Statistical Target and Sample Design Simulations for Consumer Price Sub-Indices

Carlo De Gregorio *
Senior Researcher, ISTAT, via Cesare Balbo 16, 00184, Rome, Italy (cadegreg@istat.it)

This paper provides a comparison of the sample size needs to estimate Laspeyres consumer price sub-indices based on alternative sample designs, elementary aggregation and temporal targets. In a simplified consumer market it is firstly provided a definition of the statistical target. Sample size is determined with simple and stratified random designs under three distinct approaches to elementary aggregation. Alternative temporal targets are examined: the independent monthly indices, the whole set of monthly and quarterly indices, the annual average index, and the annual link. Empirical evidence is provided on the basis of microdata characterized by high volatility within and between months.

KEYWORDS: chained index; aggregation; sample size; sampling variance; air transports; package holidays

1. INTRODUCTION

In recent years there has been a growing concern for a more explicit use of the concepts and tools of statistical inference to produce estimates of consumer price indices (CPI) and, in particular, to define the targets of the estimates in a manner that is usual for statistical surveys (Kott 1984; Dalèn 2001; the pioneering work on this subject started with Banerjee (1956) and Adelman (1958)). Although this issue has never been at the core of CPI literature, a systematic work on estimating the sampling variance of the Laspeyres form of the CPI has developed since mid-eighties especially in the US and in Sweden. In Dalèn and Ohlsson (1995) the literature on this subject is briefly surveyed (see also Wilkerson 1967; Dippo and Wolter 1983; Leaver, Johnstone and Archer 1991; Baskin and Leaver 1996; Norberg 2004; for an overview of variance estimation approaches in selected countries see ILO (2004, chap.5)). In the last decade, further research focused also on specific issues, such as the properties of sample designs based on alternative formulas for elementary aggregation (Dalèn 1992; Baskin and Leaver 1996; Fenwick 1998; ILO 2004, chap.5; Balk 2008, chap.5). Stimuli for more developed statistical techniques also derived from the improvements that occurred in price collection, especially in selected consumer markets: this happened with scanner data (De Haan, Opperdoes and Schut 1997; Fenwick 2001; Koskimäki and Ylä-Jarkko 2003) and with other sources like e-commerce and administrative or private databases. In general, the availability of larger and more flexible data sets of price quotes made it necessary to set up generalized methods, fostering a greater attention on sampling. The integration with other statistical sources also favored the adoption of sampling based approaches: for instance, the availability of regularly updated business registers has been considered to improve sample design (Biggeri and Falorsi 2006).

* The period that I spent in 2007-2009 as a national expert at the "Price statistics" unit of EUROSTAT has been a very helpful guidance on CPI sampling issues. I wish to thank, in particular, Alexandre Makaronidis, head of the HICP unit, and Keith Hayes (head of HICP methodology), as well as the colleagues of the National Statistical Institutes in the Task Force on HICP sampling. I am also grateful to my colleagues at ISTAT: Alessandro Brunetti and Paola Anitori for their helpful comments and suggestions, as well as to the other colleagues of CPI unit, namely Carmina Munzi (to her memory this paper is dedicated), Patrizia Caredda, Stefania Fatello, Rosanna Lo Conte, Stefano Mosca, Francesca Rossetti and Paola Zavagnini. Nevertheless, the views expressed here are solely those of the author and do not necessarily reflect those of ISTAT or EUROSTAT.

Nevertheless, most of the approaches adopted to measure the variance of the estimates relied on the use of replication techniques, since the data available for empirical analysis derive mainly from purposive samples and quite rarely from probabilistic designs (Balk 2008, p.176). In Dalèn and Ohlsson (1995) a design based approach is introduced and then further developed (see also Kott 1984; Dalèn 1998, 2001; Ribe 2000; Dorfman, Lent, Leaver and Wegman 2006). The definition of the universe and of the target parameters appears by far as the most critical issue to be faced to define properly CPI sampling and, more in general, the CPI itself (Dalèn 1998, 2001). Several factors connected to the rapid evolution of consumer markets impair its definition: the replacement of products and retail, the changes in their characteristics, all need to be tackled, at least theoretically, in order to provide a solid foundation for the production of the estimates. Therefore, the need of a structured framework of concepts and definitions has emerged in order to reduce complexity, at least conceptually, and to provide sufficiently general and operative solutions: "*The consumer market ultimately consists of an enormous (but finite!) number of transactions, where goods and services (products) are purchased by consumers. However, it is not feasible to compare transactions directly between periods. Like the physicists who divide matter successively into molecules, atoms and nucleons, we have to bring some structure into our market universe as a prerequisite for a measurement procedure*." (Dalèn 2001, p.3).

