IMF Working Paper

Why Elementary Price Index Number Formulas Differ: Price Dispersion and Product Heterogeneity

Mick Silver and Saeed Heravi

WP/06/174

IMF Working Paper

**Why Elementary Price Index Number Formulas Differ:
Price Dispersion and Product Heterogeneity**

Prepared by Mick Silver and Saeed Heravi[1]

Authorized for distribution by Adriaan M. Bloem

July 2006

Abstract

| **This Working Paper should not be reported as representing the views of the IMF.**
The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate. |
| --- |

The *Consumer Price Index Manual* (2004) provides guidelines for aggregation formulas that are promulgated at IMF training courses and technical assistance missions. This paper develops elementary level aggregation theory to better inform users and compilers. Most countries use either the Dutot or Jevons index formula. These formulas generally give different results; advice on choice of formula matters. Using an approach based on sample estimators, and an illustration based on scanner data, the paper shows how differences in these formulas can be explained by changes in price dispersion and, in turn, by product heterogeneity. Implications for choice of formula are considered.

Contents   Page

I. INTRODUCTION

Choice of formula for the measurement of inflation does matter. In January 1999 the formula principally used for aggregating price changes for the U.S. consumer price index (CPI) at the lower level of aggregation was changed to a weighted geometric mean of price relatives. The effect of the change was estimated by the Bureau of Labor Statistics (2001) to reduce the annual rate of increase in the CPI by approximately 0.2 percentage points. This implied a cumulative additional national debt from over-indexing federal government expenditures and tax receipts of more than $200 billion over the 12-year period up to the mid-1990s (estimates based on Boskin *et al.*, 1996 and 1998).

"Lower" level aggregation formulas are typically applied to prices of finely defined goods, such as varieties of apples, sampled from outlets. Most often, these formulas use unweighted averages of price observations. They are the building blocks of a CPI.[2] The contribution of this paper lies in the analysis of the differences between these unweighted index number formulas in terms of price dispersion.

The empirical section of the paper uses highly detailed scanner data from retailers' barcode readers that amount to about 31,000 observations over 24 months on prices, characteristics, and brands of models of television sets (TVs) sold in different outlet types.

The focus of the paper is on the difference between two lower-level formulas—the ratio of unweighted arithmetic means of prices (Dutot index) and the ratio of unweighted geometric means (Jevons index). Both formulas are commonly used, both can be justified under particular circumstances, but they can give quite different results.

Sections II.A and II.B, respectively, outline the formulas and discuss their usage and relative merits. The paper then first considers what factors underlie the difference between the two index number formulas. In Section III.A, an analytical framework is derived for this purpose which benefits from being able to distinguish calculated indexes based on sample data as estimators of their population counterparts.[3] The difference between the two formulas is shown to depend on the change over time in price dispersion. Second, we discuss which of these formulas is to be preferred. From the axiomatic approach it is shown that the Dutot index should not be used for heterogeneous item groups (Section II.B). Some of the price dispersion and difference between the formulas will be due to product heterogeneity. In Section III.B, the analytical framework is extended to show the residuals from hedonic regressions to be a useful basis for compiling a heterogeneity-corrected Dutot index to further account for the difference between the Dutot and Jevons "raw" price indexes. Third, the paper considers whether

---

[2] The resulting lower level indexes of price changes are combined at the higher level using a base-period weighted arithmetic mean of price changes to form the overall index.

[3] The issue of sampling error is thus explicitly considered in the analytical framework, as suggested by Dorfman *et al.* (1999) and Greenlees (2001).

the difference due to product heterogeneity can be substantial. In the empirical section (IV), monthly scanner data on television sets are used to compile Dutot, Jevons, and heterogeneity-corrected Dutot indexes, and to explain their differences in terms of changes in their price dispersion and product heterogeneity. In Section V, consideration is given to the question of whether the Jevons index is preferred over a heterogeneity-corrected Dutot index, that is, a Dutot index undisturbed by issues of product heterogeneity. For this purpose, the approach of considering the calculated indexes as sample estimators is again adopted but, this time, as estimators of population superlative indexes.

