USES OF MODELS IN THE ESTIMATION OF PRICE INDEXES: A REVIEW

Richard Valliant, Westat
1650 Research Boulevard, Rockville MD 20850

Key Words: Geometric, Laspeyres, Törnqvist, optimal allocation, variance components.

1. Introduction

Price indexes are some of the key statistics published by national governments. The survey systems required to estimate indexes are often complex and involve solving a number of statistical problems, including sample design, variance component estimation, sample allocation to meet multiple goals, data adjustments specific to indexes, missing data imputations, and variance estimation for nonlinear estimators. This paper reviews some of the previous research in these areas and notes some of the topics that deserve future research.

A consumer price index (CPI) is a measure of how much the purchasing power of a consumer has changed from one period to another. Indexes are used in other ways such as gauging the change of the output of an economy, but this paper concentrates on statistical problems associated with estimating a CPI. Some of the methods and issues are transferable to other index problems as well.

Models can be used both to define the population index and to guide the construction of estimators. Among economists there is no agreement on which form of index is theoretically preferred and, at the same time, is practical to estimate in an ongoing program. Three approaches have been used in the past to derive indexes (see, e.g., Diewert 1987):

(1) Economic,

(2) Test, and

(3) Stochastic.

In the first, a function is defined that measures the utility to a consumer of purchasing certain quantities of goods. The cost-of-living index between times s and t ($s<t$) is then defined to be the ratio of the minimum cost at time t to the minimum cost at time s of achieving the same level of utility. In the Test Approach, originated by Fisher (1922), a series of desirable properties of an index are listed. A candidate index is examined to see how many of the tests it passes, e.g., monotonicity, circularity, and price/time reversal. With the Stochastic Approach, an attempt is made to model the behavior of prices or ratios of prices at different time periods. A parameter of the model is the overall rate of change in prices. This third approach is a model-based attempt to derive the form of an index and has been in disfavor among economists for some time (see Dorfman 1998).

2. Alternative Indexes

There are a number of indexes that have been developed using one or more of the three approaches listed above. In preparation for discussing estimation, we describe the Laspeyres, Paasche, Geometric, Fisher, and Törnqvist indexes. Assume that we have a finite population U of N items and that the prices of the items and the quantities purchased at some time t are, respectively,

$$p_{t1},\ldots,p_{tN} \text{ and } q_{t1},\ldots,q_{tN}.$$

The Laspeyres index of change between period s and a later time t is

$$L_{t,s} = \frac{\sum_{i\in U} p_{ti}q_{si}}{\sum_{i\in U} p_{si}q_{si}} = \sum_{i\in U} w_{si}r_{tsi}$$

where $r_{tsi} = p_{ti}/p_{si}$ is the price relative between times s and t and $w_{si} = p_{si}q_{si}/\sum_{i\in U} p_{si}q_{si}$ is the share of expenditures due to item i during period s. The standard application of a Laspeyres index sets s to be some base period denoted by 0.

The Paasche index for the change between s and t is defined to be

$$P_{t,s} = \frac{\sum_{i\in U} p_{ti}q_{ti}}{\sum_{i\in U} p_{si}q_{ti}} = \sum_{i\in U} w^*_{si}r_{tsi}$$

where $w^*_{si} = p_{si}q_{ti}/\sum_{i\in U} p_{si}q_{ti}$ is an expenditure share with current period quantities evaluated at base period prices.

The geometric mean or "geomean" index is equal to

$$G_{t,s} = \prod_{i\in U}(r_{tsi})^{w_i}$$

where the w_i are a fixed set of weights that sum to 1. A typical application of the geomean would use weights that are expenditure shares during some time period, e.g., $w_{si} = p_{si}q_{si}/\sum_{i\in U} p_{si}q_{si}$. If the weights are fixed and not dependent on a time period, then

$G_{t,s}$ satisfies five axioms on price indexes in Balk (1995).

Indexes that pass most of the tests listed by Fisher (1922) and Diewert (1987) are the Fisher and Törnqvist. The Fisher index is the geometric mean of a Laspeyres and a Paasche index:

$$F_{t,s} = \left(L_{t,s} P_{t,s}\right)^{1/2}$$

while the Törnqvist is

$$T_{t,s} = \prod_{i \in U} \left(r_{tsi}\right)^{\bar{w}_{tsi}}$$

with $\bar{w}_{tsi} = \left(w_{si} + w_{ti}\right)/2$, i.e., the mean of the expenditure shares at times s and t.

