Submission for the 2012 IAOS Prize for Young Statisticians

Elementary Aggregate Indices and Lower Level Substitution Bias

Mr. Joseph Winton (joseph.winton@ons.gov.uk),

Dr. Robert O'Neill (robert.o'neill@ons.gov.uk) and

Mr. Duncan Elliott (duncan.elliott@ons.gov.uk)

Office for National Statistics,
Cardiff Road,
Newport NP10 8XG,
United Kingdom

24 February 2012

Elementary Aggregate Indices and Lower Level Substitution Bias

Mr. Joseph Winton (joseph.winton@ons.gov.uk),

Dr. Robert O'Neill (robert.o'neill@ons.gov.uk) and

Mr. Duncan Elliott (duncan.elliott@ons.gov.uk)

Office for National Statistics, Cardiff Road, Newport NP10 8XG, United Kingdom[1]

Abstract

Elementary aggregates (EAs) represent the lowest level at which price indices are constructed and are the first step in calculating an index of the general level of prices. There is currently debate regarding the appropriate index formula to employ in elementary aggregates. The choice between two of these, the Jevons (a geometric mean of price ratios) and the Carli (an arithmetic mean of price ratios) indices is often characterised as being determined by whether or not consumers substitute between products as relative prices change. Under certain assumptions a Jevons may be appropriate where there is substitution behaviour whilst a Carli may be appropriate where there is none.

We estimate the elasticity of substitution using econometric and algebraic techniques suggested elsewhere in the literature and also propose an extension to the algebraic technique. We employ sub-sampling techniques to overcome obstacles faced when applying these methods to panel data on alcohol consumption. We find that estimates of substitution behaviour are insufficient for informing the choice index formula at the EA level, which in part may be due to the presence of demand side effects that are not accounted for.

[1] The analysis and views presented in this paper are those of the authors and do not necessarily represent those of the Office for National Statistics.

1. Introduction

There is considerable debate in the index numbers literature on the choice of method for aggregating prices at the Elementary Aggregate (EA) level[2]. This paper contributes to that debate with an empirical study of the nature of consumer substitution behaviour and the implications for EAs.

The UK publishes a Retail Price Index (RPI) and a Consumer Price Index (CPI)[3] for which methods used for EAs often differ. This contributes to differences in the overall rate of inflation between the two measures, described as the formula effect. As there is no consensus on the appropriate method, this is a critical area of research, particularly as the impact of the difference is highly sensitive due to the potential economic impact of consumer prices.

EAs are the first stage of index construction. Many countries, including the UK, do not collect sufficiently detailed quantity or expenditure information to calculate one of the theoretically ideal index number formulas (ILO, 2004: 357). In the absence of detailed expenditure data there are a number of alternative methods for estimating EAs, but no consensus on an optimal approach. We use detailed historical price and expenditure data for alcohol to evaluate alternative EA formulae against an ideal index. We do this by exploring the extent to which consumers' substitution of goods in response to relative price changes can be estimated and used to inform the decision for EAs. In doing this we assume a Constant Elasticity of Substitution (CES), estimated using both econometric and algebraic approaches.

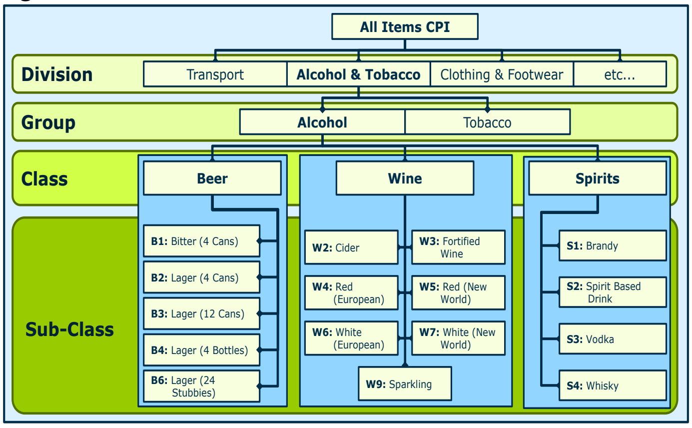
Section 2 discusses the data and methods used. Section 3 presents results, the implications of which are discussed in section 4, and section 5 concludes with areas for further research.

2. Methodology

2.1 Data

This study uses household-based purchase scanner data provided by TNS[4] and contains detailed descriptions of all commodities bought by a panel of approximately 25,000 members in one week of each month over a 34 month period: January 2003 – October 2005.

Figure 1 - Section of COICOP Structure 2005



Commodities were stratified according to the 2005 CPI Basket of Goods (Roe, 2005) using the Classification of Individual Consumption by Purpose (COICOP). This investigation focuses on the COICOP group for alcohol as this group contains a relatively small number of sub-classes and means that the impact of different EA formulae within sub-classes can be assessed at a higher level. Monthly aggregation of prices and quantities has been chosen for this investigation[5].

[2] See Leyland (2011), Roe and Fenwick (2003) and ONS (2011).