II. ELEMENTARY INDEX NUMBER FORMULAS, THEIR USE AND JUSTIFICATION

A. Elementary Index Number Formulas

The main lower-level formulas used in practice for $m=1,..M$ matched items with prices and quantities, p_m^τ and q_m^τ respectively for periods $\tau = 0, t$ are:[4]

The arithmetic mean of price relatives—the *Carli price index*, P_C:

$$P_C \equiv \sum_{m=1}^{M} \left(p_m^t / p_m^0 \right) / M \tag{1}$$

The ratio of the arithmetic means—the *Dutot price index, P_D*:

$$P_D \equiv \left(\sum_{m=1}^{M} p_m^t / M \right) \Bigg/ \left(\sum_{m=1}^{M} p_m^0 / M \right) = \sum_{m=1}^{M} \left(p_m^t / p_m^0 \right) \cdot p_m^0 \Big/ \sum_{m=1}^{M} p_m^0 \tag{2}$$

which can be seen to be a base-period price-share weighted Carli index.

The geometric mean of price relatives (which is also equal to the ratio of geometric means)—the *Jevons price index, P_J*:

$$P_J \equiv \prod_{m=1}^{M} \left[p_m^t / p_m^0 \right]^{1/M}$$

$$= \left(\prod_{m=1}^{M} \left[p_m^t \right]^{1/M} \right) \Bigg/ \left(\prod_{m=1}^{M} \left[p_m^0 \right]^{1/M} \right) = \exp \left[M^{-1} \sum_{m=1}^{M} \log(p_m^t) \right] \Bigg/ \exp \left[M^{-1} \sum_{m=1}^{M} \log(p_m^0) \right] \tag{3}$$

---

[4] The main alternative formulas to those presented in this section are the harmonic mean of price relatives, P_H, the Carruthers, Sellwood, Ward, and Dalen (P_{CSDW}) index, which is the geometric mean of P_C and P_H, and the Balk-Walsh index (see Dalen, 1992 and Balk, 2002). Their differences can all be phrased in terms of price dispersion and its change.

B. Their Use and Justification

The explanation of differences between formulas may be argued to be of cursory interest if only one formula was in general use and was superior to the others on theoretical grounds. In this section we demonstrate why the focus of this study is on two formula, the Jevons and Dutot indexes, by showing both are widely used and have their justification.

First, we consider the extent to which the formulas are used for CPI compilation. Of a sample of 37 countries,[5] 13 countries used the Dutot index, 14 the Jevons index, and 4 the Carli index, with 6 other countries primarily using Jevons indexes, but with Dutot/Carli used for specific product groups. Second, we consider their relative merits from an axiomatic (test) approach and also from economic theory.

The Carli index fails the important time reversal test such that a Carli index calculated forwards, between periods 0 and t, exceeds one calculated backwards, between periods t and 0: $P_C(p^0 , p^t) \times P_C(p^t , p^0) \geq 1$; it is *upwards-biased*.[6] Given this failing of the Carli index, and that it is hardly used, we focus on the relationship between the Jevons and Dutot indexes. The Jevons index satisfies all of the main tests. The Dutot index satisfies all of the main tests with the important exception of the commensurability test, i.e., if the units of measurement for each model in each outlet change, then the elementary index remains unchanged (Diewert, 1995 and 2004). However, for homogeneous items commensurability is not an issue and the Dutot index can then be used, as advised in the CPI Manual (ILO *et al*., 2004).

The CPI Manual notes that the economic approach provides only "weak support" for the Jevons index (ILO *et al*., 2004: 370). The support for the Jevons index from the economic approach is for product markets in which consumers substitute away from items with above average price increases. More particularly, the cross-elasticities of demand are required to be unitary—expenditure shares remain constant as relative prices change—consumers are assumed to have Cobb-Douglas preferences. If the sampling of items is with probability proportional to *expenditure shares* for the product in the base period, and base period and current period expenditure shares are equal, the Jevons index acquires the properties of an symmetric base and current period-weighted superlative index. An economic theoretic justification for the Dutot index relies on assuming Leontief preferences, for which a change in relative prices leads to no change in relative quantities consumed. Such behavior would be reflected by the sampling of homogeneous items with probabilities proportional to relative base period *quantities* for the product. However, not

---

[5] The countries were those whose methodology was reported on the IMF's Dissemination Standards Bulletin Board: dsbb.imf.org. Not all countries provided sufficient detail and an email survey of such countries clarified the position for 37 out of the 51 countries on the Website, at the time of the survey.

