Consumer Price Indices: Purposes and definitions.

A useful general approach to the construction of the indices and some measures of the divergences between them

Luigi Biggeri
Department of Statistics, University of Florence, Italy
National Committee for the Guarantee of Statistical Information, Rome, Italy

Abstract. During the recent debate on the relevance and reliability of the Consumer Price Indices (CPIs) constructed by the National Statistical Offices, the discussions and controversies on the sources of possible bias in a specific CPI originated mostly from the fact that the discussants often sought a specific «true» or «unique» formulation for the index, to be utilised for all kind of different purposes and uses.

The paper first recalls the strict relationships existing among purpose, definition, formula and actual computation of the price index numbers both from a theoretical and a practical point of view, remarking that different purposes will require different indices, to be chosen on a case by case basis. Then, the paper gives a small contribution to solve some practical issues, focusing on a general approach for the construction of CPIs that permits, case by case and in a coherent way, the choice of the most suitable formulas and giving some suggestions on simple methods to evaluate the differences (bias) between the computed index and the desired or the most satisfactory one.

1. Introduction

More then 10 years ago we stated that it is extremely important to focus our attention on the practical aspects of the construction of price indices and, especially, on the evaluation of the CPI reliability (Biggeri and Giommi, 1987). Since then, many important contributions have been published, many conferences and seminars on the subject have been organised, extending the dispute to the quality of CPIs as well (see for example, Turvey, 1989; US Bureau of Labor Statistics, 1993; Statistics Canada, 1994; EUROSTAT, 1996; Balk, 1997 and the works quoted therein).

A fruitful debate on the relevance and reliability of the Consumer Price Indices (CPIs) constructed by National Statistical Offices, and on the evaluation of the possible errors exhibited by them, is underway (see for example the so-called Boskin report (US Senate Commission, 1996) and Diewert, 1996). However, many discussions and controversies on the sources of possible bias in a specific CPI are mostly originated because the discussants often sought a specific «true» or «unique» formulation for the index to be utilised for all the different purposes and uses and compared this theoretical index with the computed one.

Instead, it is well known that the published consumer price indices often have different meanings and are useful for different purposes, taking into account also that there many potential requirements and different situation and data availability. And besides, there is not any "ideal" consumer price index valid for all the purposes, because the different theories and approaches to the construction of price indices no one can give alone the right reply to all the purposes for which the CPIs are used. Moreover, especially from a practical point of view, we have to consider that the CPI computation is not a direct statistical estimation but a very complex construction and we have to find coherent procedures to do it with reference to its purpose. Finally, it is obvious that we must try to find the most appropriate measure of the concepts that the economic theory is asking to measure, but at the following conditions: (i) the request and the purpose of the measure must be clearly defined; and we have to clearly reply (ii) what is possible actually to measure with enough accuracy, and (iii) which is the validity of what we have measured, providing information and at least partial measures of possible biases or of the validity errors.

Therefore, we feel that, taking into account the main points of the current debate, the subject of the purposes and definitions of CPIs are worth a further visit; in particular we intend:

a) first, to recall the different approaches to the construction of price index numbers to show the reasons why an «ideal» all-purpose consumer price index is quite impossible to justify both from a theoretical and a practical point of view (this will be presented in the Section 2);

b) second, to present a general practical framework for the construction of a price index to remark that the purpose of the index affects both the underlying concepts and definitions and the process (or procedures) of construction (Section3);

c) third, to recall the main purposes and definitions used to construct the consumer price indices by the National Statistical Offices and to illustrate an approach to the construction of price index numbers, the so-called «equivalence conditions» approach, that permits, case by case and in a coherent way, the choice of the most suitable formulas and of the different steps in the production process of a specific CPI (Section 4);

d) finally, to present a framework for all the aspects of the accuracy and validity of a price index and to give small contribution to solve practical issues specifying some partial measures of the divergence (bias imputed to the validity errors) between the actually calculated index and the desired or the most satisfactory one (Section 5).

2. The different approaches to the construction of price index numbers: different indices for different purposes

In an invited paper presented at the XXXII Scientific Meeting of the Italian Statistical Society (Biggeri, 1984), we reviewed the different approaches for the construction of price index numbers with particular reference to consumer price indices (for other reviews see for example: Allen, 1975 and Diewert, 1995). Our review of the *statistical* (or classical) approach, where prices and quantities are generally considered independent variables, and of the *economic* (or functional) approach, in which we assume certain relations between these variables, allowed, among other things, for the clarification of some concepts and the generalisation of some of the approaches to the subject that will be useful to correctly appraise the different views of today's debate. We also showed that the indices obtained from the different approaches are sometimes identical, and above all we underlined the fact that there is not any "ideal" consumer price index valid for all the purposes and every procedure or formula satisfies particular needs, so that the validity and the choice of the different index numbers can be judged only with reference to the purposes for which they are used. It follows, in our opinion, that we must rely on a general setting that will permit the proper choices for the particular problem we are dealing with and for the data at hand, stressing in particular the necessity and possibility of integrating the economic and the statistical theory of index numbers and, in any case, the use of the theory essentially as a guide for the construction of CPI. Let us summarise here the main results of our updated analysis.

2.1. The «statistical» approach.

Let us assume we want to measure, by means of a synthetic index ${}_0P_t$, the variation of a vector of prices between time *0* and time *t* (binary comparison). Let us assume, furthermore, that, with reference to the purpose for which we want to make such a measurement, there are two correctly defined price vectors, $(p_{10}, \ldots, p_{k0}, \ldots, p_{n0})$ and $(p_{1t}, \ldots, p_{kt}, \ldots, p_{nt})$, of n goods and services at time *0* and at time *t* and that the two vectors are technically comparable, that is, the characteristics of the prices with regard to the type and the quality of the goods, to the place of exchange, to the market type, etc., are the same at time *0* and at time *t*. Let us suppose, finally, that we have correctly observed the two price vectors or that we hold the elementary price indices $({}_0P_{1t}, \ldots, {}_0P_{kt}, \ldots, {}_0P_{nt})$, where ${}_0P_{kt} = p_{kt}/p_{k0}$.