Note that each of these indexes, as formulated above, assumes that the same set of items is available for pricing in the two time periods being compared. In a dynamic economy this is unrealistic since the types of items available for purchase may change rapidly. Electronics and computer equipment are extreme examples, but other commodities like women's apparel also undergo enough change that pricing the same item for extended periods of time is impossible. Practical work-arounds are, thus, required when any of these indexes is implemented in practice.

3. Estimation of the Indexes

The indexes defined in the last section are population values that must be estimated from samples. How complicated the sample design must be depends on the type of commodity and on what sorts of lists are available for sampling. In a large country like the United States, it may be necessary to select the sample in several stages in order to identify items whose characteristics are specific enough that the items can be priced over a number of time periods. For pricing most commodities other than housing, the U.S. Bureau of Labor Statistics (BLS), for example, samples geographic areas, retail outlets, and items within the outlets for pricing. In this section we assume that a probability sample of items has been selected and review how estimators of the various indexes can be constructed, using auxiliary or explanatory data.

One of the standard and most flexible estimators in finite population sampling is the generalized regression (GREG) estimator (Särndal, Swensson, and Wretman 1992). The GREG allows auxiliary data to be easily incorporated into an estimator and is motivated by a linear model. Many standard estimators, including the Horvitz-Thompson and the post-stratified estimators, are special cases of the GREG. The types of auxiliary data available for index estimation are usually qualitative, e.g., region of the country, type of retail outlet that sells the item (e.g., department store, discount store, grocery, etc.), season of the year, type of item, (food, clothing, electronics, etc.).

Since each of the indexes described in section 2 can be expressed in terms of price relatives or their logs, it is natural to attempt to frame models that describe the behavior of one or the other. Let y_{tsi} be either r_{tsi} or $\ln(r_{tsi})$. Thus, the Laspeyres, Paasche, the components of the Fisher index, and the logs of the Geometric and Törnqvist indexes can all be expressed in the form $\sum_{i \in U} \tilde{w}_{tsi} y_{tsi}$ for appropriate definitions of $\tilde{w}_{tsi}$ and y_{tsi}.

Consider the working model

$$y_{tsi} = \mathbf{x}_{ti}' \boldsymbol{\beta}_t + \varepsilon_{ti} \tag{1}$$

where $\mathbf{x}_{ti}$ is a $p \times 1$ vector of explanatory or auxiliary variables, $\boldsymbol{\beta}_t$ is a $p \times 1$ parameter vector, and the errors have zero mean. Errors for different items may be correlated, but detailed specification of the variance/covariance structure is not needed for the discussion in this section. Empirical research by many authors suggests that the logs of the price relatives are often nearly normally distributed than the price relatives themselves and are usually the preferred function for modeling. However, for indexes like the Laspeyres and Paasche, modeling price relatives is clearly most convenient.

Suppose that a probability sample of n items is selected with the inclusion probability of item i being π_i. Let $\mathbf{X}_{tN}$ be the $N \times p$ population matrix of explanatory variables at time t and $\mathbf{X}_{tn}$ be the corresponding matrix for the sample items. Taking the weights, $\{\tilde{w}_{tsi}\}_{i \in U}$, as known, the GREG estimator of $\theta_{t,s} = \sum_{i \in U} \tilde{w}_{tsi} y_{tsi}$ is

$$\hat{\theta}_{t,s} = \sum_{i \in s} \tilde{w}_{tsi} \frac{y_{tsi}}{\pi_i} + \left(\tilde{\mathbf{w}}_N' \mathbf{X}_{tN} - \tilde{\mathbf{w}}_n' \Pi^{-1} \mathbf{X}_{tn}\right) \mathbf{A}_{\pi n}^{-1} \sum_{i \in s} \mathbf{x}_{ti} \frac{y_{tsi}}{\pi_i} \tag{2}$$

where $\tilde{\mathbf{w}}_N$ is the vector of weights for items in the population, $\tilde{\mathbf{w}}_n$ is the vector for the sample items, $\Pi = diag(\pi_i)$, and $\mathbf{A}_{\pi n} = \mathbf{X}_{tn}' \Pi^{-1} \mathbf{X}_{tn}$. The g-weights associated with the GREG in this case have the form

$$g_{tsi} = \tilde{w}_{tsi} + \left(\tilde{\mathbf{w}}_N' \mathbf{X}_{tN} - \tilde{\mathbf{w}}_n' \Pi^{-1} \mathbf{X}_{tn}\right) \mathbf{A}_{\pi n}^{-1} \mathbf{x}_{ti} \tag{3}$$