[6] Fisher (1922) famously commented: "In fields other than index numbers it is often the best form of average to use. But we shall see that the simple arithmetic average produces one of the very worst of index numbers. And if this book has no other effect than to lead to the total abandonment of the simple arithmetic type of index number, it will have served a useful purpose" (pp. 29-30).

all CPIs will have such well determined sampling schemes to justify the Jevons index and not all product markets will have a priori grounds for expecting the appropriate substitution behavior.[7]

III. DIFFERENCES BETWEEN THE JEVONS AND DUTOT FORMULAS

A. Jevons, Dutot and Price Dispersion

The difference between the elementary aggregate index formulas, in terms of changes in the variances of the prices, is generally considered for sample indexes by means of a Taylor approximation (see Dalen, 1992; Diewert, 1995 and 2004; and Balk, 2002 for details). The statistical exposition here differs somewhat from these earlier approaches because it first, provides an exact explanation of the difference between population Dutot and Jevons indexes and, in the empirical section (IV), we find a very close correspondence for sample indexes. Second, the statistical exposition draws attention to the distinction between a calculated index based on a sample of the target population, as an index estimator of the population index, as advised by McClelland and Reinsdorf (1999), Dorfman *et al.* (1999) and Greenlees (2001). McClelland and Reinsdorf (1999) draw attention to the small sample bias in the sample Jevons index as an estimator of its population counterpart.

The sample Carli index, P_C, in equation (1), as an arithmetic mean of the sample price relatives, is a consistent and unbiased estimator of the population Carli index, $I_C = \mathrm{E}(p^t/p^0)$.

The sample Dutot index, P_D, in equation (2), as a ratio of two sample arithmetic means of prices, is a consistent, but not unbiased, estimator of the ratio of population means, the population Dutot index,

$$I_D = \frac{\mathrm{E}\left[p^t\right]}{\mathrm{E}\left[p^0\right]} = \frac{\mu^t}{\mu^0}. \tag{4}$$

The sample Jevons index, P_J, in equation (3), as a ratio of the exponents of two sample means of log prices, is a consistent estimator of the population Jevons index,

$$I_J = \frac{\exp\left[\mathrm{E}\left[\log\left(p^t\right)\right]\right]}{\exp\left[\mathrm{E}\left[\log\left(p^0\right)\right]\right]} = \frac{\exp\left(\eta^t\right)}{\exp\left(\eta^0\right)} = \exp\left(\eta^t - \eta^0\right) \tag{5}$$

where $\eta^t = \mathrm{E}\left[\log\left(p^t\right)\right]$ and $\eta^0 = \mathrm{E}\left[\log\left(p^0\right)\right]$.

Since the exponential function cannot be taken through expected values:

---

[7] While much of the logic for the adoption of the geometric mean index for the U.S. CPI was based on the economic approach, it was still only applied to 61 percent of expenditure-weighted product groups on the grounds of expected substitution behavior.

$$\mu^\tau \equiv E\left[p^\tau\right] \neq \exp\left[E\left[\log\left(p^\tau\right)\right]\right] = \exp\left(\eta^\tau\right) \qquad \text{for } \tau = 0,t , \tag{6}$$

and by Jensen's inequality:

$$\mu^\tau \equiv E\left[p^\tau\right] > \exp\left[E\left[\log\left(p^\tau\right)\right]\right] = \exp\left(\eta^\tau\right) \qquad \text{for } \tau = 0,t . \tag{7}$$

As such the numerator of I_D will exceed the numerator of I_J, as will the denominator, making it impossible to determine which effect will dominate, without making a further distributional assumption.

We introduce the distributional assumption:

$$\log\left(p^\tau\right) \square \ Normal\left(\eta^\tau, \varepsilon_\tau^2\right) \qquad \text{for} \quad \tau = 0,t . \tag{8}$$

It follows from the properties of a lognormal distribution that:

$$\mu^\tau = \exp\left(\eta^\tau + \varepsilon_\tau^2 / 2\right). \qquad \text{for} \quad \tau = 0,t . \tag{9}$$

Substituting μ^τ, for $\tau = 0,t$, in equation (4) by equation (9) and using equation (5) gives a relationship between the population Dutot and Jevons indexes in terms of the difference in the variances of log-prices between periods 0 and t:

$$I_D = \frac{\mu^t}{\mu^0} = \frac{\exp\left(\eta^t + \varepsilon_t^2 / 2\right)}{\exp\left(\eta^0 + \varepsilon_0^2 / 2\right)} = I_J \cdot \frac{\exp\left(\varepsilon_t^2 / 2\right)}{\exp\left(\varepsilon_0^2 / 2\right)} = I_J \cdot \exp\left[\left(\varepsilon_t^2 - \varepsilon_0^2\right)/2\right]. \tag{10}$$

