



Can dissimilarity indexes resolve the issue of when to chain price indexes?

Lorraine Ivancic^a, Kevin J. Fox^{a,b,*}

^a Centre for Applied Economic Research, University of New South Wales, Sydney 2052, Australia

^b School of Economics, University of New South Wales, Sydney 2052, Australia

ARTICLE INFO

Article history:

Received 8 September 2010
 Received in revised form
 18 May 2012
 Accepted 13 September 2012
 Available online 19 September 2012

JEL classification:

C43
 E31

Keywords:

Index numbers
 Price indexes
 Chain drift
 Dissimilarity

ABSTRACT

Chaining is used in index number construction to update weights and link new items into an index. However, chained indexes can suffer from, sometimes substantial, drift. The Consumer Price Index Manual (ILO, 2004) recommends the use of dissimilarity indexes to determine when chaining is appropriate. This study provides the first empirical application of dissimilarity indexes in this context. We find that dissimilarity indexes do not appear to be sufficient to resolve the issue of when to chain.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Chaining is used in index number construction to update weights and to link new items into the index. An issue with the use of chained indexes is that they may be prone to drift. The extent of drift appears to be magnified the more price (and quantity) bouncing is captured in the data. A number of authors have shown that with the use of high frequency data (or scanner data) the impact of chain index drift can be quite extreme (Ivancic et al., 2011; Reinsdorf, 1999; Feenstra and Shapiro, 2001). As a result, it is important to know when the use of chained indexes is appropriate.

The Consumer Price Index Manual (ILO, 2004) states that ‘chaining is advisable if the prices and the quantities pertaining to adjacent periods are more similar than the prices and the quantities of more distant periods, since this strategy will lead to a narrowing of the spread between the Paasche and Laspeyres at each link’ (p. 281). The ILO (2004) recommends the use of a dissimilarity index to establish the degree of dissimilarity of prices and quantities in any two periods. From this information a decision can then be made about whether the use of a chained or direct index is appropriate. We apply dissimilarity indexes to a scanner data set to examine how well these indexes work. To

our knowledge this is the first empirical application of dissimilarity indexes in the price index context.

2. Dissimilarity indexes

Measures of dissimilarity can be applied to both the price and quantity vectors. Diewert (2002) showed that there were many different functional forms that a dissimilarity index could potentially take; see also Fox et al. (2004). These indexes can also take the form of either absolute or relative measures of dissimilarity. The difference between absolute and relative dissimilarity indexes, where there are two price vectors, p^1 and p^2 , is described by Diewert (2002) as follows:

‘An absolute index of price dissimilarity regards p^1 and p^2 as being dissimilar if $p^1 \neq p^2$ whereas a relative index of price dissimilarity regards p^1 and p^2 as being dissimilar if $p^1 \neq \lambda p^2$ where $\lambda > 0$ is an arbitrary positive number.’ (p. 2).

See the Appendix for desirable axioms to be satisfied by absolute and relative dissimilarity indexes; as a referee points out, these axioms do not explicitly include minimizing the frequency of mistakes in determining when to chain. Based on an axiomatic approach to index choice, Diewert (2002) ‘tentatively’ recommended the use of the weighted asymptotically linear index of relative dissimilarity for prices and the weighted asymptotically linear index of absolute dissimilarity for quantities. These two dissimilarity indexes are used in our empirical application.

* Corresponding author. Tel.: +61 2 9385 3320; fax: +61 2 9313 6337.
 E-mail address: K.Fox@unsw.edu.au (K.J. Fox).

The weighted asymptotically linear index of relative dissimilarity for prices, D_{PAL} , is defined as:

$$D_{PAL} = \sum_{i=1}^I \left(\frac{1}{2} \right) (s_{i1} + s_{it}) \left[\left(\frac{p_{it}}{p_{i1}P(p_1, p_t, q_1, q_t)} \right) + \left(\frac{p_{i1}P(p_1, p_t, q_1, q_t)}{p_{it}} \right) - 2 \right], \quad (1)$$

where $P(p_1, p_t, q_1, q_t)$ is any superlative index number formula (Diewert, 1976), $p_t = (p_{1t}, \dots, p_{it})$ is a vector of prices for item $i = 1, \dots, n$ in period t , and s_{it} = the expenditure share of item i in period t .

Eq. (1) captures the extent to which the price change (between periods 1 and t) for an individual item, i , differs from the overall measure of price change (which is measured here by a superlative index). For example, if the estimated price change for item i is the same as the overall rate of price change then the amount of ‘dissimilarity’ for item i captured by the relative dissimilarity index will be zero. The dissimilarity indexes were calculated using the Fisher index as our superlative index of choice. Two other superlative price indexes (Walsh and Törnqvist) were also used to calculate the dissimilarity indexes, but as the use of different superlative indexes had little impact on the estimates of dissimilarity (and no impact on the conclusions reached), results presented in the next section are based on the Fisher index only.