and the GREG can be written as

$$\hat{\theta}_{t,s} = \sum_{i \in s} g_{tsi} \frac{y_{tsi}}{\pi_i}.$$

If model (1) holds, then $E_M\left(\mathbf{A}_{\pi n}^{-1} \sum_{i \in s} \mathbf{x}_{ti}\, y_{tsi}/\pi_i\right) = \boldsymbol{\beta}_t$ and $E_M\left(\hat{\theta}_{t,s} - \theta_{t,s}\right) = 0$, i.e., the GREG is model-unbiased. Also, whether the model holds or not, as $n \to \infty$, if $\mathbf{A}_{\pi n}^{-1} \sum_{i \in s} \mathbf{x}_{ti}\, y_{tsi}/\pi_i$ is bounded, then the GREG is design-consistent and approximately design-unbiased. Having both model-based and design-based justification is desirable for indexes produced by government programs.

Thus, more specifically, if linear model (1) holds for price relatives, then $\hat{\theta}_{t,s}$ is a reasonable estimator for the Laspeyres, Paasche, and the components of the Fisher indexes. If a linear model holds for logs of price relatives, then appropriate estimators of the log of the Geometric index and the log of the Törnqvist index can be constructed with the GREG.

A key issue, of course, is whether the GREG is practical since it involves the weighted universe total, $\tilde{\mathbf{w}}_N' \mathbf{X}_{tN}$. For quantitative explanatory variables, such weighted totals may have no interpretation and would not be available. More plausible choices of auxiliaries may be stratifiers that identify the type of outlet in which the item is sold, the location of the outlet (urban, suburban, or rural), or the region of the country. The regression basis of the GREG also permits interactions of main effects to be easily incorporated, which would simply be ways of defining domains if the main effects in the model are stratifiers. In these cases, $\tilde{\mathbf{w}}_N' \mathbf{X}_{tN}$ consists of expenditure totals for various domains and can be estimated from a separate expenditure survey of households.

Calculation of the sample expenditure estimate, $\tilde{\mathbf{w}}_n' \Pi^{-1} \mathbf{X}_{tn}$, is also an issue because the values of $\tilde{w}_{tsi}$ may be unknown for individual items—even those in the sample. When estimating a Laspeyres index, the U.S. BLS deals with the problem by selecting items with probabilities proportional to expenditure. It is assumed that the probabilities are also proportional to the desired values of $\tilde{w}_{tsi} = w_{0i}$ (defined in section 2), so that when $\tilde{w}_{tsi}/\pi_i$ is computed, certain unknowns cancel (see Leaver and Valliant 1995, pp. 549-550).

4. Combining Data Over Time

Estimating change over time is the primary goal of price indexes, and both long-term change and short-term change are important. By long-term change, we mean the change from a specified base period to the current time. The base period is often 10 years or more in the past. Short-term changes are for intermediate periods like a month, a quarter, or a year. One of the cosmetic goals of some index programs is to have the short-term changes compatible with the long-term change. Multiplying the annual changes together for the last 10 years produces the published change for the whole 10 year period, for example.

One method of combining estimates of change over time by directly multiplying the estimators of short-term change to estimate long-term changes. Let $I_{s,0}$ be an index of change from the base period 0 to a later time s. We have at least two alternatives for defining a 1-period price change from time s to time s+1. One is by using the price relatives, $r_{s+1,s,i}$ in the different index formulas in section 2 to obtain an index we can denote as $I_{s+1,s}$. The other is to take the ratio of the time s+1 long-term change to the time s change:

$$\tilde{I}_{s+1,s} = \frac{I_{s+1,0}}{I_{s,0}}. \qquad (4)$$

These 1-period changes can be multiplied together to obtain a long-term change as

$$\tilde{I}_{t,0} = \prod_{s=0}^{t-1} \tilde{I}_{s+1,s} = \prod_{s=0}^{t-1} \frac{I_{s+1,0}}{I_{s,0}} \qquad (5)$$

where, by convention, we set $I_{0,0} = 1$. Because of serial cancellation in (5), the far right-hand side reduces exactly to $I_{t,0}$. The construction in (5) can be used for any of the indexes defined in section 2.

By definition in (5), the 1-period short-term changes, $\tilde{I}_{s+1,s}$, are compatible with the long-term change, $\tilde{I}_{t,0}$, in the sense that that the product of the short-term changes equals the long-term change. In contrast, if $I_{s+1,s}$ is the 1-period change, then $I_{t,0} \neq \prod_{s=0}^{t-1} I_{s+1,s}$, in general.

For example, with the Laspeyres index and t=2, we have

$$I_{2,0} = \sum w_{0i} r_{20i}$$
$$\neq \left(\sum w_{0i} r_{10i}\right)\left(\sum w_{1i} r_{21i}\right)$$
$$= I_{1,0} I_{2,1},$$

i.e., the fixed base index $I_{2,0}$ differs from the chained Laspeyres index, $I_{1,0} I_{2,1}$.

