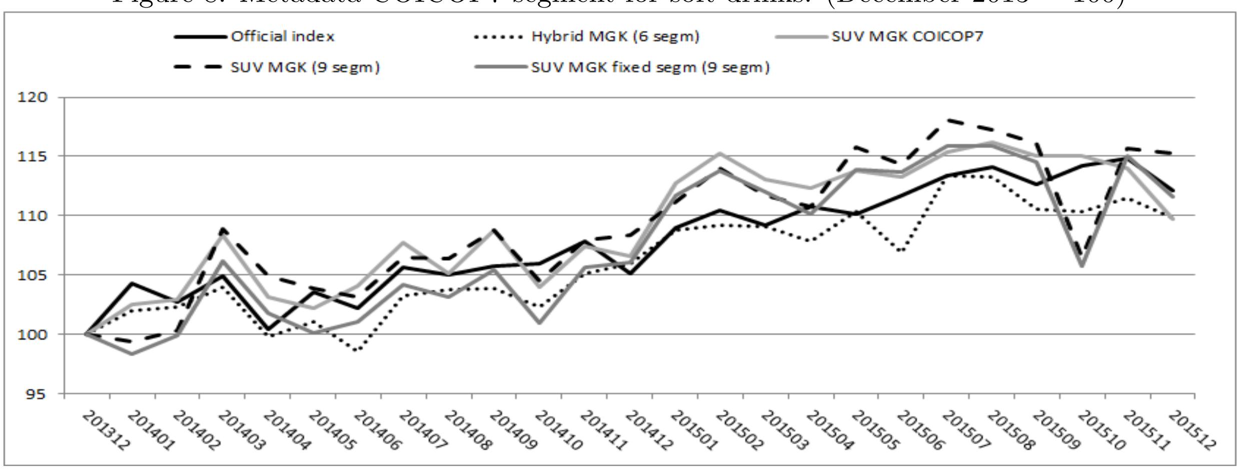


Remark More detailed investigation is also needed to gain a better understanding of bilateral vs. multilateral segmented indices. Take e.g. the fixed segment method, where by construction the segments are 'matched' over time. The two MGK indices differ only in terms of the reference SUV prices used in the quantity index. The bilateral index will be lower than the multilateral counterpart, provided the reference SUV prices calculated over $R_B = \{0, t\}$ are higher than those calculated over $R_M = \{0, 1, ..., t\}$. This is the case, e.g. provided the reference SUV prices follow a convex movement from 0 to t, which can be checked empirically. For the RQP indices to be close to each other at the same time, the bilateral RQ index must be higher than the multilateral RQ index, which is the case e.g. provided the reference segment quantities follow a convex function.

4.5 Metadata segmentation for soft drinks

Segmentation based on available metadata is tested for COICOP6-domain soft drinks. The reason is that data about package and unit volume can be extracted from the metadata available. The segments formed based on such metadata are referred to as COICOP7-segments. The resulting segmented index is given in Figure 8, together with the other segmented and hybrid indices. All the segmented indices are more volatile than the official index at this level, which is as expected. The COICOP7-segmented index is above the official index most of the time in the two-year period, and shows similar movements as the SUV-MGK index based on dynamic segmentation. The hybrid index may seem somewhat closer to the official index than the rest.

Figure 8: Metadata COICOP7-segment for soft drinks. (December 2013 = 100)



Remark It is seen that price-segmentation can produce results similar to metadata-segmentation. To further reduce volatility at a low level, such as COICOP6-domain soft drinks here, one needs to explore various low-cost means for enhancing the automatic item-matching. Appropriate ways of combining the various indices seems also worth considering in future.

4.6 A concluding remark

We find it very challenging to make a choice among the indices in practice, even in situations where they empirically compare to each other in accordance to ones theoretical understanding.

The ultimate reason for this, of course, is that we do not have a theoretically ideal index to target at, hence no ideal benchmark for the empirical results. As a concluding remark we therefore would like to stress the importance of developing ***a set of shared explicit empirical criteria for well-behaving/acceptable indices in practice***. We plan to work on this in future.

5 Building an index series over time

Over time an index series will have to be *indirect*, in the sense that the index between two periods t_1 and t_2 in the series are not necessarily calculated directly for the comparison universe (U_{t_1}, U_{t_2}). In the context of CPI, the index for any two periods more than one year apart is usually an indirect index. There arises thus a question of how to update (or chain) the direct indices over time.