As is well known, the problem of determining a synthetic index ${}_0P_t$ can be solved either by utilising some «aggregation» function of the prices, or of the elementary indices, in the two periods, in the following manner:

$${}_0P_t = F(p_{1t}, \ldots, p_{kt}, \ldots, p_{nt}) / G(p_{10}, \ldots, p_{k0}, \ldots, p_{n0}) \quad (1)$$

$${}_0P_t = H({}_0P_{1t}, \ldots, {}_0P_{kt}, \ldots, {}_0P_{nt}) \quad (2)$$

The problem is then solved by the choice of the functions *F* and *G* or *H* and by the derived expressions. Such a choice is certainly conditioned by the hypothesis we make regarding the characteristics of the price vectors and by the purpose we have in mind in constructing the synthetic index, since we must be able to give it a precise economic meaning.

In relation to the suggested hypothesis and formulas, a distinction is usually made between the so-called «stochastic» (or atomistic) approach, or with non weighted means and the so-called «aggregative» approach, or with weighted means.

The first line of thought has as its main objective, the study of the variation of the general price level, or of the value of money, and is based on the hypothesis that a unique, common, monetary cause affects, proportionally and in the same direction, the prices of all goods and services. It considers, instead, on the whole ineffective the market causes of price variations of single commodities. The deviations of the relative price variations of single commodities from the general variation (mean) due to the monetary cause, are considered observational errors and, hence, the distribution of the elementary indices ${}_0P_{kt}$ is hypothesised normal (measurement errors) or lognormal (evaluation errors). The synthetic index is then given by the arithmetic mean or the geometric mean of the elementary indices.

The objective of the aggregative school (which became popular in the second half of the nineteenth century with the well-known Laspeyres' and Paasche's formulas, but which can be dated back to Lowe in 1822) is to measure the variation of cost (expenditure) due to the price variations of a specific aggregate (basket) of n goods and services referred to a specific group of economic agents (for example, expenditures of a particular category of consumers or producers and/or for particular categories of goods and services), aggregate defined by a vector of quantities (q_1,...., q_k ,....,q_n).

Indices of this type are usually employed to measure the purchase power of the income of a particular group of families or of other monetary flows such as pensions, rents etc., or to measure the variation of the monetary value of an aggregate of goods sold by a term contract. In any case, in this approach the synthetic index is generally obtained as the ratio between the values of aggregates (expenditures), real and/or assumed**.** This means that we implicitly attach a different importance to the price variations of the various goods and services, and in fact we can always obtain the same index by using a weighted arithmetic mean of elementary indices, having chosen suitable weights.

Following the two approaches, it is possible to obtain the main formulas of synthesis proposed by the various authors, from Dutot, to Laspeyeres and Paasche and Tornqvist. Anyway, we feel it is important to make two general considerations.

First of all, it should be recalled that the stochastic (monetary) approach was strongly criticised (Gini, 1924; Keynes, 1930) because it does not seem to allow for an economic definition of the general level of prices, thought to be an abstract concept or, as Keynes asserts, an «inconceivable» concept. In any case, even if it is defined as the marginal utility of money, the hypothesis of independence of marginal utilities of single goods, that is implicit in this approach, would be unacceptable from an economics point of view. Furthermore, the statistical observation of the real world excludes the hypothesis of normality for the distribution of the elementary indices, or of their logarithm, that is the basis of this approach. On the other hand, the indices obtained by the aggregative school lack a direct reference to economic theory, but they generally have a simple economic interpretation though their meaning is often too narrow (cost variation of a fixed basket of goods). This is the reason why these indices, in particular those of the Laspeyres type, continue in practice to be the most used today.

Secondly, it should be noticed that the stochastic indices can be obtained by means of a procedure which is similar to that which is generally used for finding expected values (Parenti, 1948), imposing approximation conditions (of least squares). In fact, generalising the works of Bowley (1928) and Balk (1980), any synthetic index can be obtained through a simple linear regression model, of the form

$$g({}_0P_{kt}) = g({}_0P_t) + e_k \qquad (k = 1,........,n) \qquad (3)$$

making suitable hypotheses on the variance of the random variable e_k, in addition to the usual $E(e_k) = 0$. If, for example, we assume $g(x) = x$ and $Var(e_k) = \sigma^2$, the estimation of ${}_0P_t$ is formally obtained as an arithmetic mean. If, with the same hypothesis regarding g, we assume $Var(e_k) = \sigma^2/w_{k,00}$ (where $w_{k,rs} = p_{kr}\ q_{ks}/\ \Sigma\ p_{kr}\ q_{ks}$ with r and s referring to time) we have a Laspeyres index. If we assume $g(x) = log(x)$, we get a geometric mean, and so on. In any case, to get the most part of known synthetic formulas, we have to assume that $Cov(e_k, e_j) = 0, (k, j = 1,.......,n)$, a hypothesis seldom acceptable from an economics point of view. Recently, this approach has been also further developed (Selvanthan and Rao, 1994). It is certainly interesting from a methodological point of view, since it enables us to estimate the variance of ${}_0P_t$ and, if the assumptions regarding the e_k are acceptable, it permits the determination of confidence intervals for the parameters and the application of the tests of hypothesis on index numbers. However, as also Diewert (1995) and other authors observed, the justifications presented for the variance assumptions in the new stochastic approaches are rather weak and are not consistent with the observed behaviour of prices. Moreover, the considerations made earlier for the first stochastic approach apply, for the most part, to the new stochastic approaches too and the fundamental critic of Keynes is still valid. Finally, from an empirical point of view, this approach does not provide elements which justify the choice of one formulation rather than another.