On the other hand, with the Geometric index we might define the short-term change as either $G_{s+1,s} = \prod_{i \in U} \left(r_{s+1,s,i}\right)^{w_i}$ or $\tilde{G}_{s+1,s} = G_{s+1,0} / G_{s,0}$ with $G_{u,0}$ (u = s+1 or s) defined by the standard geometric formula using the long-term relatives $r_{u,0,i}$. Expression (5) is then

$$\tilde{G}_{t,0} = \prod_{s=0}^{t-1} \tilde{G}_{s+1,s} = \prod_{s=0}^{t-1} \frac{\prod_{i \in U} \left(r_{s+1,0,i}\right)^{w_i}}{\prod_{i \in U} \left(r_{s,0,i}\right)^{w_i}} = \prod_{i \in U} \left(r_{t,0,i}\right)^{w_i}. \quad (6)$$

The other alternative for the long-term change is

$$G_{t,0} = \prod_{s=0}^{t-1} \prod_{i \in U} \left(r_{s+1,s,i}\right)^{w_i}. \quad (7)$$

If the universe is constant so that the products over time and items can be interchanged, then serial cancellation leads to expression (7) equaling (6). With a changing universe, however, this equality would not hold.

When population indexes are replaced by estimators in (5), the serial cancellation feature will usually be lost because of changes in the sample across time. The long-term estimate will, however, still be compatible with the short-term estimates if we combine them using formula (5).

How much different countries value this "compatibility" varies. Sweden, for example, allows the product of monthly changes within a year to be different than the annual change. But, for publishing changes across several years, Sweden does multiply the estimates of annual change (Dalén 1992).

Let $\hat{I}_s(u)$ denote an estimator of the long-term index $I_{s,0}$ based on the sample at time u. An estimator of the index in (5) is then

$$\hat{\tilde{I}}_t = \prod_{s=0}^{t-1} \frac{\hat{I}_{s+1}(s+1)}{\hat{I}_s(s+1)} = \hat{I}_t(t) \prod_{s=1}^{t-1} \frac{\hat{I}_s(s)}{\hat{I}_s(s+1)}. \quad (8)$$

Expression (8) is sometimes referred to as a product estimator and is the type used by the BLS in estimating Laspeyres indexes in the Consumer Price Index, the Producer Price Index, and the International Price Index.

A poor statistical feature of an estimator constructed by substituting estimators into (5) is that its variance tends to increase over time as more and more estimators are chained together. In the context of Laspeyres estimators, Valliant and Miller (1989, denoted as V&M below) noted that both the variance and the relvariance of (8) are increasing as long as prices are rising. Leaver (1990) empirically confirmed this phenomenon with U.S. CPI estimates. The problem of increasing variance affects product estimators more generally (see, e.g., Hansen, Hurwitz, and Madow 1953, sec. 11.7). We can also expect the phenomenon of increasing variance to occur when (8) is applied with Geometric, Törnqvist, or the other indexes noted earlier.

The form of (8) does suggest an alternative formulation that will have better variance properties. The numerator and denominator of the ratio $\hat{I}_s(s) / \hat{I}_s(s+1)$ both estimate the long-term index $I_{s,0}$ based on the samples at two time periods. The ratio is, thus, an estimator of the constant 1. An obvious modification of (8) is then

$$\hat{\tilde{I}}_{t\gamma} = \hat{I}_t(t) \prod_{s=1}^{t-1} \left[\frac{\hat{I}_s(s)}{\hat{I}_s(s+1)}\right]^{\gamma_{ts}} \quad (9)$$

where γ_{ts} is selected optimally. A linear approximation to $\hat{\tilde{I}}_{t\gamma}$ is

$$\hat{\tilde{I}}_{t\gamma} \cong \hat{I}_t(t) + \sum_{s=1}^{t-1} \gamma_{ts} \tilde{I}_{t,s} \left[\hat{I}_s(s) - \hat{I}_s(s+1)\right] \quad (10)$$

where $\tilde{I}_{t,s} = I_{t,0} / I_{s,0}$ as in (5). Next, if we define

$$\mathbf{b}_t = \left[\gamma_{t1} \tilde{I}_{t,1}, \ldots, \gamma_{t,t-1} \tilde{I}_{t,t-1}\right]' \text{ and}$$
$$\mathbf{Z}_t = \left[\hat{I}_1(1) - \hat{I}_1(2), \ldots, \hat{I}_{t-1}(t-1) - \hat{I}_{t-1}(t)\right]',$$