Ribe (2000) and the most recent methodological developments of the chained Laspeyres index adopted in the EU (the Harmonized Index of Consumer Prices, hereafter HICP) propose a structuring of transactions into homogeneous partitions based on the concepts of product-offer and consumption segments. Starting from these approaches, this paper proposes a definition of the target universe under quite restrictive hypotheses on the functioning of consumption markets, and provides a tool to measure the sample size needs in order to estimate HICP sub-indices under alternative sample designs. The results obtained with simple and stratified random designs are compared, taking into account the use of different criteria for segmentation, elementary aggregation and temporal targets for the estimates. Two case studies are also developed. In a simplified consumer market, and on the basis of the annually chained Laspeyres formula used for the HICP, a definition of the statistical target for a monthly index is firstly provided (Section 2). Given a desired precision level, the sample size is determined in a simple random design under three distinct approaches to elementary aggregation, namely the Carli, the Jevons and the Dutot formulas. This approach is then replicated with stratified random designs. For each type of design, alternative temporal targets are examined: namely, monthly indices, quarterly and annual indices, and the annual link (Section 3). The approaches are then tested on an experimental ground on artificial populations simulated on the basis of the microdata relating to two sub-indices of the HICP - air transports and package holidays - both characterized by high volatility of price dynamics within and between months (Section 4).

2. THE STATISTICAL TARGET

2.1. Some Aspects of the Construction of the HICP

The HICP is a monthly Laspeyres index whose fixed base is given by the average of a reference year (*yr*): at present, *yr=2005*. It is built as a chained index by linking together the monthly price indices of the current year *y* - based on the link month of *December y-1* - and the fixed base index *H* of *December y-1*. By iteration, in the reporting month *m* of year *y* the aggregate HICP is derived from the product of three elements: a fixed base index (*H*), the product of the *(y-yr-1)* annual links, and the link index of the reporting month (*I*). In formulas (EUROSTAT 2001, p.175-197):

$$H^{y,m} = H^{yr,12}\left(\prod_{x=yr+1}^{y-1} I^{x,12}\right) I^{y,m} \qquad (1),$$

where, to simplify notation, the basis of all indices has been set to 1 instead of the usual 100. The link of the reporting month $I^{y,m}$ is compiled as the weighted mean of the sub-indices referred to an exhaustive set of disjoint aggregates *j* of the total consumer expenditure in the weight reference year:

$$I^{y,m} = \sum_j I_j^{y,m} w_j^y \quad (2),$$

where the expenditure weights *w* sum up to unity and in principle they change every year, since they are referred to the consumption expenditure of year *y-1*: furthermore, the weights are price-updated from the weight reference period (*y-1*) to the price reference period (*December y-1*) (EUROSTAT 2001, p.188-190; Hansen 2006; ILO 2004, chap.9). The construction of any HICP aggregate follows a hierarchical procedure: the aggregation of the link indices comes first, and the result is then chained to the fixed base index of the same aggregate. The sub-indices $I_j^{y,m}$ can be therefore interpreted as the primary components of the HICP and, as a consequence, they represent the statistical targets (Ribe 2000, p.1): hereafter we shall refer to the problem of estimating these sub-indices. Notice also that expression (2) can be applied to any exhaustive partition of the target consumption expenditure: we choose in particular to deal with one of these partitions, namely the groups of COICOP-HICP classification, corresponding to the lowest level of breakdown actually used for HICP dissemination. It consists of almost 100 sub-indices (EUROSTAT 2001, p.253-68).