As is well known, various approaches have been proposed for the choice of the right formula to be used in an empirical problem, in particular the so-called «Fisher's tests» approach or the «assiomatic» approach (Eichhorn and Voeller, 1976). There is no doubt that a specific test may be a useful tool for judging the validity of an index to satisfy some specific purposes; unfortunately, it has been proved that the system of tests taken together, is inconsistent and that some of the tests in the system are in contrast with accepted economic theory. Moreover, following these approaches, one has to admit the existence of an «ideal» index which is to be preferred no matter what the purpose is and what empirical problem you are dealing with. From a purely mathematical point of view, the approach is certainly appealing, but too often it is in contrast with the empirical research requirements of the index numbers that, as stated by Allen (1975), «are essentially practical constructs».

2.2. The economic approach: the «true» cost of living indices

During the thirties many scholars, almost contemporaneously, felt the need to give an economic meaning to price index numbers which, in their opinion, could not be attached to aggregative indices (expenditure indices). These, in fact, assume that the «basket» of quantities (purchased, consumed, produced, etc.), with respect to which the price variation is measured, remains constant in the *0-t* interval and, hence, that the price variation, in the same interval, does not influence the quantities. This hypothesis, even if acceptable for very short time intervals, is certainly considered extreme by economic theory which normally assumes that the economic agent «reacts» to a price variation with a modification of the structure of quantities (purchased, consumed, produced, etc.).

Konus, regarded as the father of this approach, developed the analysis in the field of consumer economic theory and, therefore, in a consumer price context: his objective was the measurement of cost of living changes, that is the change in the cost necessary to maintain the same level of welfare measured by a welfare or utility function.
Assuming that the consumer has preferences over n goods and services, a general economic price index, Cost of Living Index, derived from this approach consists of the ratio between two expenditures born in different situations (of prices) but referred to the same level of utility. The (economic) index is, hence, obtained in the following manner:

$$ {}_0P_t^E = C\,(p_{1t},\dots, p_{kt},\dots, p_{nt}; U) / C\,(p_{10},\dots, p_{k0},\dots, p_{n0}; U) \qquad (4)$$

where C is a cost function and U the level of utility.

In order to compute the index, we have to assume, for a hypothetical consumer and for any quantities $(q_1,\dots,q_k,\dots,q_n)$ of n goods and services, the knowledge of his utility function, $U = u(q_1,\dots,q_k,\dots,q_n)$, the income R (expenditure) available to him for consumption, together with the system of prices $(p_1,\dots,p_k,\dots,p_n)$ prevailing at time 0 and at time t. The consumer will determine the quantities to be consumed by maximising his utility function for a fixed income or by minimising his expenditure R for a fixed level of utility. Hence, we can compute the expenditures at times 0 and t to reach the same level of utility, and use them in (4) to calculate the Cost of Living Index (COLI) at constant utility.

Obviously, the index will change if we change the time (and/or the position) at which the utility refers to. For example, if we refer to the utility level at time 0 (maximising the utility and spending all the income R_0) the following price index, called of Konus-Laspeyres, between time 0 and t, is obtained:

$$ {}_0P_t^{K,L.} = (\Sigma\, p_{kt}\, q^*_{kt\,(U0)}) / (\Sigma\, p_{k0}\, q_{k0}) \qquad (5)$$

where $q^*_{kt(U0)}$ are the estimated quantities that at time t give to the consumer the same utility of the time 0 with the minimum expenditure (if the utility level refers at the time t, a so-called Konus-Paasche index is obtained).

As is well known, this kind of price index takes into account the substitution effects in the quantities of goods and services caused by the relative prices changes.

However, there are some interpretative limitations due to the fact that the theoretical framework refers to a single consumer and to a fixed level of satisfaction. Actually, this means that, also in this approach, we require some simplifying assumptions, like, for example, that of invariance of tastes and real income, since their variation would undoubtedly bring about a change in the level of satisfaction (utility).

The most critical point, however, arises when the definition and the specification of the cost of living indices are extended to a group of individuals or to the entire population. Some authors (Allen, 1975), maintain that these extensions require an act of faith since, in order to attribute a meaning to them from the point of view of economic theory, we would need to hypothesise the existence of one or more «average» indifference surfaces. Now, there is no doubt that such extensions present some risks and that the meaning of the cost of living indices is more valid if limited to short time intervals and/or to groups of families with a low and relatively stable income, for which the hypothesis of homogeneous behaviour is more acceptable.

It is, however, from the empirical point of view, that we find the greatest difficulties in the construction of indices of the Konus type. The most delicate phase is that of choosing a suitable utility function (specified either directly or indirectly) since the lack of objective elements of knowledge makes this choice arbitrary, at least in part. In addition to this difficulty, one should consider the problems of estimating the vector of quantities and the scarcity of adequate statistical information.

On the other hand, the basis in economic theory for this approach can serve as a guide for empirical testing, permitting us to evaluate whether or not the hypotheses we make on consumer behaviour, specifying suitable utility functions, are able to explain the real behaviour at an aggregate level in a given historical period. If the answer is affirmative, the hypotheses may, at least temporarily, be accepted and we can attach a precise meaning to the derived cost of living index.