then the linear approximation can be written as

$$\hat{\tilde{I}}_{t\gamma} \cong \hat{I}_t(t) + \mathbf{Z}_t' \mathbf{b}_t.$$

The variance of the linear approximation is

$$\operatorname{var}\left(\hat{\tilde{I}}_{t\gamma}\right) \cong \operatorname{var}\left(\hat{I}_t(t)\right) + \mathbf{b}_t' \operatorname{var}(\mathbf{Z}_t) \mathbf{b}_t + 2 \operatorname{cov}\left(\hat{I}_t(t), \mathbf{Z}_t\right) \mathbf{b}_t$$

where the variance can be taken with respect to either a model or a sample design.

The optimal vector $\mathbf{b}_t$ is the one that minimizes the approximate variance and is equal to

$$\mathbf{b}_{t,opt} = -\left[\mathrm{var}(\mathbf{Z}_t)\right]^{-1} \mathrm{cov}\left(\hat{I}_t(t), \mathbf{Z}_t\right)$$

as shown in V&M. If $\mathrm{cov}\left(\hat{I}_t(t), \hat{I}_s(s) - \hat{I}_s(s+1)\right)$ decreases as t and s move farther apart and the covariance is negative, then $b_{t,t-1} > b_{t,t-2} > \cdots > b_{t1}$ with $b_{ts} = \gamma_{ts}\tilde{I}_{t,s}$. From (10) this implies that the influence of past data on the current period estimate dies out. Under the simple autoregressive model

$$r_{tsi} = \mu_t + \varepsilon_{ti}, \quad \varepsilon_{ti} = \rho\varepsilon_{t-1,i} + u_{ti}$$

with $E(u_{ti}) = E(u_{ti}u_{t'i'}) = 0$, $E\left(u_{ti}^2\right) = \sigma_u^2$, and $|\rho| < 1$, the optimal value of γ_{ts} for estimating a Laspeyres index is

$$\gamma_{ts,opt} = \tfrac{1}{2}(\mu_s/\mu_t)\rho^{t-s}.$$

In times of inflation, $\mu_s < \mu_t$ and with ρ positive, the effect of earlier samples is damped out.

The product estimator is optimal in the restrictive case of $(\mu_t/\mu_s) = \frac{1}{2}\rho^{t-s}$ for all $s = 1,\ldots,t-1$. When ρ is positive, this condition implies that $(\mu_t/\mu_s) \le \frac{1}{2}$, an unrealistic case of severe deflation. Since $\gamma_{ts} = 1$ in the product estimator, we can see from (10) that the effect of old data does not diminish. In fact, the old data may become relatively more influential as t increases, because $\tilde{I}_{t,1} > \tilde{I}_{t,2} > \cdots > \tilde{I}_{t,t-1}$ with inflation.

Using expression (10), we can make a simple comparison to linear composite estimation. Suppose that at time t we have an estimator of a total denoted by $\hat{y}_t$. At time t+1 let the composite estimator of the total be

$$\hat{y}_t^* = (1-w)\hat{y}_t + w\left(\hat{y}_{t-1}^* + \Delta_{t,t-1}\right)$$

where $0 < w < 1$, $\hat{y}_t$ is a direct estimator based on the time t sample, $\hat{y}_{t-1}^*$ is the composite at time t-1, and $\Delta_{t,t-1}$ is an estimator of the difference between time t-1 and t based on the sample units common to the two times. Substituting recursively into this formula, we have

$$\hat{y}_t^* = \hat{y}_t + \sum_{s=1}^{t-1} w^{t-s}\left(\hat{y}_s^{**} - \hat{y}_s\right)$$

where $\hat{y}_s^{**} = \hat{y}_{s-1} + \Delta_{s,s-1}$. That is, we can write the composite as the current period estimator $\hat{y}_t$ plus an estimator of 0. Since $w < 1$ and is raised to the power w^{t-s}, the data from earlier time periods is progressively damped out, just as in the V&M optimal estimator. Thus, linear composite estimation acts in much the same way as the optimal long-term index estimator and the opposite of the way the product estimator does.

Similar analyses can be carried out for optimal short-term index estimators of the form $\hat{\tilde{I}}_{t_2\gamma}/\hat{\tilde{I}}_{t_1\gamma}$ where $t_1 < t_2$. For Laspeyres indexes, V&M found that the optima depend on the endpoints, t_1 and t_2, and are not practical choices.

Analyses of optimal index estimators for other indexes like the Paasche, Geometric, Fisher, and Törnqvist have not been conducted. Another open question is whether it is possible to structure short-term index estimators that are compatible with long-term estimators in ways other than using a product estimator.