[3] The CPI is the European Harmonized Index of Consumer Prices for the UK and the UK government's preferred measure of inflation.

[4] Data supplied by TNS UK Limited. The use of TNS UK Ltd data in this work does not imply the endorsement of TNS UK Ltd. in relation to the interpretation or analysis of the data. All errors and omissions remain the responsibility of the authors.

[5] Aggregation by region and type of shop, as well as over weeks and quarters was also investigated but no significant evidence for differences in price and quantity behaviour between these groups was found and this has not been pursued further.

As the data contains the same products purchased at different prices within the unique time periods of interest, we adopt the approach described in the ILO manual (ILO, 2004: 356) of a narrowly defined unit value as our observed price in a given time period[6]. As the data is based on a sample, we use appropriate population based estimates of quantity and expenditure to derive monthly unit prices and expenditure shares for each product, using weights supplied in the TNS dataset.

2.2 Estimation of the Constant Elasticity of Substitution

The following section describes the methods employed in this paper to estimate the elasticity of substitution using low level data. All approaches described here are based on the constant elasticity of substitution (CES) utility function:

$$U(q) = \left(\sum_{m=1}^{M} b_m^{1/\sigma} \left(q_{mt} \right)^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)} \quad [1]$$

Where $q = (q_{1t}\ ,...,\ q_{Mt})'$ is the vector of quantities of the M goods being consumed, b_m is a taste parameter and σ is the CES. CES measures the willingness of the consumer to substitute as relative prices of goods change. A value of $\sigma = 0$ indicates that the consumer does not substitute in response to price changes, a value of $\sigma = 1$, indicates that consumers substitute to maintain a constant expenditure and values of $\sigma > 1$ indicate a higher degree of substitution behaviour.

The majority of EAs in the UK CPI are estimated with a geometric mean of price relatives (a Jevons index), which can be shown to compensate for unit elasticity of substitution (Silver and Heravi, 2006: 5). The remaining EAs in the UK CPI use ratio of averages (Dutot index) where it is thought that substitution is less likely. The geometric mean is a form of the generalised mean with elasticity set to one and the arithmetic mean (Carli index, used in many RPI EAs) is one with the elasticity set to zero (Balk, 2002: 17).

If σ can be estimated then it may be possible to either assess the validity of the current assumptions, or to improve the calculation of average price change at the EA level. In this section algebraic and econometric approaches to estimating σ are outlined, both of which are used in our empirical investigation.

Algebraic Approach

The algebraic approach described in Balk (1999) involves equating two forms of a CES index formula as in equation [2]. The left hand side term is the base period expenditure share weighted Lloyd-Moulton index, the right hand side is a CES index with current period expenditure share weights. In the case in which consumer behaviour is perfectly described by [1] then [2] holds.

$$\left[\sum_{m=1}^{M} s_m^{t-h} \left(\frac{p_m^t}{p_m^{t-h}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} = \left[\sum_{m=1}^{M} s_m^{t} \left(\frac{p_m^t}{p_m^{t-h}} \right)^{-(1-\sigma)} \right]^{-\left(\frac{1}{1-\sigma}\right)} \quad [2]$$

Where:

p^t_m is the price of commodity *m* at time *t*; and

s^t_m is the expenditure weight of commodity *m* at time *t* calculated over all commodities *m = (1,..,M)* in both periods, *t* and *t-h.*

h is the lag. In Balk (1999) $h = 1$

M is the number of products in the subset used to estimate σ

[6] Unit value is estimated as the total amount spent divided by the total quantity purchased in that period.

We use estimated unit prices and expenditure shares and identify the value of σ which brings about the equality described in [2]. We produce two estimates via the algebraic approach. The first follows the method described by Balk (1999) i.e. $h = 1$ – hereafter the consecutive period approach – and the second uses all possible values of h, (*i.e.* $h = 1, 2, \ldots T\text{-}1$) – hereafter the multiple lag algebraic approach.

In order to provide measures of accuracy for the estimates, we use a sub-sampling approach in estimating σ. We select 50 subsamples of products for each of the sub-classes, with $M = N \times f$, where N is the number of products available at all time points in our dataset and f is a sampling fraction, which for the purposes of this study equals 0.6[7]. Duplicate samples were removed in order to prevent bias in our measures of accuracy. Hence in practice there were n_{SC} ($\leq$50) samples for each sub-class (SC)[8].

We then estimate σ using both the consecutive period and multiple lag algebraic approaches in each sub sample. As a result we end up with $n_{SC} \times (T\text{-}1)$ estimates of σ for the consecutive period approach and $n_{SC} \times (T\text{-}1)!/(2(T\text{-}3)!)$ for the multiple lag approach. Whilst there is a lot of volatility in the estimates (within and across subsamples), there is no evidence of an upward or downward trend over time. There was sporadic evidence of variation in the average level of σ by h. Our within sample estimate of σ is simply the median, chosen because there were some extremes among the estimates[9]. Below, the weighted means[10] of these estimates are presented along with the standard deviations as a guide to the accuracy of the estimates. The consecutive period results are then used in constructing Lloyd-Moulton indices.