B. Jevons and Dutot Indexes and Hedonic Heterogeneity-Controlled Price Dispersion.

It is apparent from equation (10) that as product heterogeneity and price dispersion decreases, so too will the difference between the two indexes. The above exposition carries over to indexes that control for observable product heterogeneity through hedonic regressions. Consider a regression, using data on $m = 1,...,M$ matched models for periods $\tau = 0,t$, of the log of price, $\ln p_m^\tau$, on a dummy variable D^t which takes the value of 1 in period t and zero in a base period 0, and on $k = 2,...,K$ quality characteristics, z_{km}^τ :

$$\ln p_m^\tau = \beta_0 + \beta_1 D^t + \sum_{k=2}^{K} \beta_k z_{km}^\tau + u_m^\tau \tag{11}$$

where u_m^τ is assumed to be normally distributed with mean and variance δ^τ and ξ_τ^2 respectively. The hedonic (quality-adjusted) estimated Jevons index is given by:

$$P_J^* = \exp\left(\hat{\beta}_1\right) \tag{12}$$

which, since matched models are used, is equal to the Jevons index in equation (3). However, the Dutot index failed the commensurability test and is thus itself determined by the extent of price dispersion. A consistent estimator of the hedonic (quality-adjusted) Dutot index is given by:

$$\hat{I}_D^* = \exp\left(\hat{\beta}_1\right) \cdot \exp\left[\left(\xi_t^2 - \xi_0^2\right)/2\right] = P_J^* \cdot \exp\left[\left(\xi_t^2 - \xi_0^2\right)/2\right]. \tag{13}$$

where the * denotes heterogeneity-adjusted and where ξ_τ^2, for $\tau = 0,t$, are the variances of the residuals of observations in periods 0 and t respectively. Thus the difference

between the Jevons index and the Dutot *hedonic* price index is related to the change in the variance of the residuals over time. If $\left(\hat{\xi}_t^2 - \hat{\xi}_0^2 \right) < \left(\varepsilon_t^2 - \varepsilon_0^2 \right)$ (from (13) and (10) respectively) then the discrepancy between the Dutot and Jevons indices in (10) will be greater than the discrepancy between the heterogeneity-controlled Dutot and the Jevons index in (13). Note that first, for $\tau = 0, t$, as $\hat{\xi}_\tau^2 \to 0$, $P_D^* \to P_J^*$. Second, $\left(\hat{\xi}_t^2 - \hat{\xi}_0^2 \right) < \left(\varepsilon_t^2 - \varepsilon_0^2 \right)$ if the hedonic regression controls for the same proportion of price variation in each period, that is $\hat{\xi}_\tau^2 = \delta_\tau \varepsilon_\tau^2$ for $\tau = 0, t$ where $\delta_0 = \delta_t < 1$. Minimizing dispersion from product heterogeneity should account for some of the difference between the Dutot and Jevons indexes.

The use hedonic regressions in equation (13), for the purpose of this paper, is not because of changes in the quality of models compared *over time*. The indexes are all calculated over $m = 1,..,M$ matched models, in order that measured price changes are not distorted by quality changes. The concern of this paper, is to reduce the cross-sectional dispersion in the prices by abstracting out quality differences arising from product heterogeneity. As a result, the failure of the commensurability test by the Dutot index is less of a concern, and the heterogeneity-controlled Dutot index in equation (13) should provide a better estimate of price changes to that of equation (10). Moreover, some of the difference between the Jevons and Dutot indexes will be explained by the reduction in price dispersion. In the empirical section that follows, we first outline the data, then account for the differences between Dutot and Jevons indexes in terms of the change in variances as postulated by equation (10), and finally, compile a heterogeneity-controlled Dutot index using equation (13) for comparison with the Dutot and Jevons indexes.

IV. EMPIRICAL WORK

A. Data and Variables

The empirical work utilized monthly scanner data on the prices of U.K. television sets (TVs) from January 1998 to December 1999. The scanner data were supplemented by data from price collectors for outlets without barcode readers, although this factor was negligible. Each observation is a model of a TV in a given month sold in one of four different outlet types: multiple chain stores, mass merchandisers (department stores), independent stores, and catalogue stores. For example, an observation in the data set for January 1998 was the unit value (price) (£275.80), volume (5,410 transactions) and quality characteristics (see the Data Annex) of the Toshiba 2173DB 21 inch "model" of TV sold in multiple chain outlets only. For the 24 months of January 1998 to December 1999 there were 31,352 observations which covered 4.6 million transactions worth £1.7 billion. The price indexes were calculated over matched identical models sold in both periods; 20,574 observations were used to make matched price comparisons for the period January 1998 to December 1999.