5. Sample Design Issues

As in most surveys, sample design is critical to properly covering the universe and to producing estimates with acceptable precision. In this section, we discuss some of the statistical problems in variance component estimation and sample allocation. Proper coverage of the universe is, of course, a critical issue but is governed in large part by survey field operations. The lag in including new products was, in fact, one of the sources of bias in the U.S. CPI cited by Boskin, et.al. (1996).

An index estimation program usually has multiple goals that must be considered when allocating a sample of geographic areas, outlets, and items to be priced. An overall index across all types of items is usually published as are many subindexes for food, clothing, housing, medical expenses, and other groups of items. Indexes may also be published for regions of a country.

Obtaining an overall index that has a small variance may be quite important for budgetary reasons. According to Boskin, et.al. (1996), more than half of federal spending is attributable to entitlements and other mandatory spending programs that are indexed. The Congressional Budget Office (O'Neill 1995, Table 1) estimated that a hypothetical reduction in the CPI of 0.5 percentage points would result in a total contribution to federal deficit reduction of about $26.2 billion in fiscal year 2000. The reduction counts declines in federal outlays,

increases in revenues, and decline in debt service. In the years 1987-1991 the standard error of the estimate of 12-month price change in the CPI was about 0.144 percentage points (Leaver and Valliant 1995, Table 28.1). Thus, the width of a 95% confidence interval corresponds to roughly $30 billion.

The cost of living adjustment to federal programs is applied multiplicatively each year. Thus, over a period of years the cumulative, multiplicative adjustment is made using a product index. As noted in section 4, a product index generally has a relvariance that increases over time. Consequently, the dollar amount of budget increase due to indexing is more and more affected by statistical variation as time passes.

The estimator in (8) is a complex combination of products of ratios with a concomitantly complicated variance structure. In the U.S. CPI the all-items estimator is actually more complex since it is a weighted, linear combination of estimates like (8).

When allocating a sample, variance components are needed for the different stages of sampling that are used. The standard approach is to linearize the estimator and then compute the variance of the linear approximation in component form. As seen in (10) even the linear approximation is complicated since it involves data from t different time periods. In the U.S. CPI there are four stages of sampling:

- geographic primary sampling units
- entry level items (which are groups of similar items)
- outlets
- individual items within outlets

5.1 Variance Component Estimation

A linear approximation to the variance will have components due to each stage of sampling and coefficients on the components that depend on the particular form of Laspeyres index estimator that is used (Leaver, et.al. 1987). More generally, if another form of index, like the Geometric, is used, the linear approximation will be different as will the variance components, even if the stages of sampling are the same. The U.S. CPI now uses a Laspeyres index for Housing and geometric indexes for Commodities and Services. Thus, a different linearization is needed in the different parts of the index.

A simplification is to calculate the approximate, anticipated variance (AV), i.e., $E_\pi \operatorname{var}_M\left(\hat{\bar{I}}_t\right)$ where the subscript π denotes expectation with respect to a sample design and M is with respect to a model. Sampling with replacement may have to be assumed at some stages to make the derivation of the variance tractable. Valliant and Gentle (1997) used the AV approach in an allocation problem for estimating an index for wages. Because (8) or (9) uses sample data from t time periods, some strong assumptions about constancy of variance components across time and, possibly, on the rate of inflation are needed to make the variance formula tractable. Baskin and Johnson (1995), for example, used the following, random effects model for the price relative $r_{hijk} = p_{thijk}/p_{shijk}$ between times t and s for item k in primary sampling unit (*psu*) h, outlet i, and entry level item j:

$$r_{hijk} = \mu + \alpha_h + \beta_{hi} + \gamma_{hj} + \varepsilon_{hijk} .$$

We suppress subscripts for time to simplify the notation. The term μ is a mean common across all items. The random errors α_h, β_{hi}, γ_{hj}, and ε_{hijk} are assumed to have mean 0 and to be mutually independent. Separate models are needed within large groupings of commodities like food, apparel, housing, and transportation. An additional complication would be to model the correlation structure of the errors over time.