In relation to the practical difficulties encountered in the construction of Konus indices, there has been an attempt to verify whether the most well-known «statistical» indices could be considered approximate measures of the COLI or «true» indices and, at the same time, to analyse the relationships between the two types of indices. Research on this aspects, which tries to integrate the statistical and the economic approach to index numbers, produced many contributions on the so-called «exact» and «superlative» index numbers.

Konus himself, in his work of 1924, pointed out the relationships between the indices he proposed and the Laspeyres' and Paasche's indices, attempting to determine the bounds within which the «true» index should be found, reaching the conclusion that, in periods of substantial stability of consumption, the index should be well approximated by the Paasche index (from below) and by the Laspeyres index (from above).

Furthermore, Konus and many other authors showed that each «statistical» index, computed as a mean of elementary indices, corresponds to a demand function derived under specific hypotheses (at each statistical index corresponds a particular utility function). It follows that, under these hypotheses, statistical indices can be interpreted as economic indices and vice-versa, pointing out the strict connection between the two approaches. Various researches have also pointed out, however, that from different utility functions it is possible to obtain the same index and, vice-versa, different indices, of the Laspeyres and Paasche type, can be justified with the same utility function (the same consumer behaviour).

Diewert, and many other scholars after him, maintain that the problem of the choice of the right index number could be solved by defining the notion of «exact» index and of «superlative» index. Denoting by I^S the statistical index and by I^E the economic index based on a generic homothetic function f, if $I^S = I^E$, then I^S is called an «exact» index with respect to the chosen function f. Furthermore, if f is a function that provides a second order approximation (identical second order derivatives) of an arbitrary linear homogeneous utility function, I^S is called a «superlative» index with respect to f.

Actually, since many «superlative» indices exist, we could conclude that most of the statistical indices are in some way good approximations of the «true» index, but, if this were so, the problem of the choice of the index is then left undefined.

We are not surprised so much by the problem of indeterminacy in the choice of the index, as by the fact that someone still tries to find the mirage of the «true» or «unique» index, giving very little importance to the concrete economic meaning of the index and, above all, not worrying at all about verifying the empirical validity of the hypothesised behavioural models and, hence, of the proposed statistical tools.

Therefore, we could state that the difference among the different approaches is more at the conceptual than at the empirical level: the necessary integration of the economic and the statistical theory of index numbers and the empirical research ask for more attention and developments.

As a consequence of what we have said before, it is evident that the theory provides important and useful guide for the desired characteristics of the indices, but from a practical point of view, the validity and the choice of the different index numbers can be judged only with reference to the purposes for which they are computed and used, and therefore the choice must be made on a case by case basis, and we must be able to evaluate the accuracy and the validity of the computed indices.

3. A general framework for the construction of price indices with different purposes

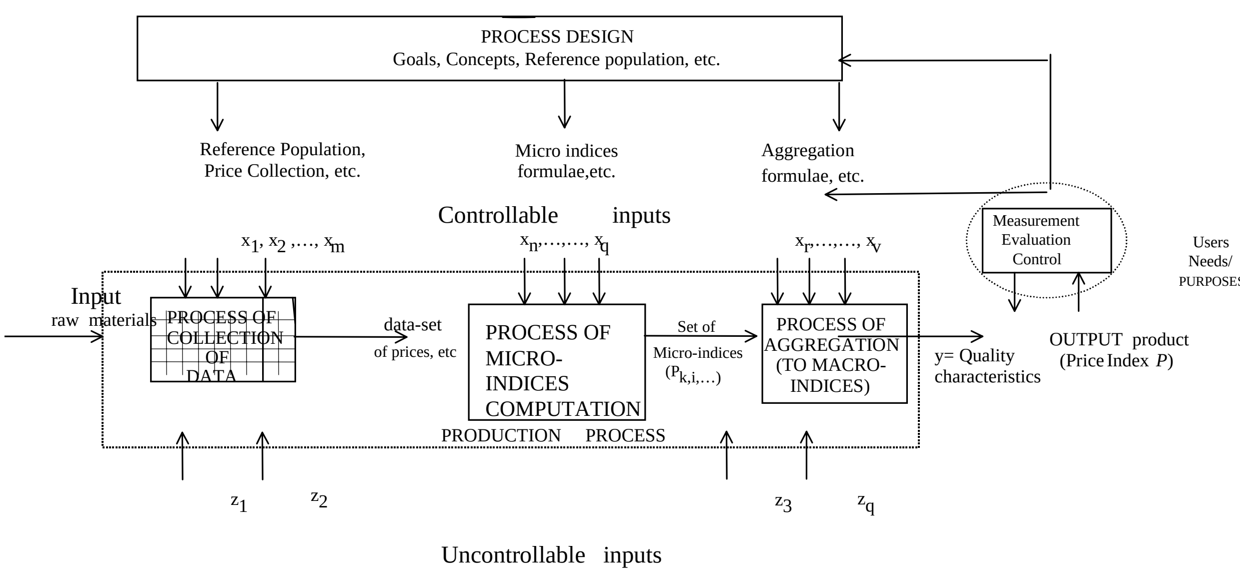
Given that prices indices are used for a wide variety of purposes - including adjustment of purchasing power of incomes, analysis of inflation, indexation by government, prices, wage and salary adjustment in contracts, derivation of estimated constant price values, international studies, and so on - both the choice of the most appropriate measure(s) for each use and the evaluation of the accuracy and validity of the computed indices must be made on a case by case basis.

In order to do it, in the construction of a specific index, we have to refer to a clearly defined production process (Biggeri, 1996) that starts from the *user needs* and should take into account the input process**.** The input process should consider the different steps and decisions – conceptual and operational - and the specific operations to be carried out for the construction of the price index *P* with specific quality characteristics.