Econometric Approach

The econometric approach to estimating σ is also applied in Ivancic *et al*. (2009). This approach uses Seemingly Unrelated Regression (SUR) to estimate the parameters of the model [3] which is derived from the CES utility function, [1][11].

$$\ln\left(\frac{s_m^t}{s_\bullet^t}\right) = \gamma_m + r\ln\left(\frac{p_n^t}{p_\bullet^t}\right) + \varepsilon_m^t \qquad m = 1,\ldots,M;\ t = 0,\ldots,T \qquad \textbf{[3]}$$

Where:

$s^t_\bullet$ is the geometric mean of expenditure shares over the M goods at time t.
$p^t_\bullet$ is the geometric mean of prices over the M goods at time t.

Hence the model is a system of equations with $M+1$ parameters. The parameter labelled r in [3] is equal to $1\text{-}\ \sigma$, hence estimates of σ can easily be obtained from the model.

Based on the CES utility function there is a necessary restriction that $\sum_{m=1}^{M}\gamma_m = 0$ and it is suggested in Ivancic *et al*. (2009) that one of the products be omitted from the system of equations and the constant from that equation can be assumed to be a balancing figure which fulfils the constraint.

---

7 f=0.6 was subjectively chosen to ensure n_{SC} is not too small given some small values of N for some sub-classes.

8 When selecting 50 subsamples with f=0.6 and with N<8 there will with certainty be replicates.

9 Ivancic *et al.* (2009) apply the algebraic approach of Balk (1999). However, it is difficult to compare the volatility of our results to that found their results as our data has 33 bilateral periods compared to 4 in their results.

10 The weights for the consecutive period approach are equal whilst the multiple period approach weights (w_h) are $w_h = \frac{T-h-1}{\sum_k T-h-1}$. In the multiple period approach, within sample estimates are made for each lag (h) and therefore over σ_h . The weights above are used in the estimation of the mean and as the estimates by h are clearly not independent so the estimated standard errors include the appropriately weighted covariances between estimates.

11 A full derivation of equation [3], beginning with the utility function described in [1] can be found in Ivancic *et al*. (2009).

There are problems in implementing this approach due to the characteristics of our data set. It is known from the literature on SUR (see for example Griffiths *et al*., 2001) that there is a sample size requirement for SUR models that $T \geq max(M, kmax)$ where M is the number of equations and $kmax$ is the maximum number of explanatory variables in any of the equations ($kmax$ = 2 in [3]). Hence in our model we encounter estimation problems when $M > T$, which occurs for several of the sub-classes in our investigation, assuming $M = N$.

In order to circumvent the problems caused by not fulfilling the sample size restriction we implement a sub-sampling approach to the estimation of the parameters in [3]. The first step in this approach is to select a sample of products of size $M = min(N\text{-}1, T\text{-}\rho)$, where $T\text{-}\rho$ is the more stringent restriction on sample size, accounting for the rank of the stacked regressor matrix (ρ) discussed by Griffiths *et al*. (2001: 6). The model in [3] is then estimated over M equations and an estimate of r (and so σ) is obtained[12].

We have experimented with varying M from 3 to $min(N\text{-}1, T\text{-}1)$[13] and have seen that as M increases the sample standard errors of the estimates decrease and that the values of σ tend to converge to a point where adding more products does not cause the estimated value to change noticeably. The greatest gains in accuracy tend to be when increasing M from three to about five or six. We draw 50 simple random samples without replacement of size M, principally to assist with the estimation of σ where $N > T\text{-}\rho$, but also to provide additional estimates of accuracy for σ (which we find to be slightly more conservative than the estimated standard errors for r from the model). The final estimate of σ we employ in constructing our index numbers is then the mean of the σ estimated over all samples for that sub-class.

2.3 Calculation of Price Indices

Using the TNS data, price indices have been calculated using eleven index formulae. These include standard weighted, un-weighted and superlative indices, definitions of which can be found below. Lloyd-Moulton indices[14] have also been calculated using σ estimated from the consecutive period algebraic and econometric methods described above. Equal weighted Lloyd-Moulton indices have also been calculated by setting $s^t_m = 1/M$ for all m; as these could be calculated in practice given only an estimate of σ.

Figure 2 – Common Price Index Formulae

$$P_{Carli,t} = \frac{1}{M}\sum_{m=1}^{M}\left(\frac{p_m^t}{p_m^0}\right) \qquad P_{Dutot,t} = \frac{\frac{1}{M}\sum_{m=1}^{M} p_m^t}{\frac{1}{M}\sum_{m=1}^{M} p_m^0} \qquad P_{Jevons,t} = \prod_{m=1}^{M}\left(\frac{p_m^t}{p_m^0}\right)$$

$$P_{Laspeyres,t} = \sum_{m=1}^{M}\left(s_m^0 \frac{p_m^t}{p_m^0}\right) \qquad P_{Paasche,t} = \frac{1}{\sum_{m=1}^{M}\left(s_m^t \frac{p_m^0}{p_m^t}\right)} \qquad P_{Törnqvist,t} = \prod_{m=1}^{M}\left(\frac{p_m^t}{p_m^0}\right)^{\frac{s_m^0 + s_m^t}{2}}$$

Where $P_{x,t}$ is a price index for period t, with period 0 as the base period.