Various approaches can be used to estimate the variance components themselves. Baskin (1992, 1993) and Baskin and Johnson (1995) experimented with analysis of variance (ANOVA), hierarchical Bayes estimation using Gibbs sampling, and restricted maximum likelihood estimation (REML). General discussion of variance component estimation in hierarchical (or multilevel) models can be found in Goldstein (1995). The Bayesian approach is discussed in Gelfand and Smith (1990) and Gelfand, Hills, Racine-Poon, and Smith (1990). Baskin and Johnson (1995) encountered a number of problems, including negative ANOVA estimates. The Bayes estimates are guaranteed to be nonnegative, but in the CPI data set would not consistently converge and were often trapped at 0, even when 0 was not a reasonable value. Baskin and Johnson (1995) found that REML estimates performed the best but that the orders of magnitude of components were somewhat different from the ANOVA estimates that had been previously used in the CPI. The REML estimates of the individual-item component, in particular, were relatively larger than the ANOVA estimates. This

would lead to a larger sample of items allocated to individual outlet/ELI combinations, and, thus, has some important sample design implications.

5.2 Sample Allocation

To achieve the multiple goals of an index program, the allocation of the sample to *psu*'s, outlets, and groups of items must be a compromise between one that minimizes the variance of the overall index, meets precise goals for subindexes, and respects any cost constraints. Leaver, et.al. (1987, 1996) used constrained, nonlinear optimization methods for the U.S. CPI.

Creating an accurate cost function is a critical step in determining the sample allocation. In the U.S. CPI, there are costs associated with initiation data collection and processing, personal visit and telephone pricing, and data processing costs. To be useful, these must be decomposed into parts associated with the stages of sampling. Leaver, et.al. (1996) give extremely detailed breakdowns for costs, some of which are:

- outlet related costs of initiation (i.e., recruiting an outlet into the pricing survey), including field personnel and data processing.
- individual item related costs of initiation, including personnel time for sampling items onsite.
- on-going price data collection costs, including travel and personnel time distributed between personal visits and telephone pricing.

A key step is to set up a cost-tracking system that captures the quantities needed to develop the cost function. Generally, the costs will have to be recorded in more detail than needed to develop an annual survey budget.

A general optimization problem for index estimation can be stated in words as follows.

> *Find the sample sizes of psu's, outlets within psu's, ELI's within psu's, and items within ELI's that*
>
> - minimize $\mathrm{var}\left(\hat{\bar{I}}\right)$
> - subject to
>   - an upper bound on total cost
>   - upper bounds on the variances of a set of important subindexes
>   - lower bounds on the numbers of items selected from particular groups of important items
>   - workload constraints in each *psu*

This is just one of many ways in which the problem could be formulated. We might, for example, (a) minimize total costs subject to constraints on variances and workloads or (b) minimize a weighted combination of variances of indexes for different groups of items subject to cost and workload constraints.

6. Areas for Research

Index estimation is a fertile area for both economic and statistical research. We reiterate a few of the possibilities here that were noted in earlier sections and introduce some additional topics. Some of these issues require both economic and statistical thinking to resolve. Others require clever survey operations to handle. Areas previously noted are:

- **Variance component estimation**
- **Cost function construction**: This may be facilitated by on-line cost accounting systems that allow field office personnel to record expenses in new ways.
- **Sample allocation optimization**

Other areas not previously mentioned are:

- **Model construction for use in estimators**: Some items have price structures that are determined nationally—Internet sales, catalog and phone sales, for example. Other items have regional and/or seasonal variations—fruits, vegetables, gasoline. Prices also have a complicated time series structure. Local economic conditions—unemployment rate, wage & salary levels—also affect price changes. How to account for this structure in constructing index estimators is an open question.
- **Model bias of estimators under model-misspecification:** Any estimator can be analyzed with respect to an underlying model whether one favors design-based or model-based inference. The models that underlie price populations are likely to be extremely complicated. Although accounting for this structure in estimation may not be feasible, model-based analysis may pinpoint characteristics of samples (e.g., balance on certain variables) that are important in producing model-unbiasedness. Different types of highly controlled

systematic samples (see, e.g., Dorfman and Valliant 2000) may approximately achieve such balanced samples.

- **Universe coverage and frame construction:** Internet sales have created new and difficult problems of universe coverage. As this type of transaction begins to occupy a larger share of the economy, coverage of on-line auction sales and of other items with variable on-line, transaction prices becomes more important. Retail purchases from off-shore vendors will also become more prevalent.
- **Folding in prices for new items:** An item often becomes unavailable for pricing because its specifications change so radically that it cannot be considered as the same item between two periods. Computer equipment is an obvious example. How best to introduce these new items is an issue because the new item may involve a large price change. When a prescription drug goes off patent, for example, its generic equivalent may be sold at a far lower price.
- **Estimation of aggregation weights:** Estimates like (8) and (9) are usually made for classes of items and then weighted together. The aggregation weights may be estimated from a consumer expenditure survey conducted independently from the pricing survey. Ghosh and Sohn (1990) and Lahiri and Wang (1992) investigated the use of empirical and hierarchical Bayes methods for improving estimates of aggregation weights. These methods were promising but have not been implemented in a national index program.
- **Quality Adjustment using Models:** The standard econometric procedure, known as hedonic regression, for doing this is to predict the price of an item based on its characteristics and then compare that to the actual price. Improvements in current practice are needed. In fact, Moynihan (1999) in his Presidential Address to the 1998 Joint Statistical Meetings felt that this was one of the most important statistical issues facing the U.S. government.