Following the mentioned proposal, we presented here a general practical framework for the construction of a price index that permits, case by case and in a consistent way, the choice of the most suitable formulas and of the different alternatives in the different steps of the production process of a specific index and to take under control its quality. Looking at the framework it will be easy to remark that the purpose of the index affects both the underlying concepts and definitions and the process (or procedures) of construction.

In Fig. 1 a sketch for a very *simplified production process* of the price indices as a system of inputs and an output is presented.

Fig. 1. A simplified production process for price indices



Looking at the Fig. 1, starting from the side of the *inputs*, it is evident that in this process the raw material most important is the information on prices and expenditures and/or consumptions, for each commodity (item), that theoretically could be collected from the respondents. The cells inside of the left box represent the fact that these elementary information could be classified by type of commodity, distribution channel or point-of-purchase, territorial area, category of household that make the purchase or the consumption and so on.

The collection of data, carried out by a unique survey or more separate surveys, produces an *intermediate output*: the data-set, on prices and expenditures and/ or consumptions and/or weigths, that is used for the computation of the elementary or *micro-indices* P_k (where k denotes the generic modality of any classification of commodities, of point-of-purchase, etc.), as for the process denoted in the second box. Then, the set of micro-indices, the second intermediate output, is in its turn used in the aggregation process to arrive to the price index P (or price indices), the final output of the process.

The inputs or *process variables* (factors) can be considered the different steps and decisions - conceptual and operational - and the specific operations carried out for the complex construction of the micro-indices P_k and of price index P with specific *quality characteristics*. That is, the process inputs and design will therefore be dependent, for example, on:

- the definition of the *purpose* for which the index is constructed;
- the definition of the operational *concepts* to be measured, with reference to the specified purpose;
- the choice of the *reference population* (prices, commodities and services);
- the specification of the manner in which *prices* and *weights* are to be *collected* (purposive or probabilistic sampling of point-of-purchase and commodities);
- the specification of the treatment of *non-market services*, *seasonal commodities*, *special markets*, changes in the *quality of commodities*, etc.;
- the choice of the *formulae* for calculating the indices at micro and at aggregated levels and the system of weights, including the reference base (it is obvious that the choices could be equal or different at the two levels).

These phases and decisions are overlapping and influence each other, taking precise shapes also for the related operations only with reference to concrete cases.

In any case, we can, in general, consider some inputs x as *controllable* factors. For example in our case $x_1,\ldots,x_m$ could be the characteristics of the survey(s) carried out for the collection of data; $x_n,\ldots,x_q$ the characteristics of the computation formulae of the micro-indices and $x_r,\ldots,x_v$ the characteristics of the aggregation formulae. Other inputs $z_1,z_2,\ldots,z_q$ can be considered *uncontrollable* (or noise) factors; for example, environmental factors. For each *controllable* factor the producer can decide the *level* of the input and the combination of levels of different factors that he considers the optimum for his product. In the case of production of a price index, the levels of controllable factors can be considered the different choices that can be made with respect to the characteristics of its method of construction (e.g. the selection of a sample of points-of-purchase carried out using various different methods). It is therefore evident that with different decisions and methods of construction (the combination of the levels of the factors) the results obtained may be different.

The *output product*, specified by the variable *y*, has one or more observable *quality characteristic(s)* or response, which is a measure of the quality of the process. In our case, the response variable y is equal to the estimated index value *P* and, as for other economic statistical information, the required *quality characteristics* of a price index depend on the request of the users and on the purpose of the index, but in any case they should be at least: (i) real economic significance; (ii) relevance; (iii) timeliness; (iv) accuracy; and (v) accessibility.

It is evident that the outline of the process of producing price indices presented here is extremely simplified, useful for general analysis and observations. The different steps and operations necessary for the complete analysis as a Total Quality Management for the price indices production process – i.e. *Conceptualisation*, *Planning*, *Design*, *Development*, *Implementation*, and *Validation*, must be specified in projects or *sub-processes* that can be subdivided into activities fitting the plan, the design, and so on. Flowcharts illustrating the major processes and products of the operations will be very useful for describing the activities involved and for checking their consistency.

Anyway, as we have already said before, all the phases and decisions depend on the purpose of the index; they are often overlapping and influence each other, taking precise shapes for the related operations only with reference to concrete case of the construction of price indices for the different purposes.

4. The different purposes and definitions of the construction of Consumer Price Indices: the «equivalence conditions» approach

There are many recent good examples of the procedure followed by the National Statistical Offices for the construction of CPI, and usually they explain the main purpose for which the index is computed. The Australian Bureau of Statistics and the Statistics New Zealand, for example (Australian Bureau of Statistics, 1997b and 1997c; Cook and Lewington,1997), classify the various CPI uses (purposes) as follows:

- as a measure for income adjustment process;
- for general indexation of public and private sectors contracts;
- as a measure for the deflation of consumption in the national accounts;
- as a measure of inflation;
- as a measure of change in the cost of living

that are critical to any decisions on the methods of construction and the items to be included in the index, because no single index can adequately satisfy all the above mentioned purposes.

These offices also note that there are three alternative conceptual approaches to identify the basket of items to be used in the index (acquisition, cost of use and actual outlays (or payments) approaches), that everytime are most appropriate for the different purposes of the CPI. The suggestion of the two mentioned offices, together with many others (see EUROSTAT, 1996 and Astin and Sellwood, 1997), is to clearly state the principal (primary) purpose and produce an index strictly devoted to it. This would make the meaning and the methods of construction of the index more understandable and would avoid confusion and lack of confidence from the users.

To satisfy the purposes for which the compiled principal index is not adequate, we could then construct *satellite indices or sub-indices* (by groups of commodities and services, by groups of households, by territorial area, etc.) on the basis of the other purposes.