2.2. The HICP conceptual framework

Ribe (2000) defines a structured definition of the universe for HICP, as part of the studies which contributed to the advancements in HICP methodological framework and legal basis (see Commission Regulation (EC) no. 1334/2007; a collection of the early legislation concerning HICP can be found in EUROSTAT (2001)). Ribe (2000) ultimately derives the target parameter for the annual links of a monthly Laspeyres CPI as the ratio of two simulated consumption expenditures, obtained throughout the mapping of the universe of transactions in the weight reference period (year *y-1*) into the sets of offers available in the price reference period (*December y-1*) and in the reporting period (month *m* of year *y*). To define this re-pricing of transactions, the concept of product-offer is introduced and more precisely defined in EC Regulation 1334/2007: "*product-offer means a specified good or service that is offered for purchase at a stated price, in a specific outlet or by a specific provider, under specific terms of supply, and thus defines a unique entity at any one time*". As a matter of fact, product-offers are the observation units in CPI sampling and they determine a partition of total transactions. Nevertheless they represent a rapidly changing stock: they change as the characteristics of the goods and services evolve, or as they are replaced, or as retail evolves, or simply as prices change. In order to provide stable entities on which to base price comparisons, the sets of all the transactions and product-offers in the statistical universe are exhaustively clustered into consumption segments, where each segment identifies homogeneous product-offers as concerns marketing targets, consumption purposes and characteristics. Consumption segments represent the fixed objects to be followed by the Laspeyres index. In particular: "*(...) 'consumption segment' means a set of transactions relating to product-offers which, on the grounds of common properties, are deemed to serve a common purpose, in the sense that they: are marketed for predominant use in similar situations, can largely be described by a common specification, and may be considered by consumers as equivalent. (….). The notion of consumption segments by purpose is therefore central to sampling and to the meaning of quality change and quality adjustment. However, an ambiguity in this concept concerns the level of aggregation at which it is defined and applied. (...) The range of product-offers will change over time as products are modified or replaced by retailers and manufacturers. The HICP requires the representation of all currently available product-offers within the consumption segments by purpose selected in the reference period in order to measure their impact on inflation. This applies particularly to new models or varieties of previously existing products*" (EC Regulation 1334/2007).

This structured framework for the definition of HICP statistical universe remains at present a theoretical tool open to a wide range of possible solutions, while methodological and empirical research is still needed in order to test its applicability as a statistical tool. One key point is given by the definition of consumption segments. HICP regulation itself recognizes that ambiguities still concern the level of aggregation with which consumption segments are defined and applied. Furthermore, the characteristics of the mapping functions for re-pricing remain undetermined and the replication of the transactions in the price reference period and in the reporting period is open to several alternatives. More concretely, consumption segments need to be specified case-by-case and this fosters the strategic role of consumer markets analysis, as concerns for example the structure of supply and demand, the marketing approaches and the segmentations adopted by producers and dealers.

2.3. The Definition of the Target Parameter

Given the Laspeyres formula, and following the approach set up in Ribe (2000) and in HICP legal basis, the point of departure for the definition of each target sub-index is given by the set of all the transactions in the weight reference year *y-1* concerning the COICOP-HICP group *j*. It is assumed a perfect knowledge of all the information which is necessary to compute the indices. In particular, each transaction in the weight reference period is tracked; it concerns the purchase of a product-offer, where each product-offer is attributed to a specified consumption segment. Product-offers are defined by the combination of two sets of characteristics.

A first set consists of a vector g_i of variables describing the product, the outlet and the corresponding consumption segment (*h*): as shortcut, we shall refer to such a vector with the term "product". Naming with $G = \{g_i \mid i = 1,...,N\}$ the set of all available products for the consumption purpose *j*, it is exhaustively divided in *M* disjoint consumption segments G_h, with *h=1,…,M*. This partition (Γ_M) can be expressed as follows (to economize notation, hereafter we omit in formulas the corresponding suffix *j*):

$$\Gamma_M = \{G_h, h = 1,...,M\},$$

where $\bigcup_h G_h = G$ and $G_{h'} \bigcap G_{h''} = 0$ for every $h' \neq h''$. In order to simplify the definition of the universe and of the target parameter, some restrictive assumptions on the available product-offers are introduced to reduce the dynamics of the universe only to price changes and to provide a simplified framework for the definition of the statistical target (Dalèn 2001; Balk 2008, chap.5). The elements of the set *G* are assumed to be fixed and time-invariant: the number and the characteristics of the available offers do not change, and the outlets and providers remain also unchanged.

The second set of characteristics is given by a vector of data concerning the act of purchase: namely, the time and the price (*p*). In principle, the definition of product-offer is fixed to a point in time, so that a specific model of a product in a given outlet today is a different product-offer from the same model in the same outlet next month. Discrete (monthly) pricing policies are assumed, where the prices of each element of *G* are eventually changed only at the beginning of each month. Given these assumptions, the generic element of the *(Nx13)* matrix Ω^{y-1} of all the product-offers in year *y-1* is given by:

$$\left(g_i; \quad p_i^{y-1,1} \quad ... \quad p_i^{y-1,12}\right).$$

The consumption expenditure (*E*) in the weight reference year *y-1* we have:

$$E^{y-1}\left(\Omega^{y-1}\right) = \sum_h \overline{p}_h^{y-1} T_h^{y-1} \qquad (3),$$

where T labels the number of transactions and $\overline{p}_h^{y-1} = \sum_m p_h^{y-1,m} \left(\frac{T_h^{y-1,m}}{T_h^{y-1}} \right)$ is the annual average price actually paid for the transactions in the consumption segment h.