REFERENCES

Balk, B. (1995), "Axiomatic Price Index Theory: A Survey, *International Statistical Review*, **63**, 69-93.

Baskin, R. (1992), "Hierarchical Bayes Estimation of Variance Components for the U.S. Consumer Price Index," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 716-720.

Baskin, R. (1993), "Estimation of Variance Components for the U.S. Consumer Price Index via Gibbs Sampling," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 808-813.

Baskin, R. and Johnson, W. (1995), "Estimation of Variance Components for the U.S. Consumer Price Index," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 126-131.

Boskin, M., Dulberger, E., Griliches, Z., and Jorgensen, D. (1996), "Toward a More Accurate Measure of the Cost of Living," *Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index*, Washington DC.

Dalén, J. (1992), "Computing Elementary Aggregates in the Swedish Consumer Price Index," *Journal of Official Statistics*, **8**, 129-147.

Diewert, W.E. (1987), "Index Numbers," in *The New Palgrave: A Dictionary of Economics*, Vol. 2, London: Macmillan, 767-780.

Dorfman, A. (1998), "Price Indexes as Quasi-Longitudinal Studies," *Survey Methodology*, **24**, 139-145.

Dorfman, A. and Valliant, R. (2000), "Stratification by Size Revisited," *Journal of Official Statistics*, **16**, in press.

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

Gelfand, A.E., Hills, S.E., Racine-Poon, A., and Smith, A.F.M. (1990), "Illustrations of Bayesian Inference in Normal Data Models Using Gibbs Sampling," *Journal of the American Statistical Association*, **85**, 972-985.

Gelfand, A.E. and Smith, A.F.M. (1990), "Sampling Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, **85**, 398-409.

Ghosh, M. and Sohn, S. Y. (1990), "An Empirical Bayes Approach Towards Composite Estimation of Consumer Expenditure," unpublished report, *Washington DC: Bureau of Labor Statistics.*

Lahiri, P. and Wang W. (1992), "A Multivariate Procedure Towards Composite Estimation of Consumer Expenditure for the U.S. Consumer

Price Index Numbers," *Survey Methodology*, **18**, 279-292.

Goldstein, H. (1995), *Multilevel Statistical Models*, 2nd edition, London: John Wiley.

Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953), *Sample Survey Methods and Theory*, Vol. I, New York: John Wiley & Sons.

Leaver, S.G. (1990), "Estimating Variances for the U.S. Consumer Price Index for 1978-1986," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 290-295.

Leaver, S.G., Johnson, W.H., Baskin, R., Scarlett, S., Morse, R. (1996), "Commodities and Services Sample Redesign for the 1998 Consumer Price Index Revision," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 239-244.

Leaver, S.G. and Valliant, R. (1995), "Statistical Problems in Estimating the U.S. Consumer Price Index," Chapter 28, pp. 543-566, in *Business Survey Methods*, B. Cox, D. Binder, N. Chinnappa, A. Christianson, M. Colledge, P. Kott eds., New York, John Wiley.

Leaver, S.G., Weber, W.L., Cohen, M.P., and Archer, K.P. (1987), "Item-Outlet Sample Redesign for the 1987 U.S. Consumer Price Index Revision," *Proceedings of the 46th Session*, Vol. LII, Book 3, International Statistical Institute, 173-185.

Moynihan, D.P. (1999), "Data and Dogma in Public Policy," *Journal of the American Statistical Association*, **94,** 359-364.

O'Neill, J. (1995), "Prepared Statement on the Consumer Price Index," Hearings before the Committee on Finance, U.S. Senate, 194th Congress, First Session, Washington DC: U.S. GPO.

Särndal, C.-E., Swensson, B., and Wretman, J. (1992), Model Assisted Survey Sampling, New York: Springer-Verlag.

Valliant, R. and Gentle, J. (1997), "An Application Of Mathematical Programming to Sample Allocation," *Computational Statistics and Data Analysis*, 337-360.

Valliant, R. and Miller, S.M., (1989), "A Class of Multiplicative Estimators of Laspeyres Price Indexes," *Journal of Business and Economic Statistics*, **7**, 387-394.