With reference to these issues we have to point out two general observations:

First, the choice of the most appropriate measure(s) for each use is, obviously, quite important and it should be done with reference to a system of price indices designed with reference to an analytical framework (see for example Biggeri, 1984; Biggeri and Giommi, 1987; Turvey, 1989 and Australian Bureau of Statistics, 1997a). The last paper presents a good proposal for the construction of a family of CPIs.

Second, due to ease of computation, timeliness and clear meaning, the most widely used formula for calculating CPIs are of the Laspeyres type; these have the following structure:

$$ {}_rP_t = \Sigma_k \, w_k \; {}_rP_{k,t} \qquad (6) $$

where k denotes the generic modality of any classification (for example of commodities), r $(r = l,..\,.,T)$ and t $(t = l,...,T)$ indicate, respectively, the base and the current period. These indices are then obtained by weighted arithmetical averages of relative or partial indices with weights w_k, such that $\Sigma_k \, w_k = 1$. Obviously, this kind of indices have different meanings depending on the choices and decisions mentioned above.

Because of theoretical and practical constraints, the *defined* , and even more so for the *computed*, index could loose some specificity that reduces its relevance for the designed purpose and use. For this reason, it is necessary to clarify the quality and relevance of the computed index and to give some, at least approximate, measures of divergence between it and the desired index, that is, to show how much it is a reasonable surrogate for the requested model.

Given the aim of this paper, we will not dwell on the specific choices and definitions made by the various statistical offices in constructing their CPIs, since they are moulded after very concrete cases (moreover, they will be taken up by others papers in this seminar). Our objective is, instead, to try an answer to the following questions: how is it possible to judge if the formula used to construct the index is suitable for the CPI purposes, taking into account the difficulties of the choices mentioned above?; how is it possible to make the subsequent decisions in a coherent way?

We feel that these objectives can be, in some way, achieved using the following general approach in the construction of price indices, the so-called «equivalence conditions» approach, that permits, case by case and in a coherent way with reference to the particular problem we are dealing with and for the data at hand, the choice of the most suitable formulas and of the different steps in the production process of a specific CPI.

The proposed procedure, which is similar to that which is generally used for finding suitable averages and other measure of location (Parenti, 1948), is based on «equivalence (invariance) conditions», such as:

$$F\,({}_rP_{1t},\dots,{}_rP_{kt},\dots,{}_rP_{nt}) = F\,({}_rP_t), \tag{7}$$

by which the average synthetic index is computed in such a way not to alter the measurement of the relative variation in the function *F* (invariant) with respect to which the judgement of equivalence is made.

If, for example, the function *F* is specified as $\Sigma\, g\,({}_rP_{k,t})\,a_k$, where the a_k are generic non negative weights and, if $g^{-1}(\cdot)$ represents the inverse function of $g(\cdot)$, the equivalence condition becomes

$$\Sigma\, g\,({}_rP_{k,t})\,a_k = \Sigma\, g\,({}_rP_t)\,a_k \tag{8}$$

from which we get

$${}_rP_t = g^{-1}\ (\Sigma\, g\,({}_rP_{k,t})\,a_k\,/\,\Sigma\, a_k\,) \tag{9}$$

and, with the appropriate specification of g, we get most of the statistical indices mentioned in section 2.

This general approach to the problem is very useful from an empirical point of view since, defining the equivalence condition with reference to the purpose for which the index is being constructed, it has the advantage of guiding the choice of the type of mean we need to compute. In order to give an economic meaning to an index constructed in this fashion, it is of course necessary that the invariant express a concept that is concretely and economically relevant.

Regarding price variations, it is easy to see that the most suitable invariant function is most often the expenditure necessary to purchase a fixed basket of goods and services whose structure depends on the «position» (in terms of type of expenditure and behaviour) of the economic agent (real or hypothetical) who has an interest in the computation of the price index. Imposing this type of invariance condition leads to synthetic indices which are weighted means of elementary indices, with a weight structure that depends precisely on the «position» of the economic agent. Hence we find, as special cases, the aggregative indices. If the economic agent, either consumer or producer, has, for example, a conservative behaviour, that is, if between time *0* and time *t* he does not modify the structure of his consumption or of his inputs, and maintains that which was set at time *0*, the invariant quantity will be $\Sigma\, p_{kt}\, q_{k0} = \Sigma\, {}_0P_{kt}\,(p_{k0}\, q_{k0})$ and, hence, we get a Laspeyres type index. If, by contrast, the economic agent has a speculative behaviour and, anticipating the price variations, he modifies the structure of his consumption, and his expenditure, immediately after time *0*, will already refer to the quantity q_{kt} , the invariant quantity will be $\Sigma\, p_{kt}\, q_{kt} = \Sigma\, {}_0P_{kt}\,(p_{k0}\, qkt)$ producing a Paasche type index. These two expressions, therefore, correspond to two extreme forms of behaviour. Following this line of thought, we should then have, at least theoretically, as many indices as there are intermediate consumer (or producer) positions. It is not guaranteed, then, that an index constructed to satisfy the needs of a hypothetical economic agent, with his specific position, will satisfy the needs of another agent in a different position.

Finally, it is easy to observe that also the economic indices can be obtained following this approach, imposing an equivalence condition in which the invariant quantity is the expenditure necessary to obtain a specified level of utility.

We spent quite a long time on this approach because, even if the expressions we got are old and well known, we think that it permits, case by case and in a coherent way, the choice of the most suitable formulas on the basis of equivalence conditions which justify - with reference to the purpose and character of the research - the substitution of the «global variation» for the price variations of the single goods and services.