Table 1 provides monthly Dutot and Jevons price indexes for *matched* models of television sets for January 1998 to December 1999 (January 1998=100). For March 1998, for example, only matched models sold in both January 1998 and March 1998 are included in the calculation. This follows the practice of CPI compilation and ensures that like is compared with like; the price change is untainted by changes in quality. Matched model comparisons are also apparent from equations (1) to (3), in which the aggregation in each period is over the same $m = 1,...,M$ items. Since models available in March 1998 may no longer be sold in, say, in December 1998, the matched models for January 1998, in the March 1998 index, will differ from those for January 1998 in the December 1998 index.[8] Note that since the matched sample varies each month, so too does the calculated variance for January 1998. For example, from Table 1, for matched models in February 1998 compared with January 1998, the variance for January 1998 was 0.495; however, for matched models in December 1998, the January 1998 variance was 0.500. The sample is renewed in January 1999 and comparisons between January 1999 and subsequent months are similarly undertaken on a matched basis, and linked to the January 1998 = 100 reference period.

The explanatory variables used for estimating the hedonic regression in equation (11) included 37 brand dummies, 3 outlet-type dummies, 19 screen size dummies, 6 tube-type dummies, and 23 further characteristics as outlined in the Data Annex. Hedonic regressions were estimated for each month between January 1998 and December 1999. The estimated coefficients were almost invariably statistically significant and their signs accord with *a priori* expectations.[9] The average $\bar{R}^2$ for the estimated equations was 0.91 with a maximum of 0.93 and a minimum of 0.89.

B. Dutot and Jevons Indexes

Prices can be seen from Table 1 to fall over the two years by 26.4 and 28.4 per cent respectively for the Dutot and Jevons indexes. Columns 6 and 7 compare the ratio of the Dutot to Jevons index, with and without the adjustment in equation (10) for the calculated variances. In December 1999 the Dutot index is 2.7 percent higher than Jevons. Unity in column 7 of Table 1 reflects the fact that the difference between the Dutot and Jevons indexes is perfectly accounted for by the equation (10) variance change term, and the results are invariably very close to one demonstrating the successful application of the analytical framework to the data. Yet there is a slight divergence from unity in column 7 of Table 1. This is because Equation (10) is based on population parameters and an assumption of a log-normal distribution of prices. The very slight divergence from unity

---

[8] For example, in January 1998 and February 1998 there were, respectively, 1,320 and 1,270 models of TVs (observations) sold. However, the calculations were based on the 1,093 matched models available in both months. By December 1998 there were 1,326 models sold, however, only 798 were available in both periods and, thus, could be used in the calculation. The sample was renewed in January 1999. This follows standard CPI practice, but incurs selectivity bias as considered in Silver and Heravi (2005) and Triplett (2004).

[9] Details are available from the authors.

reflects the fact that sample, rather than population, indexes are calculated and that the distributional assumptions may not hold. In spite of this, the departure from unity is very small.

C. Heterogeneity-Controlled Dutot Indexes

The sample used is the same matched sample used for the Dutot and Jevons indexes calculated in columns 2 and 3 of Table 1. There is a price fall of 27.2 percent given by the heterogeneity-controlled Dutot index, compared with 26.4 percent for the Dutot and 28.4 percent for the Jevons index.[10] The heterogeneity-controlled Dutot is based on much smaller price variances, as shown in the last two columns of Table 1, compared with the variances in columns 4 and 5. Of note is that the relatively effective control of heterogeneity, in terms of the reduction of the variance via the hedonic regression, serves to bring the (heterogeneity-controlled) Dutot index closer to the Jevons index, accounting for over half of the disparity between the two raw price indexes in Columns 2 and 3.

V. IMPLICATIONS FOR CPI COMPILATION

That different countries use different formulas for calculating elementary aggregate indexes for their CPIs begs a question as to whether they will give different results and, if so, the factors underlying such differences. We provided a framework that showed in equation (10) that the difference between the Dutot and Jevons indexes can be explained in terms of the change in the dispersion of prices, as was confirmed by the empirical work in section IV. There is then the question as to how much of the difference between the results of the two indexes can be reasonably attributed to the Dutot index's failure of the commensurability test. Equation (13) showed how hedonic regressions, used to control for price dispersion arising from product heterogeneity, could be used to further explain the difference between the Jevons and Dutot indexes. In the empirical work we found that this reduction in price dispersion accounted for a large part of the difference between the Jevons and Dutot indexes.