The Fisher price index is defined as the geometric mean of the Laspeyres and Paasche Price Indices.

The indices have all been calculated as annually linked direct indices with January of each

[12] Note that for each of the sub-samples we re-calculate the geometric means of expenditure shares and prices in [3] so that they relate only to the sub-sample.

[13] Due to space restrictions further evidence regarding the performance of the sub-sampling approach is not included here, however further information is available on request from the authors.

[14] The left hand term of [2]

year as the base period. Direct index calculation was chosen as it broadly represents the construction used in the UK CPI; annual linking of the series makes the best use of the data as there are a larger number of common products within years than across the whole period[15].

2.4 Evaluation of results

In order to provide some measure of the performance of the indices described above, they are evaluated against a target index. In each case we have ξ estimates of an index, that is we have $P_{x,t}$ ($x = 1,...,\xi$), eg. $P_{x,t}$ may be the Jevons index number at time t. We then select one of these indices as our target index, which we would like our index numbers to closely approximate. Typically the target index will be a superlative index as they are known to approximate the true level of inflation.

Having selected our target index we evaluate the performance of the ξ-1 remaining indices against it. In this paper we use three loss functions to measure the performance of our indices against our target: Mean Squared Error (MSE); Mean Absolute Deviation (MAD); and QLIKE[16]. Hence assuming that the ξ^{th} index is our target we calculate:

$$MSE_{SC,x} = \frac{1}{T}\sum_{t=1}^{T}\left(P_{\xi,t} - P_{x,t}\right)^2$$

$$MAD_{SC,x} = \frac{1}{T}\sum_{t=1}^{T}\left|P_{\xi,t} - P_{x,t}\right|$$

$$QLIKE_{SC,x} = \frac{1}{T}\sum_{t=1}^{T}\left(\ln\left(\frac{P_{x,t}}{P_{\xi,t}}\right) + \frac{P_{\xi,t}}{P_{x,t}} - 1\right)$$

Where T is the total number of time periods.

We use multiple loss functions as it is possible for results to be inconsistent across loss functions; hence by using a variety we restrict the possibility that our results are a product of the way in which we have chosen to measure accuracy. We then calculate the average loss across all subclasses to evaluate the performance of each index.

We evaluate the performance of the formulae in predicting the index levels as well as the one and twelve month growth rates. For the growth rates, we also provide a measure of the number of times that the growth of $P_{x,t}$ is in the same direction as the growth of $P_{\xi,t}$.

3. Results

3.1 CES Parameters

Estimates of the CES parameter σ have been calculated using both the algebraic and econometric methods described in section 2. A summary of the results is given in Table 1 below; the estimated values of σ are reasonably consistent between the two approaches[17] and for the majority of sub-classes the interpretation of consumer substitution behaviour, if not the estimated value is comparable.

Estimates of σ suggest that within alcohol there is substitution behaviour present in many of

[15] Consecutive monthly indices were considered to maximise the usable data, however this was discounted due to problems of chain drift in high frequency data.

[16] The QLIKE is included as, like the MSE, it can be shown to be a robust loss function. Robust here mean that when comparing forecasts to a proxy for a latent variable, such as inflation, the ranking resulting from the loss function is not affected by noise in the proxy (which is here the Fisher index.) See Patton (2011) for further details.

[17] Ivancic *et al*. (2009) find some variation in estimates within the algebraic approach and also compared to the econometric approach. Our results for the algebraic approach are not directly comparable in terms of the differences that they find as we are averaging over time to arrive at an estimate of σ; however by estimating σ over 33 bilateral periods as opposed to the 4 in Ivancic *et al*. (2009) we have a clearer picture of the message underneath the volatility may expect greater similarity between the algebraic and econometric approaches.

the sub-classes.

Where there are large differences between the econometric and consecutive period approach, estimates (such as B3 and W3), the mulitpile lag approach gives algabraic estimates of σ which are closer to the econometric approach as both combine information from all periods.