The main advantage of this approach, is, in our opinion, the fact that the reference to an equivalence condition (implicit in any synthetic index), forces us to consider the purpose for which the index was computed and the meaning of the «position» to which we are referring. And therefore, the equivalence condition affect also the reference population and all the phases of the index computation. It also compels us to evaluate, with an acceptable degree of approximation, whether or not different conditions can satisfy our needs and it forces us , in any case, to be coherent with such condition, in the various phases of the construction of the index.

5. Some measures of the validity errors: the components of the divergence between computed and desired indices

In order to specify some partial measures of the divergence (bias attributed to the validity errors) between the actually computed index and the desired or the most satisfactory one, it is important to refer to a framework that considers all the aspects of the accuracy and validity of a price index.

This framework has already been presented by the author in a previous paper where the Total Quality Management Approach for quality improvement and harmonisation of price indices was analysed. In that case, the production process of the index focused on the user, and hence on the purpose of the index (Biggeri, 1996). Some applications, following this approach, to improve the quality of CPIs have already been carried out by some National Statistical Offices (see for example Haworth, Fenwick, and Beaven, 1997) with good results**.**

To evaluate the results of the production processes with different sets of controllable factors – that is, with various alternative methods of construction – a definition of a *performance measure* for the output product, that is for the price index P, is needed. Such measure is usually based on an error model.

Different approaches may be followed in building errors models. Clearly the errors that may be made in estimating a CPI are closely related to the complexity of its method of construction. Many contributions on this subject have been presented at various international seminars (see for example: Turvey, 1989; EUROSTAT, 1996; Balk, 1997). We could conceptualise the problem, starting from the following mean-square error model where the total error is defined as the difference between the estimated (computed) value and the value of the desired index considering its purpose (Biggeri and Giommi, 1987):

$$MSE = E(P- P')^2 + (P'- P'')^2 + (P''- P^*)^2 + 2\,(P'- P'')\,(P''- P^*) \qquad (10)$$

where: the indices have the structure as in formula (6) and *P* is the calculated survey index; *P'* is the expected value of the *index (E(P))*; *P''* is the operational defined index (the index defined taking into account the operative definitions) and *P** is desired index, taking into account the user's needs and /or the requirement of the theory.

The components of the MSE, that is the various differences included in (10), represent: (P – P') is the *sampling errors*, *(P' – P'')* is the *measurement errors* or bias, *(P'' – P*)* is the *validity errors* and *(P'- P'') (P''- P*)* is the interaction of the bias terms. This framework can be used to classify the several possible CPI errors, since each difference represents various kind of specific errors which can be committed in the realisation of the production process of the index (Biggeri and Giommi, 1987). Since P is not in general obtained from a single survey, the sampling and non-sampling errors, being related to all the surveys used for the construction of the index, cannot be easily specified by a single complex model. However, it is often possible to develop partial measures, in which only the effect of a specific single source of error is quantified. Such measurements, often obtainable only through artificial experiments and by the introduction of some very restrictive hypotheses, are nevertheless very useful and must be further developed in the future - even if they yield only rough estimates of the errors.

As far as the evaluation of some measurement errors and validity errors, to which we are interested here, we might mention the errors due, for example, to the difficulties of operational definition of the reference population (coverage errors), to the arrangement of a representative sample of the items included in the basket, to the inappropriateness of the indirect indicators eventually used for the weights, to the non adequacy of the index formula used with the respect to the desired one, etc.. Regarding these errors, it is necessary to keep in mind that, if the *indices are calculated as weighted means*, in many cases it is possible to have an approximate measure of the bias and analyse the components of them.

In fact, as is well known, starting from the Bortkiewicz's theorem, it is possible to decompose the divergence between prices indices associated with different systems of weights and this can be done by different factors and elements (see for example Schultz, 1997).

Along the same lines, it seems to us that it is also possible to consider some slightly different decomposition of the divergence between the actually computed index (P^c) and the most satisfactory (or desired) one (P^d) referred to the same interval of time, taking into account the fact that most of the mentioned bias have effect on the system of weights. The divergence between the two indices would then depend only on the differences in the weighting system and could be decomposed attributing a precise meaning to each of its components.

If we denote by $d_k = w_k^c – w_k^d$, the difference between the standardised weights used to calculate the two indices, by s_p and s_d respectively the standard deviations of elementary (or partial) indices and of differences between weights d_k and by R_{pd} the linear correlation coefficient between the elementary (or partial) price indices P_k and the difference in the corresponding weights (that is $R_{pd} = \Sigma \ (P_k – P^m) \ (d_k – d^m) / n \ s_p \ s_d$, where n is the number of commodities and services (or commodity groups), P^m and d^m are the arithmetic means of P_k and d_k), then the difference between the two indices, the calculated and the satisfactory one, can be decomposed as follows:

$$P^c – P^d = \Sigma_k P_k d_k = n \ s_p \ s_d \ R_{pd} \quad . \tag{11}$$

It is important to emphasise the fact that the difference between the two indices vanishes when there is no relationship between the price variations of the single commodities and the differences between the weights attributed to them, and when one of the standard deviations of elementary indices or of the differences between weights is equal to zero.

The decomposition presented here does not refer to the relationship between price and quantity relatives, as is usually done in other decompositions. Our proposal, even if it does not have the important economic meaning of the other decompositions, is particularly interesting, for two main reasons: first, from an empirical point of view, it is generally more difficult to get detailed data on quantities rather than to obtain information on the structure of the weighting system; second, it is easier, making suitable hypotheses, to estimate a conjectural weighting system in order to judge whether the computed index is reasonable in comparison to the desired one, particularly when we do not have enough data.

Furthermore, following this approach, it is possible to get a measure of the coverage and representation errors.