5.1 Some basic constructions

Period-to-period chaining (PPC) index The PPC index is given by

$$P_{PPC}^{0,t} = \prod_{r=0}^{t-1} P^{r,r+1} = P^{0,1}P^{1,2}\cdots P^{t-1,t} \tag{7}$$

where each component $P^{r,r+1}$ is calculated over the reference universe $R(r, r+1) = U_r \cup U_{r+1}$. The reference universe for $P^{0,t}$ is $U_R = U_0 \cup \cdots \cup U_t$. It seems unacceptable to chain over periods beyond $R = \{0, ..., t\}$, which requires switching the direction of chaining, such as $P^{0,2} = P^{0,3}P^{3,2}$, or $P^{0,2} = P^{0,3}P^{3,1}P^{1,2}$. Thus, in practice PPC always refers to successive periods.

Window splicing (WS) index The WS index is given by

$$P_{WS}^{0,t+1} = P^{0,1}P_{U_{R(1,t+1)}}^{1,t+1} \tag{8}$$

where $P_{U_{R(1,t+1)}}^{1,t+1}$ is calculated based on the updated reference universe for $(1, t+1)$. In case t amounts to 12 months, the index series is updated by a year-on-year movement. Repeating the operation at every period yields then $P^{0,t+2} = P^{0,2}P_{U_{R(2,t+2)}}^{2,t+2}$, $P^{0,t+3} = P^{0,3}P_{U_{R(3,t+3)}}^{3,t+3}$, and so on. The reference periods $R(1, t+1)$ can be $\{1, t+1\}$, or a ***rolling window (RW)***: $R(t+1; t) = \{1, 2, ..., t+1\}$, which dates backwards from the current $t+1$, including another t periods.

RW period-to-period (RWPP) index The RWPP index is given by

$$P_{RWPP}^{0,t+1} = P^{0,t}P_{U_{R(t+1;d)}}^{t,t+1} \tag{9}$$

where $P_{U_{R(t+1;d)}}^{t,t+1}$ is calculated based on the updated RW $R(t+1; d)$ at $t+1$, e.g. the most recent 12 months including $t+1$ if $d = 11$. Regardless of the length of RW, the index series is updated by a period-to-period movement. Repeating the same operation at every period yields then $P^{0,t+2} = P^{0,t+1} P^{t+1,t+2}_{U_{R(t+2;d)}}$, $P^{0,t+3} = P^{0,t+2} P^{t+2,t+3}_{U_{R(t+3;d)}}$, and so on.

Cyclic multi-period (CMP) index The CMP index is given by

$$
\begin{aligned}
&P^{0,t+1}_{CMP} = P^{0,t} P^{t,t+1}_{U_{R(t,t+1)}}, \quad P^{0,t+2}_{CMP} = P^{0,t} P^{t,t+2}_{U_{R(t,t+2)}}, \; \cdots, \quad P^{0,t+m}_{CMP} = P^{0,t} P^{t,t+m}_{U_{R(t,t+m)}}, \\
&P^{0,t+m+1}_{CMP} = P^{0,t+m} P^{t+m,t+m+1}_{U_{R(t+m,t+m+1)}}, \quad P^{0,t+m+2}_{CMP} = P^{0,t+m} P^{t+m+1,t+m+2}_{U_{R(t+m,t+m+2)}}, \dots
\end{aligned} \tag{10}
$$

where $P^{r,t}_{U_{R(r,t)}}$ is calculated based on the updated reference universe at t. The scheme over $\{t, ..., t+m\}$ is repeated in cycles and, within each cycle, the length of the movement increases each time till the time of the next cycle. For instance, the cycle can be a year, in which case a movement within a cycle can also be up to a year. This is a common pattern in many existing CPI series.

5.2 Some choices in practice

In each index series, one can identify a **base series** that is preserved at each period t, which has been calculated before, and a **movement index** that updates the index series. Let s denote **base period**, which is the last period of the base series, and is usually the period of base prices for comparison with the current statistical period t. Due to the indirect nature of an index series, we simply denote the index of the base period as $\tilde{P}^s$, without specifying the period at which the index series is fixed at 100%. The questions are then how to update s and how to calculate the movement index $P^{s,t}$, given which the index of period t is given by

$$\tilde{P}^t = \tilde{P}^s P^{s,t} \tag{11}$$

Period-to-period updating Let the base period $s = t - 1$ be updated at every period, and $P^{s,t} = P^{t-1,t}$. Period-to-period index based on $R(t-1,t) = \{t-1,t\}$ could maximise the number of matched items, except when item universes exhibit strong seasonal variation. Using the PPC index to update $P^{t-1,t}$ means the identity test can not be use to prevent drifting, because the comparison universe is never separated from each other by some intermediate periods. Using the RW-PPM index to calculate $P^{t-1,t}$ means e.g. to use data from at least 10 months ago for the two months that are being compared, which does not seem compelling intuitively.