The weighted asymptotically linear index of absolute dissimilarity for quantities, D_{QAL} , is defined as:

$$D_{QAL} = \sum_{i=1}^N \left(\frac{1}{2} \right) (s_{i1} + s_{it}) \left[\left(\frac{q_{it}}{q_{i1}} \right) + \left(\frac{q_{i1}}{q_{it}} \right) - 2 \right], \quad (2)$$

where $q_t = (q_{1t}, \dots, q_{it})$ is a vector of quantities for item $i = 1, \dots, n$ in period t , and s_{it} = the expenditure share of item i in period t .

The absolute dissimilarity quantity index captures the extent to which the quantities purchased of an item, i , vary between period 1 to t . If the same quantities of item i are purchased in period 1 and period t then the amount of dissimilarity captured by the absolute dissimilarity index is zero.

To calculate the direct dissimilarity indexes Eqs. (1) and (2) were applied exactly as specified. To calculate the chained dissimilarity, dissimilarity indexes were calculated between each of the links in the chain. From these, the average dissimilarity across all links was calculated. So to compare the dissimilarity between the chained and direct indexes we in fact compare the dissimilarity of the direct indexes with the average dissimilarity between the links in the chained indexes.

The criterion used to determine when chaining is appropriate was defined as follows:

1. If both the chained absolute quantity index and relative price dissimilarity index are found to be less than their direct counterparts then chaining is recommended.
2. If both the direct absolute quantity and relative price dissimilarity indexes are found to be less than their chained counterparts then chaining is not recommended.
3. If the direct dissimilarity index is less than the chained dissimilarity index on only one dimension (either price or quantity) then there is no clear evidence about whether chaining is appropriate.

Dissimilarity indexes were calculated over a one year time period using, in turn, weekly, and monthly time aggregation over prices and quantities. Items were, in turn, treated as different items if they were not located in the same store (i.e. no item aggregation over stores) or treated as the same good no matter which store they were found in (i.e. item aggregation over stores).

3. Data

We use an Australian scanner data set containing 65 weeks of data, collected between February 1997 and April 1998. The data set contains information on 110 stores which belong to four supermarket chains located in one of the major capital cities in Australia. These stores accounted for over 80% of grocery sales in this city during this period (Jain and Abello, 2001). The data set includes information on 19 supermarket item categories. The item categories and number of observations available for each item category are as follows: biscuits (2,452,797), bread (752,884) butter (225,789), cereal, (1,147,737) coffee (514,945), detergent (458,712), frozen peas (544,050), honey (235,649), jams (615,948), juices (2,639,642), margarine (312,558), oil (483,146), pasta (1,065,204), pet food (2,589,135), soft drinks (2,140,587), spreads (283,676), sugar (254,453), tin tomatoes (246,187) and toilet paper (438,525).

Information on each item includes the average weekly price paid for each item in each store in each week, the total quantity of that item sold in each store in each week, a short product description (including information on brand name, product type, flavor and weight), a unique numeric identifier for each item (that allows for the exact matching of items over time) and information on which store an item was sold in.

4. Dissimilarity index results

The ILO (2004) states that chaining is appropriate when ‘the prices and quantities pertaining to adjacent periods are more similar than the prices and quantities of more distant periods’ (p. 281). For the direct indexes, the basket of goods is held fixed, but for the chained indexes we use a “flexible” basket which allows goods to exit and enter over time. This is closer to actual statistical agency practice when chaining. Results from using a fixed basket for the chained indexes are qualitatively similar, and are excluded here for brevity.

Results are presented in Tables 1 and 2. We find that when the chained dissimilarity indexes are compared with their direct counterparts there are very few circumstances – in total only 9 out of 76 – where both the direct price and quantity dissimilarity indexes are less than the chained price and quantity dissimilarity indexes. Chaining was found to be appropriate in the majority of cases – 47 out of 76 cases. In the remaining 20 cases there was no clear evidence on the issue of chaining. It is somewhat reassuring to find that for a number of item categories where index number estimates show huge amounts of drift (i.e. margarine, soft drinks and toilet paper with weekly time aggregation and no item aggregation over stores) the dissimilarity index results do not recommend the use of chaining. However, there are a number of cases where the dissimilarity indexes indicate that chaining is appropriate but the relevant index number estimates suggest that chaining is not reasonable. For example, for the item category ‘pasta’ (with no item aggregation over stores and chaining at a weekly frequency) the Laspeyres and Fisher indexes are $(539.63 - 100) = 439.63\%$ and $(81.79 - 100) = -18.21\%$ for the sample period respectively. The corresponding direct indexes are $(104.01 - 100) = 4.01\%$ and $(101.56 - 100) = 1.56\%$, which appear to be much more reasonable. Results such as these give cause for concern about the use of dissimilarity indexes to determine when to chain.

Download English Version:

<https://daneshyari.com/en/article/5060325>

Download Persian Version:

<https://daneshyari.com/article/5060325>

[Daneshyari.com](https://daneshyari.com)