If there was no preference for the Jevons or Dutot index by other criteria, the Jevons index would be preferred to the "raw" Dutot, for this empirical illustration, on the grounds of the failure of the commensurability test by the latter. Product heterogeneity was shown to sizably affect the Dutot index. If the regressions are taken to have adequately corrected for product heterogeneity, then the test and variance analysis cannot inform the choice between the Jevons index and the heterogeneity-controlled Dutot index.

A preference for the (heterogeneity-controlled) Dutot or Jevons index can also be considered in terms of which index best approximates an index number formula with

---

[10] The Jevons index does not fail the commensurability test and is thus unaffected by the use of residuals from matched samples.

desirable properties, that is, a superlative index such as the Fisher, Törnqvist, or Walsh index. Yet superlative index number formulas make use of base and current period quantities, which the Dutot and Jevons indexes do not—they are unweighted. Balk (2002) and Diewert (2004) consider the assumptions required for the Dutot and Jevons indexes to be sample estimators of superlative indexes. Quantity/expenditure weights are implicitly introduced if the sampling is with probability proportionate to size, where the measure of "size" is the quantity/expenditure share.

Under an assumption of first, sampling of prices with probability proportionate to base period *expenditure* shares, and second, unity elasticity of substitution, and thus equal base and current period expenditure shares, the Jevons index is a sample estimator of a Törnqvist index. Under an assumption of sampling with probability proportionate to *quantity* shares in the base period (current period) the Dutot index can be shown to be a sample estimator of a population Laspeyres (Paasche) index, and with a further assumption of equal base and current period quantity shares, the Dutot index is a sample estimator of a Fisher index. Only with these assumptions can preferences for the Jevons or (heterogeneity-corrected) Dutot index extend beyond the variance and axiomatic considerations of this paper. The assumptions required for the Jevons index are not as limiting as those for the Dutot index, though their veracity in real CPI practice will often be questionable.

The analysis has been conducted for matched samples, following the practice largely used by statistical offices in the compilation of their CPIs. However, for product areas with a large product turnover, such as domestic appliances and personal computers, "new" unmatched models introduced in period t, but not existing in period 0, and "old" unmatched models existing in period 0, but no longer in period t, are unlikely to have residuals of the same magnitude as those for matched models. Silver and Heravi (2005) examined such residuals for cameras, dishwashers, television sets, vacuum cleaners, and washing machines. They found for each product the mean of the residuals for unmatched new models to be *above* the corresponding mean for unmatched old ones. The resulting increase in dispersion would effect a larger difference between the Dutot and Jevons indexes than reported for conventional CPI matched models measurement, as undertaken in this study.

VI. SUMMARY

This paper has provided an improved analytical framework for establishing the difference between the Dutot and Jevons indexes. Previous such work was based on a Taylor approximation. Our framework provides an exact expression, and one that further benefits from expressing indexes as sample estimators of their population counterparts. It is not usual in index number theory that their sampling basis is recognized. The paper then decomposes the difference between the two formulas into that due to product heterogeneity, and that essentially due to the fact that the index formulas are different types of averages.

So what advice can be given to statistical offices? Which is better?

(i) Jevons can be said to be better because of the difference due to product heterogeneity, since the Dutot index fails the commensurability (units of measurement) test and the Jevons index does not. In our empirical example we found product heterogeneity to account for over half the difference between the two formulas. This would argue for the Jevons index. But if product heterogeneity was responsible for only a small proportion of the difference; on what grounds could we then choose between the two index numbers?

(ii) As noted in the previous section, the choice between formulas would be on the basis of their sample design. Under an assumption of 1) sampling of prices with probability proportionate to base period *expenditure* shares, and 2) unity elasticity of substitution, and thus equal base and current period expenditure shares, the Jevons index is a sample estimator of a desirable Törnqvist index. Under an assumption of sampling with probability proportionate to *quantity* shares in the base period (current period) the Dutot index can be shown to be a sample estimator of a population Laspeyres (Paasche) index, and with a further assumption of equal base and current period quantity shares, the Dutot index is a sample estimator of a desirable Fisher index. Thus statistical offices might use different formulas for different product areas depending on the likely validity of assumptions about consumer behavior and the sample design used.