Table 1 – Estimates of CES Parameter

Sub-Class		Econometric Approach		Consecutive Period Algebraic Approach		Multiple Lag Algebraic Approach	
		Estimate of σ	Standard Error	Estimate of σ	Standard Error	Estimate of σ	Standard Error
B1	Bitter (4 Cans)	0.9	0.3	1.3	0.3	1.0	0.2
B2	Lager (4 Cans)	1.0	0.3	0.8	0.5	0.7	0.4
B3	Lager (12 Cans)	3.7	1.1	1.8	2.8	2.7	0.9
B4	Lager (4 Bottles)	0.5	0.6	0.8	0.9	0.0	0.8
B6	Lager (24 Stubbies)	2.9	1.3	0.9	1.3	1.9	1.0
S1	Brandy	0.5	0.4	0.4	2.0	0.4	0.9
S2	Spirit Based Drink	1.3	0.4	0.9	1.0	1.2	0.5
S3	Vodka	5.7	0.6	1.3	0.5	4.8	1.3
S4	Whisky	2.4	0.1	2.7	1.1	3.0	0.3
W2	Cider	0.2	0.4	1.6	1.2	-0.2	0.6
W3	Fortified Wine	0.6	0.4	1.8	0.5	1.8	1.0
W4	Red Wine (European)	1.0	0.4	0.9	0.7	1.0	0.2
W5	Red Wine (New World)	3.8	0.4	4.8	0.5	3.9	0.4
W6	White Wine (European)	1.3	0.3	1.8	0.7	1.6	0.3
W7	White Wine (New World)	3.0	0.4	4.2	0.7	3.6	0.5
W9	Imported Sparkling Wine	1.1	0.4	2.2	0.5	2.0	0.8

3.2 EA Indices

Table 2 shows the performance of each index as described in 2.4. The Törnqvist always performs significantly better than the other indices; this result is expected as both Fisher and Törnqvist are superlative indices and both approximate the true rate of inflation.

Table 2 - Performance against Fisher Index

Index	Performance against Fisher (rank in Brackets)					
	MSE		MAD		QLIKE × 1000	
Törnqvist	0.52	(1)	0.22	(1)	0.02	(1)
Base Weighted Lloyd-Moulton (Econometric)	5.10	(2)	1.64	(3)	0.22	(2)
Base Weighted Lloyd-Moulton (Algebraic)	5.33	(3)	1.60	(2)	0.27	(3)
Dutot	11.33	(4)	2.41	(4)	0.51	(5)
Carli	11.34	(5)	2.44	(5)	0.49	(4)
Paasche	15.02	(6)	3.05	(7)	0.70	(7)
Laspeyres	17.09	(7)	3.20	(8)	0.67	(6)
Jevons	18.55	(8)	2.89	(6)	0.82	(8)
Equal Weighted Lloyd-Moulton (Econometric)	29.74	(9)	3.90	(9)	1.38	(9)
Equal Weighted Lloyd-Moulton (Algebraic)	57.76	(10)	4.44	(10)	2.82	(10)

An interesting result is the poor performance of the equal weighted Lloyd-Moulton indices. As Figures 3 & 4 show, the equal weighted Lloyd-Moulton indices are often much lower than the target indices and it can be seen in Table 2 that on average these indices perform very poorly. In fact, the widely used unweighted indices (Carli, Jevons and Dutot) are all closer to the target[18]. It is interesting to note that the Carli and Dutot Indices both outperform Jevons, despite most estimates of σ being greater than 0.5, and many greater than 1. Evaluating their performance by subclasses (table 5) Carli performs best for 8 out 16 subclasses whilst Jevons and Dutot perform best for 4 subclasses each.

Interestingly, the Weighted CES indices perform very well and are the best performers (after the Törnqvist), showing that if estimates of σ were available then the CES indices would be a

[18] This makes sense given the often large values for σ used in the calculation of the equal weighted Lloyd-Moulton indices. The Carli is an equal weighted Lloyd-Moulton index with $\sigma = 0$ and the Jevons is the limit of the equal weighted Lloyd-Moulton index as σ approaches 1.

good approximation of a Fisher Index without the need for current period expenditure weights[19]. It was also found for notable disparities in the algebraic and econometric estimates of σ, the weighted Econometric CES index performs better than its algebraic equivalent.

As expected, for sub-classes with estimates of σ close to 0 the un-weighted CES index is very similar to the Carli, likewise for estimates of σ close to 1 the un-weighted CES index is similar to the Jevons; we also found for sub-classes with estimates of σ close to 2 the un-weighted CES index is very similar to the Harmonic Mean of price relatives.

Table 3 - Performance of Growth Rates across Sub-Classes

Index	Performance against Fisher (MAD)[20]		Average Number of Correct Growth Directions	
	1 Month Growth	12 Month Growth	1 Month Growth (32 Total)	12 Month Growth (22 Total)
Törnqvist	0.18 (1)	0.25 (1)	31.4 (1)	20.8 (1)
Base Weighted Lloyd-Moulton (Econometric)	0.99 (2)	1.41 (2)	28.2 (3)	18.6 (2)
Base Weighted Lloyd-Moulton (Algebraic)	1.03 (3)	1.57 (3)	28.4 (2)	18.4 (3)
Paasche	1.08 (4)	1.74 (4)	27.9 (5)	17.9 (4)
Laspeyres	1.09 (5)	1.77 (5)	28.2 (3)	17.8 (5)
Dutot	1.61 (6)	2.30 (6)	24.9 (8)	16.2 (8)
Carli	1.62 (7)	2.45 (8)	24.7 (10)	16.4 (6)
Jevons	1.63 (8)	2.32 (7)	24.9 (8)	16.3 (7)
Equal Weighted Lloyd-Moulton (Econometric)	1.65 (9)	2.59 (9)	25.0 (7)	16.1 (9)
Equal Weighted Lloyd-Moulton (Algebraic)	1.75 (10)	2.99 (10)	25.1 (6)	14.6 (10)

In Table 3, the performance of the growth rates in the different indices is evaluated. From this table, it is clear that expenditure weights are key when estimating growth, as all weighted indices perform best. Similarly to the levels, Dutot and Carli outperform Jevons; however there is little difference between the two un-weighted arithmetic indices.