Indirect slicing Let the base period s be updated once a year, say, in December. The WS index gives rise to an indirect index from s to T by slicing, i.e.

$$\tilde{P}^{s,t} = \frac{\tilde{P}^t}{\tilde{P}^s} = \frac{\tilde{P}^{t-12} P^{t-12,t}_{R(t-12,t)}}{\tilde{P}^s} \tag{12}$$

One can e.g. set $R(t-12,t) = \{t-12,t\}$, and use the bilateral MGK or RQP index, which satisfies the identity and fixed-based tests. However, the number of matched items may be reduced over a year. To increase the amount of data for the death and birth items between $t-12$ and t, one may use $R(t;12) = \{t-12, \cdots, t\}$, although it is unclear whether this has any theoretical advantage.

Cyclic updating Let the base period s be updated once a year, say, in December. The CMP index uses (11) directly. For $P^{s,t}$, if one set $R(s,t) = \{s, \cdots, t\}$, then the index of a period closer to s is calculated using less data than one later. If one set $R(s,t) = R(t;12)$, then each movement index is calculated using a rolling yearly window. However, more outdated data is used for a period closer to s than one later. Of course, one may also use the bilateral index directly.

A Multilateral GEKS index

The so-called ***GEKS index*** from 0 to r, for $0 < r \leq t$, over $U_{R_M} = U_0 \cup \cdots \cup U_t$ is given by

$$P_{GEKS}^{0,r} = \left(\prod_{s=0}^{t} P^{0,s} P^{s,r} \right)^{\frac{1}{t+1}} = \left((P^{0,r})^2 \prod_{s \neq 0,r} P^{0,s} P^{s,r} \right)^{\frac{1}{t+1}} \tag{13}$$

(Ivancic et al, 2011). For any $r < t$, it involves indirect comparisons via the periods outside $\{0, ..., r\}$. For example, given $R = \{0,1,2\}$, one would have $P^{0,1} = \big((P^{0,1})^2 P^{0,2} P^{2,1} \big)^{\frac{1}{3}}$ where both $P^{0,2}$ and $P^{2,1}$ are only available at period 2 but not 1. This is unacceptable in practice, as it would require waiting till a later date in order to calculate the price index for the current period. Thus, the GEKS index that can be implemented in practice is *always* the one with $r = t$ in (13). The GEKS can take any of the indices discussed above as the *constituent* index. Now that the reference universe of the GEKS index $P_{GEKS}^{0,t}$ necessarily extends beyond $R(0,t) = R_B$, it generally does not pass any other tests except the responsiveness test.

Remark In the direction of time, the GEKS index is intransitive since, for any $0 < r < t$,

$$\begin{aligned} P_{GEKS}^{0,t} &= \Big((P^{0,t})^2 \prod_{0<s<t} P^{0,s} P^{s,t} \Big)^{\frac{1}{t+1}} = \Big((P^{0,t})^2 P^{0,r} P^{r,t} \prod_{0<s<t; s \neq r} P^{0,s} P^{s,t} \Big)^{\frac{1}{t+1}} \\ \neq P_{GEKS}^{0,r} P_{GEKS}^{r,t} &= \Big((P^{0,r})^2 \prod_{0<s<r} P^{0,s} P^{s,r} \Big)^{\frac{1}{r+1}} \Big((P^{r,t})^2 \prod_{r<s<t} P^{r,s} P^{s,t} \Big)^{\frac{1}{t-r+1}} \end{aligned}$$

For example, given $R = \{0,1,2\}$, one would have $P^{0,2} \neq P^{0,1} P^{1,2}$, where

$$P^{0,2} = \big((P^{0,2})^2 P^{0,1} P^{1,2} \big)^{\frac{1}{3}} \qquad P^{0,1} = \big((P^{0,1})^2 \big)^{\frac{1}{2}} = P^{0,1} \qquad P^{1,2} = \big((P^{1,2})^2 \big)^{\frac{1}{2}} = P^{1,2}$$