 If such assumptions and sampling systems do not hold, there is little further we can say about the two formulas, at least from economic theory or axiomatic tests. The geometric mean is known to be more robust to outliers, but less easy to explain than the arithmetic mean.

(iii) The analysis has been conducted for matched samples, following the practice largely used by statistical offices in the compilation of their CPIs. However, for product areas with a large product turnover the case for a geometric formulation is stronger.

(iv) Much of the analysis of this paper carries over to producer price indexes, with the exception of the economic analysis where the assumption about cost-minimizing consumer behavior differs from that relating to revenue-maximizing producer behavior (see Diewert, 2004).

Data Annex

The variable set *on each observation* included: **price,** the unit value of a model across all transactions in a month/outlet-type and **volume,** the number of transactions during the period in the outlet type - many of the models sold had relatively low sales. There were 38 **brands**—37 dummy variables benchmarked on Sony (the excluded brand in the regression). The **characteristics** included:

   (i) size of screen—dummy variables for about 19 screen sizes;
   (ii) possession of Nicam stereo sound;
   (iii) wide screen;
   (iv) on-screen text retrieval news and information panels from broadcasting companies, in order of sophistication: teletext, fastext and top fastext—3 dummy variables;
   (v) 6 types of reception systems—5 dummy variables;
   (vi) continental monitor style;
   (vii) Dolby Pro, Dolby SUR/DPL, Dolby Digital sound—3 dummy variables;
   (viii) Flat & Square, Super-Planar tubes—2 dummy variables;
   (ix) s-vhs socket;
   (x) satellite tuner, analogue/digital—2 dummy variables;
   (xi) digital;
   (xii) DVD playback or DVD recording—2 dummy variables;
   (xiii) rear speakers;
   (xiv) without PC-internet/PC+internet;
   (xv) real flat tube;
   (xvi) 100 hertz, doubles refresh rate of picture image;
   (xvii) vintage; and
   (xviii) DIST—the percentage of outlets in which the model was sold.

 Outlet-types are multiple chains, mass merchandisers (department stores), independents and catalogue stores.

REFERENCES

Balk, Bert M., 2005, "Price Indexes for Elementary Aggregates: The Sampling Approach," *Journal of Official Statistics*, Vol. 21, No. 4, pp. 675–99.

Bureau of Labor Statistics, 2001, "The Experimental CPI Using Geometric Means (CPI-U-XG)," last modified, October 16, 2001. Available via the Internet at http://www.bls.gov/cpi/cpigmtoc.htm.

Boskin, Michael. S. (Chair), 1996, Advisory Commission to Study the Consumer Price Index, "Towards a More Accurate Measure of the Cost of Living," Interim Report to the Senate Finance Committee, Washington, D.C.

Boskin Michael, S., Ellen R. Delberger, Robert J. Gordon, Zvi Griliches, and Dale W. Jorgenson, 1998, "Consumer Prices in the Consumer Price Index and the Cost of Living," *Journal of Economic Perspectives* , Vol. 12, No. 1, pp. 3-26.

Dalen, Jorgen, 1992, "Computing Elementary Aggregates in the Swedish Consumer Price Index," *Journal of Official Statistics*, Vol. 8, No. 2, pp. 129-47.

Diewert, W. Erwin., 1995, "Axiomatic and Economic Approaches to Elementary Price Indexes," *Discussion Paper* No. 95-01 (Vancouver: Department of Economics, The University of British Columbia).

_____, 2004, "Elementary Indices," in *Consumer Price Index Manual: Theory and Practice*, (Geneva: International Labour Office) Chapter 20, pages 355-70. Available via the Internet at www.ilo.org/public/english/bureau/stat/guides/cpi/index.htm.

Dorfman, Alan H., Leaver, Sylvia, and Janice Lent, 1999, "Some Observations on Price Index Estimators," *Statistical Policy Working Paper* Vol. 29, No. 2 (Washington: Bureau of Labor Statistics). Available via the Internet at http://www.bls.gov/ore/abstract/st/st990080.htm.

Fisher, Irving., 1922, *The Making of Index Numbers* (Boston: Houghton Mifflin).

ILO/IMF/OECD/UNECE/Eurostat/World Bank, 2004, *Consumer Price Index Manual: Theory and Practice* (Geneva: International Labour Office). Available via the Internet at www.ilo.org/public/english/bureau/stat/guides/cpi/index.htm.