Figure 3 - CES Index Series for S1: Brandy[21]

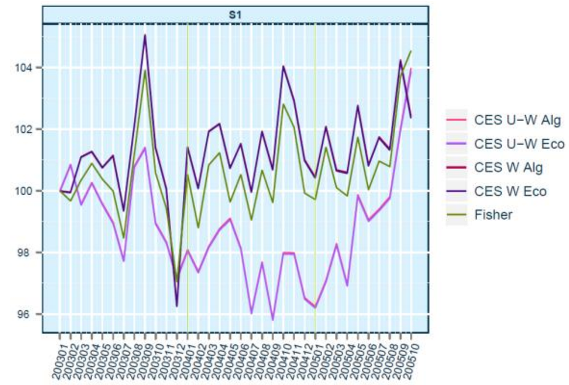
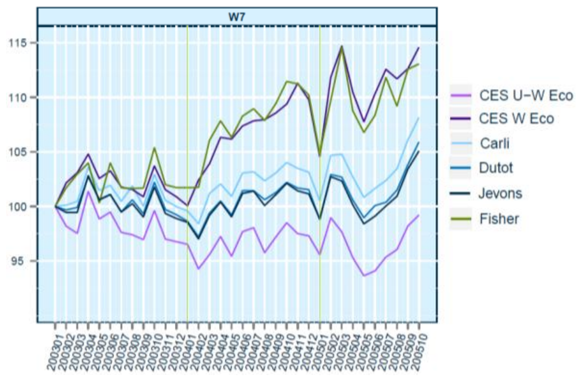


Figure 4 – Index Series for W7: New World White Wine



3.3 Higher Level Indices

Higher level indices have been constructed for each index formula by weighting together EA series using 2005 CPI weights; these have then been evaluated as above[22]. The results for class (averaged over all three classes) and group (all alcohol) levels can be found in Table 4.

We find similar results to that of the lower level results. Weighted CES Indices consistently perform well achieving ranks of 2nd and 3rd at both the class and group level.

[19] This is a more useful result for higher level indices where expenditure shares are available without the high costs associated with low level data. This would also rely on the assumption that $\boldsymbol{\sigma}$ is constant over time, which we have not yet explored.

[20] Results are consistent between MSE and MAD so only one has been presented here.

[21] Note that due to very similar values of σ for the two approaches in the sub-class Brandy, the index series are almost identical.

[22] It was decided that all EA indices should be aggregated in this way. Whilst some, like Fisher do not have consistency of aggregation, the focus of this investigation is on the EA level so it was judged to be an effective way to assess the impact of different EA formulae at higher levels, since this is the process followed in UK consumer price indices.

Figure 5 – Aggregate Index series for Spirits

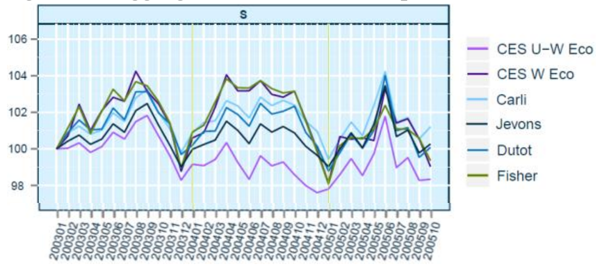
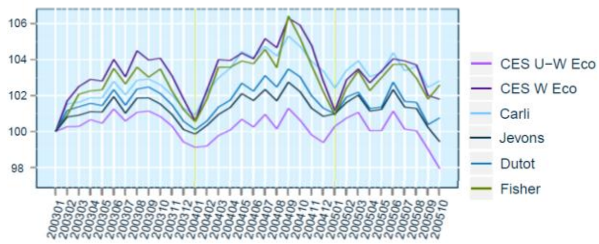


Figure 6 – Selected Index Series for All Alcohol



Surprisingly, at the group and class levels Carli[23] pulls away from Dutot in terms of MAD. It is also interesting that at the all alcohol level, the Jevons outperforms the Paasche quite considerably in terms of the level of the index.