B Weighted geometric-means price index

A ***weighted geometric-means (WGM)*** index is given by

$$P_{WGM}^{0,t} = \frac{\prod_{i \in U_t} (p_i^t / p_i)^{w_i^t}}{\prod_{i \in U_0} (p_i^0 / p_i)^{w_i^0}} = \Big(\frac{\prod_{i \in U_t} (p_i^t)^{w_i^t}}{\prod_{i \in U_0} (p_i^0)^{w_i^0}} \Big) / \Big(\frac{\prod_{j \in U_t} p_j^{w_j^t}}{\prod_{j \in U_0} p_j^{w_j^0}} \Big) \tag{14}$$

with the weights $\sum_{i \in U_t} w_i^t = 1$ and $\sum_{i \in U_0} w_i^0 = 1$, and p_j a reference price of $j \in U_R$.

Remark When $R = \{0, t\}$ and $U = U_0 = U_t$, setting $p_j \equiv p$ and $w_j^0 = w_j^t = \frac{1}{2}(q_j^0 p_j^0 / \sum_{i \in U} q_i^t p_i^t + q_j^t p_j^t / \sum_{i \in U} q_i^t p_i^t)$ reduces (14) to the Tönquist index. Setting $p_j \equiv p$ and time-specific weights $w_j^r = q_j^r p_j^r / \sum_{i \in U_r} q_i^r p_i^r$, for $r = 0, t$, reduces (14) to the ratio between two expenditure-share weighted geometric means. Setting

$$p_j = \prod_{r \in T_j} \Big(\frac{p_j^r}{P^{0,r}} \Big)^{\frac{w_j^r}{\sum_{b \in T_j} w_j^b}} \qquad \text{and} \qquad w_j^r = \frac{q_j^r p_j^r}{\sum_{i \in U_r} q_i^r p_i^r}$$

yields the time-product dummy (TPD) index (de Haan and Krsinich, 2014). Iklé (1972) provides another expenditure-share weighted reference price, where

$$p_j = \left(\prod_{r \in T_j} w_j^r \big(\frac{p_j^r}{P^{0,r}} \big)^{-1} \right)^{-1} \qquad \text{and} \qquad w_j^r = \frac{q_j^r p_j^r}{\sum_{i \in U_r} q_i^r p_i^r}$$

Test results The WGM index (14) does not satisfy the identity test, except when $R(0,t) = R_B$, where $p_j = p_j^0 = p_j^t$, since otherwise one can not ensure $p_j = p_j^0 = p_j^t$ in a dynamic universe, just like the MGK index. It does not satisfy the fixed-basket test in general because there is no direct connection to $V^{0,t}$. Under the settings of tests t3 and t4, we have $P_{WGM}^{0,t} = P_{WGM}^{0,t}(D_{0t}) = 1$, again, just like the MGK index. But we can still consider it to be responsive. It does not seem to satisfy the upper bound test T5 or the lower bound test T4 in general, as there is no direct connection to $V^{0,t}$. All in all, the lack of a direct connection to the expenditure ratio $V^{0,t}$ makes it somewhat harder to grasp the WGM index intuitively.

References

[1] Balk, B. M. (2001). *Aggregation Methods in International Comparisons: What Have We Learned?* ERIM Report, Erasmus Research Institute of Management, Erasmus University Rotterdam.

[2] Chessa, A. G. (2016). A new methodology for processing scanner data in the Dutch CPI. *Eurostat review of National Accounts and Macroeconomic Indicators*, **1**, 49-69.

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

[4] Diewert, E. W. (1999). Axiomatic and Economic Approaches to International Comparisons. In *International and Interarea Comparisons of Income, Output, and Prices*, A. Heston and R.E. Lipsey (eds.), Studies in Income and Wealth, Vol. 61, Chicago: University of Chicago Press., 13-87.

[5] Fisher, I. (1922). *The Making of Index Numbers*. Boston: Houghton-Mifflin.

[6] Iklé, D.M. (1972). A New Approach to the Index Number Problem. *Quarterly Journal of Economics*, **86**, 188-211.

[7] Ivancic, L., Fox, K. J. and Diewert, E. W. (2011). Scanner data, time aggregation and the construction of price indexes. *Journal of Econometrics*, **161**, 24-35.