Greenlees, John S., 2001, "Random Errors and Superlative Indexes," *Bureau of Labor Statistics Working Paper* No. 343, March (Washington: Bureau of Labor Statistics). Available via the Internet at http://www.bls.gov/ore/abstract/ec/ec010110.htm.

McClelland, Robert and Marshall Reinsdorf, 1999, "Small Sample Bias in Geometric Mean and Seasoned CPI Component Indexes," *Bureau of Labor Statistics Working Paper* No. 324, August (Washington: Bureau of Labor Statistics). Available via the Internet at http://www.bls.gov/ore/pdf/ec990050.pdf.

Silver, Mick, and Saeed Heravi, 2005, "Why the CPI Matched Models Method May Fail Us: Results from an Hedonic and Matched Experiment Using Scanner Data," *Journal of Business and Economic Statistics*, Vol. 23, No. 3, pp. 269-81.

Triplett, Jack E., 2004, *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*, Directorate for Science, Technology and Industry (Paris: Organisation for Economic Cooperation and Development).

Table 1. Dutot and Jevons Indexes: Their Relationship to Each Other and to Heterogeneity-Controlled Dutot Indexes

Period	Sample indexes		Sample variances[†] $a_\tau^2 = \left[var\left(log\left(p^\tau\right)\right)\right]$		Ratio of Dutot to Jevons		Heterogeneity-controlled Dutot sample index	Sample variances[†] $\xi_\tau^2 = \mathrm{var}\left(u_m^\tau\right)$	
	P_D	P_J	a_{Jan98}^2	a_t^2	$\dfrac{P_D}{P_J}$	$\dfrac{P_D}{P_J \cdot \exp\left[\left(\left(a_t^2 - a_{Jan98}^2\right)/2\right)\right]}$	$\hat{I}_D^*$	ξ_{Jan}^2	ξ_t^2
January 1998=100									
Jan-98	100.00	100.00			1.000	1.000	100.00		
Feb-98	99.72	99.27	0.495	0.503	1.005	1.000	99.45	0.034	0.038
Mar-98	97.75	97.55	0.506	0.507	1.002	1.002	97.84	0.039	0.044
Apr-98	96.46	96.73	0.497	0.493	0.997	0.999	96.72	0.039	0.039
May-98	95.75	95.92	0.498	0.502	0.998	0.997	96.00	0.039	0.041
Jun-98	94.06	93.62	0.498	0.508	1.005	0.999	94.18	0.040	0.052
Jul-98	93.88	92.99	0.496	0.516	1.009	0.999	93.21	0.040	0.045
Aug-98	91.82	91.35	0.484	0.497	1.005	0.998	91.62	0.040	0.046
Sep-98	89.20	88.69	0.508	0.526	1.006	0.997	88.83	0.040	0.043
Oct-98	88.63	88.06	0.507	0.520	1.006	1.000	88.27	0.040	0.045
Nov-98	86.50	86.51	0.492	0.496	1.000	0.998	86.75	0.040	0.046
Dec-98	84.94	83.87	0.500	0.531	1.013	0.997	84.46	0.040	0.055
Jan-99	85.25	83.90	0.509	0.540	1.016	1.000	84.94	0.031	0.056
Feb-99	84.93	83.96	0.594	0.608	1.012	1.004	84.12	0.049	0.052
Mar-99	82.59	82.09	0.580	0.587	1.006	1.003	81.82	0.050	0.044
Apr-99	82.97	81.43	0.594	0.615	1.019	1.008	81.27	0.054	0.050
May-99	80.62	79.92	0.605	0.616	1.009	1.003	79.75	0.055	0.051
Jun-99	79.61	78.17	0.582	0.610	1.018	1.004	78.29	0.057	0.060
Jul-99	79.49	77.85	0.596	0.630	1.021	1.004	77.89	0.057	0.058
Aug-99	77.42	76.26	0.593	0.613	1.015	1.005	76.28	0.057	0.058
Sep-99	75.99	74.90	0.607	0.635	1.015	1.000	75.13	0.057	0.063
Oct-99	75.51	74.39	0.626	0.652	1.015	1.002	74.33	0.057	0.055
Nov-99	76.78	75.04	0.635	0.667	1.023	1.007	75.49	0.057	0.069
Dec-99	73.59	71.64	0.653	0.704	1.027	1.001	72.75	0.058	0.088