Let us denote again by P^d a desired total index, by P^1 the price index for the commodities included in the computation, P^0 the price index for the excluded commodities and α and β the normalised weights for the two commodity groups (or sub-sets of elementary indices); the unknown desired total index would be:

$$P^d = \alpha P^1 + \beta P^0 = P^1 + \beta (P^0 - P^1) \tag{12}$$

and then

$$P^d – P^1 = \beta (P^0 – P^1). \tag{13}$$

The difference $P^s – P^1$ will depend, as is well known, on the weights of the excluded commodities and on the difference between the two indices $(P^0 – P^1)$, that is on the different evolution of the set of elementary indices included in the computed index and of the set of elementary indices excluded from the computation. Moreover, taking into account what we have showed above, (13) may be decomposed as in (11) giving interesting information on its characteristics.

From what we have said above, it should be evident the importance of studies on both, elementary indices and weights variability. It should also be clear the importance of studying the relationships between indices and weights for all commodities or for groups of commodities in order to get information on the magnitude of the different type of errors and on the bias in the CPIs. Obviously, for this kind of study we need more detailed information than that normally used for the computation of the index. But there is no doubt that the National Statistical Offices could obtain such information by the new technology used for the capture of data (for example using scanner data), at least on the occasion of the so called 'benchmark' surveys. The proposed analyses would ease the comparison between the desired and the surrogate index and would surely contribute to a better choice of CPIs, adequate for the different purposes.

References

Allen, R. G. D. (1975), *Index Numbers in Theory and Practice*, Mac Millan, New York.

Astin, J. A. and Sellwood, D. J. (1997), "Harmonization in the European Union: a Review of some Technical Issues", in Balk, B. M. (1997).

Australian Bureau of Statistics (1997a), *An Analytical Framework for Price Indexes in Australia*, Information Paper 6421.0, Camberra.

Australian Bureau of Statistics (1997b), *Issues to Be Considered During the 13th Series Australian Consumer Price Index review*, Information Paper 6451.0, W. McLennan, Australian Statistician, Camberra.

Australian Bureau of Statistics (1997c), *Outcome of the 13th Series Australian Consumer Price Index Review*, Information Paper 6453.0, W. McLennan, Australian Statistician, Camberra.

Balk, B. M., Editor (1997), *Proceedings of the Third Meeting of the International Working Group on Price Indices*, Statistics Netherlands.

Biggeri, L. (1984), "Teoria e pratica dei numeri indici: vecchi orientamenti e recenti sviluppi" (The Theory and Practice of Index Numbers: Old Approaches and New Developments), Proceedings of the XXXI Scientific Meeting of the Italian Statistical Society, Sorrento, Italy.

Biggeri, L. and Giommi, A. (1987), "On the Accuracy and Precision of the Consumer Price Indices. Methods and Applications to Evaluate the Influence of the Sampling of Households", *Proceedings of the 46th Session of the ISI* (International Statistical Institute), Book 2, Tokyo.

Biggeri, L. (1996), "Improving the Quality of Price Indices", in EUROSTAT (1996).

Cook, L. and Lewington, R. (1997), "What should the CPI Measure? A Paper for the CPI Revision Advisory Committee 1997", Draft Paper, Statistics New Zealand, Wellington.

Dalén, J. (1994), "Sensitivity Analyses for Harmonising European Consumer Price Indices", in Statistics Canada (1994).

Diewert, W.E. (1995), "On the Stochastic Approach to Index Numbers", Discussion Paper N. 95/31, Department of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W.E. (1996), "Sources of Bias in Consumer Price Indexes", Discussion Paper N. 96/4, Department of Economics, University of British Columbia, Vancouver, Canada.

Eichhorn, W. and Voeller, J. (1976), *Theory of the Price Index*, Springer-Verlag, Berlin.

EUROSTAT (1996), *Improving the Quality of Price Indices*: *CPI and PPP*, Proceedings of the International Seminar held in Florence, Italy, December 18-20, 1995, Luxembourg.

Gini, C. (1924), "Quelques consideration au sujet de la construction de nombres indices des prix et des questiones analogues", *Metron*, vol. IV,n. 1, Roma.

Haworth, M., Fenwick, D. and Beaven, R. (1997), "Recent Developments in the UK Retail Price Index: Quality Management", in Balk (1997).

Hill, P. (1996), "The scope of the HICPs", paper presented at the EUROSTAT Task Force meeting on Harmonization of Consumer Prices Indices, Rome, 16-17 September 1996.

Keynes, J. M. (1930), *A Treatise on Money*, Vol. 1, Mac Millan, London.

OECD (1997), "Synthesis Paper on Shortcomings on the Consumer Price Index. Measure of Inflation for the Economic Policy Purposes", Internal discussion paper, ECO/CPE/GEN(97) 3.

Parenti, G. (1948), *Lezioni di statistica economica*, (Economic Statistics), Bozzi, Genova, Italy.

Schultz, B. J. (1997), "Decomposition of the Divergence between Price Indices Associated with Alternative Baskets", Draft Paper presented at the Third Meeting of the International Working Group on Price Indices held from 16th to 18th of April 1997, Voorburg, Netherlands.

Statistics Canada (1994), *International Conference on Price Indices*, Papers and Final Report, November 1994, Ottawa.

Turvey, R. (1989), *Consumer Price Indices. An ILO Manual*, International Labour Office, Geneva.

U.S. Bureau of Labor Statistics (1993), The Consumer Price Index: Five Reprints from the Monthly Labor Review, December 1993, Washington D.C.

U.S. Senate Finance Committee, (1996), *Toward a More Accurate Measure of the Cost of Living, final report to the Senate Finance Committee from the Advisory Committee to Study the Consumer Price Index*, Washington D.C.