Table 4 - Performance of Higher Level

	Group (all alcohol level)						Average of Class levels					
	Performance against Fisher (MAD)				Correct 12 month Growth Directions		Performance against Fisher (MAD)				Correct 12 month Growth Directions	
	Index Level		12 Month Growth		Out of a possible 22		Index Level		12 Month Growth		Out of a possible 22	
Törnqvist	0.09	(1)	0.11	(1)	20	(1)	0.09	(1)	0.13	(1)	21.0	(1)
Base Weighted Lloyd-Moulton (Algebraic)	0.34	(2)	0.36	(3)	19	(2)	0.59	(3)	0.65	(3)	19.7	(2)
Base Weighted Lloyd-Moulton (Econometric)	0.52	(3)	0.28	(2)	18	(4)	0.54	(2)	0.49	(2)	19.0	(3)
Carli	0.54	(4)	0.70	(4)	18	(4)	0.89	(4)	1.02	(6)	17.7	(5)
Dutot	1.22	(5)	0.75	(5)	15	(7)	1.19	(5)	1.06	(7)	18.0	(4)
Jevons	1.66	(6)	0.81	(6)	17	(6)	1.55	(6)	1.17	(8)	17.0	(8)
Paasche	2.54	(7)	0.88	(8)	19	(2)	2.22	(7)	0.95	(4)	17.7	(5)
Laspeyres	2.65	(8)	0.87	(7)	15	(7)	2.31	(8)	0.96	(5)	17.3	(7)
Equal Weighted Lloyd-Moulton (Econometric)	2.69	(9)	1.29	(9)	12	(9)	2.37	(9)	1.56	(9)	14.7	(9)
Equal Weighted Lloyd-Moulton (Algebraic)	3.21	(10)	1.58	(10)	9	(10)	2.64	(10)	1.72	(10)	14.3	(10)

4. Influencing the choice?

The Un-Weighted CES Indices perform poorly against all targets, but whilst these formulae do not provide a solution to the choice of index formula for EAs, the estimates of σ associated with them may provide some insight into which of the widely used un-weighted indices is most appropriate.

Table 5– Relative Performance of Un-Weighted Indices

Sub - Class	B1	B2	B3	B4	B6	S1	S2	S3
Econometric σ	**0.9**	**1.0**	**3.7**	**0.5**	**2.9**	**0.5**	**1.3**	**5.7**
MAD[24] (Fisher vs Jevons)	2.25	3.46	1.01	1.51	3.11	2.11	3.64	0.95
MAD (Fisher vs Dutot)	2.47	2.54	1.08	1.60	3.20	2.14	2.40	0.80
MAD (Fisher vs Carli)	4.25	1.83	0.98	2.18	3.54	1.77	2.84	1.01
Sub - Class	**S4**	**W2**	**W3**	**W4**	**W5**	**W6**	**W7**	**W9**
Econometric σ	**2.4**	**0.2**	**0.6**	**1.0**	**3.8**	**1.3**	**3.0**	**1.1**
MAD (Fisher vs Jevons)	1.16	3.44	1.46	1.54	1.42	1.82	5.95	10.07
MAD (Fisher vs Dutot)	0.85	2.55	1.82	1.50	1.45	1.60	5.67	5.75
MAD (Fisher vs Carli)	0.96	2.66	2.25	1.13	1.31	0.83	4.33	5.95

Table 5 shows the mean absolute differences of the three widely used un-weighted indices from the Fisher index along with the econometric estimates of σ. The un-weighted arithmetic

[23] This is surprising since use of the Carli Index is not sanctioned for the Harmonized Indices of Consumer Prices as it is liable to have a significant upward bias and the axiomatic approach shows it to have some undesirable properties such as failing the time reversal test (ILO, 2004).

[24] Results are consistent between MSE and MAD so only one has been presented here.

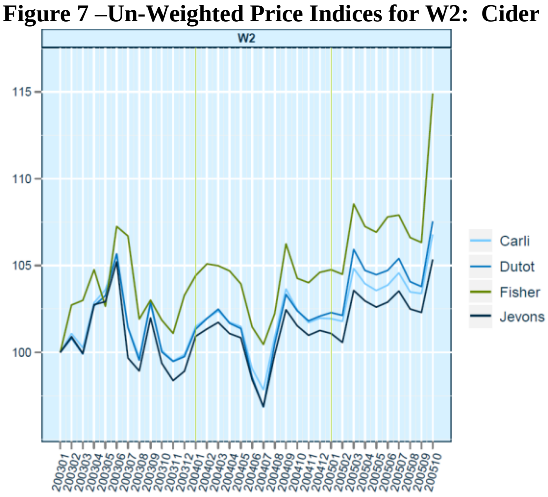
mean of price relatives (Carli) assume that no substitution occurs (σ = 0) whereas un-weighted indices using a geometric mean to combine price relatives (Jevons) assume that consumers substitute goods to maintain a constant expenditure (σ = 1). Given this, one might expect that for sub-classes with $\sigma > 1$ the Jevons Index should be closer to the target than the Carli and for sub-classes with σ close to 0, the reverse should be true.

Figure 7 shows this relationship quite clearly; the sub-class Cider has an estimated CES of 0.2 and it is clear from the graph that the arithmetic mean indices are consistently closer to Fisher.

Despite this clear relationship in certain sub-classes, Table 4 shows that more often than not, this relationship does not hold. However, where this is the case, there is often very little difference in the performance of Jevons and Dutot against Fisher.

It is important to remember that whilst substitution behaviour may influence the choice of index formula at EA level, there are also many other considerations such as ensuring homogeneity of prices through well defined EAs, and accounting for model misspecification due to the presence of demand side effects that in the above model are assumed to be nil.

Figure 7 –Un-Weighted Price Indices for W2: Cider



5. Conclusions

In this section we summarise the main conclusions from the work we have conducted above and note areas where the work might be expanded in the future.

As mentioned above it is commonly asserted in the index numbers literature that where consumers are thought to be substituting between products as prices change the Jevons index is the most appropriate elementary aggregate for use in constructing price indices. Our empirical investigation identifies several cases where elasticity is estimated to be greater than 1 but the Carli is better able to approximate the Fisher index than the Jevons. This is counter to the assertion that Jevons is the most appropriate formula when substitution behaviour is present and is a result worthy of further investigation.

The weighted Lloyd-Moulton index performs well in our example above, while the unweighted version of the index does not. We are not surprised that the weighted version of the index performs well given that the index is calculated for the same time period over which the model is fitted to estimate elasticity. A better test of this result would require a data set which covers longer time periods. This would then allow for elasticity of substitution to be estimated within a given period and then be applied to an out of sample period. If the weighted Lloyd-Moulton continues to perform well in the out of sample period this would be an important verification of the results presented in this paper.

One possible source of uncertainty within the results presented above is the models used to estimate elasticity itself. In the CES model it is necessary to assume that price changes are caused by supply side effects, thus ignoring the possible influence of demand side effects, characterised by the assumption that taste parameters are constant over time. This may not be realistic and may contribute to the result described above which sees the Carli outperform the Jevons despite the presence of substitution.

We also assume that the elasticity of substitution is stable over time and products. Our experiences with sub-sampling in both the algebraic and econometric settings would seem to imply that such assumptions are not true. Given these reservations about the model used to produce estimates of the CES it would seem prudent to attempt to investigate alternative methods of estimation in the future.

We have identified problems with the framework for estimating the CES, the performance of the un-weighted Lloyd-Moulton index and the use of elasticity as a guide for choosing between Jevons and Carli EAs. This may be evidence that rather than focusing on which formula to use it may be better to focus on determining homogenous EAs which would alleviate the need for such considerations as the difference between the two would be minimised. However such an approach may be costly and an alternative may be to estimate a weighted Lloyd-Moulton index using lagged scanner data which could represent a less expensive alternative to estimating a superlative index.

6. References

Balk, B. M. (1999) *'On Curing the CPI's Substitution and New Goods Bias'*, 5th Meeting of the Ottawa Group.

Balk, B. M. (2002) *'Price Indexes for Elementary Aggregates: The Sampling Approach'* Statistics Netherlands Research Paper, No. 0231, Voorburg, The Netherlands.

Griffiths, W.E., Chotikapanich, D. & Skeels, C.L. (2001) '*Sample Size Requirements for Estimation in SUR Models'*, Department of Economics – Working Paper Series 794, The University of Melbourne.

ILO (2004) *'Consumer Price Index Manual: Theory and Practice'*, Geneva, International Labour Office

Ivancic L., Diewert W. E. & Fox K. J. (2009) *'Using a Constant Elasticity of Substitution Index to estimate a Cost of Living Index: from Theory to Practice'*, Discussion Papers 2012-15, School of Economics, University of New South Wales.

Leyland, J. (2011) *'RPI versus CPI - The Definitive Account'*, Significance Web Exclusive Article, http://www.significancemagazine.org/details/webexclusive/1314363/RPI-versus-CPI---The-Definitive-Account.html

ONS (2011) *'CPI and RPI: increased impact of the formula effect in 2010*', Office for National Statistics Information Note. http://www.ons.gov.uk/ons/guide-method/user-guidance/prices/cpi-and-rpi/cpi-and-rpi--increased-impact-of-the-formula-effect-in-2010.pdf

Patton, A. (2011) *'Volatility Forecast Comparison Using Imperfect Volatility Proxies',* Journal of Econometrics, Vol. 160(1) pp. 246-256

Roe, D. (2005) *'Consumer Prices Index and Retail Prices Index: The 2005 Basket of Goods and Services'*, Office for National Statistics. http://www.ons.gov.uk/ons/rel/cpi/cpi-rpi-basket/consumer-prices-index-and-retail-prices-index--the-2005-basket-of-goods-and-services/2005-basket.pdf

Roe, D. & Fenwick D. (2003), '*The New Inflation Target: the Statistical Perspective*', Office for National Statistics. http://www.ons.gov.uk/ons/rel/cpi/consumer-price-inflation/new-inflation-target---full-text/new-inflation-target---full-text-.pdf

Silver, M. & Heravi, S. (2006) *'Why Elementary Price Index Number Formulas Differ: Price Dispersion and Product Heterogeneity'* IMF Working Paper WP/06/174