In order to define the "true" value of the target sub-index $I_j^{y,m}$ - and following the principle of adopting consumption segments as the fixed objects followed by HICP (EC Regulation 1334/2007) - we need to simulate the total consumer expenditure of the weight reference year (*y-1*) on the basis of the product-offers available in the reporting month m of year y (identified by the couple (*y, m*)) and in the price reference month (*December y-1,* conventionally labeled with (*y,0*)). By applying (3) we obtain:

$$I^{y,m} = \frac{E^{y-1}\left(\Omega^{y,m}\right)}{E^{y-1}\left(\Omega^{y,0}\right)} = \frac{\sum_h \overline{p}_h^{y,m} T_h^{y-1}}{\sum_h \overline{p}_h^{y,0} T_h^{y-1}} = \sum_h I_h^{y,m} w_h \qquad (4),$$

where $w_h^y = \frac{\overline{p}_h^{y,0} T_h^{y-1}}{\sum_h \overline{p}_h^{y,0} T_h^{y-1}}$ is the value weight of consumption segment h and $I_h^{y,m} = \frac{\overline{p}_h^{y,m}}{\overline{p}_h^{y,0}}$ is its price relative (non zero average prices by segment in the base month (y,0) are here assumed; on the treatment of zero prices in the HICP see EUROSTAT (2001, p. 184-5)). Average prices are derived on the basis of a mapping of the set of product-offers Ω^{y-1} into the sets $\Omega^{y,0}$ and $\Omega^{y,m}$ available in the base and current months. The hypotheses of time invariance of the set G makes it possible to assume a one-to-one mapping, driven by the elements of the set G: for each transaction involving the product g_i in *y-1*, the corresponding product-offers are $(g_i, p_i^{y,0})$ and $(g_i, p_i^{y,m})$. Different versions of the target parameter defined in (4) can be provided according to the method of aggregation that is used to calculate average prices. Two alternative approaches are here proposed, namely the weighted arithmetic (5) and the geometric mean (6):

$$\overline{p}_h^{y,m} = \frac{\sum_{g_i \in G_h} p_i^{y,m} T_i^{y-1,m}}{\sum_{g_i \in G_h} T_i^{y-1,m}} \qquad (5),$$

$$\overline{p}_h^{y,m} = \exp\left[\frac{\sum_{g_i \in G_h} T_i^{y-1,m} \ln\left(I_i^m\right)}{\sum_{g_i \in G_h} T_i^{y-1,m}} \right] \qquad (6).$$

It is well known that each approach implies specific assumptions on consumers' elasticity to price changes (see below, section 3.1). The objective is now to compare the properties of alternative standard sampling approaches adopted to provide an estimate of $I^{y,m}$.

3. SAMPLE SIZE WITH ALTERNATIVE DESIGNS AND AGGREGATION FORMULAS

3.1. Simple Random Sampling (SRS)

Assume that a simple random sample S of n products is drawn from the set G, while no other information is available on the universe of product-offers, consumption segments and transactions. This is equivalent to drawing a sample of n product-offers from the sets of available product-offers $\Omega^{y,0}$ and $\Omega^{y,m}$.

Different estimates of I^m can be produced, depending on the approach followed to aggregate the sampled quotes and to produce the estimates of the target index (hereafter, we drop the suffix labeling the year). Three alternative types of frequently used unweighted means are here compared: Carli (arithmetic mean of price relatives, labeled with "C"), Dutot (ratio of mean prices, "D") and Jevons (geometric mean, "J"). We have then, respectively:

$$\hat{I}^{m,C} = \frac{\sum_{i \in S} \frac{p_i^m}{p_i^0}}{n} = \frac{\sum_{i \in S} I_i^m}{n} \quad (7),$$

$$\hat{I}^{m,D} = \frac{\sum_{i \in S} p_i^m \Big/ n}{\sum_{i \in S} p_i^0 \Big/ n} = \frac{\sum_{i \in S} I_i^m p_i^0}{\sum_{i \in S} p_i^0} \quad (8),$$

$$\hat{I}^{m,J} = \exp\left[\frac{\sum_{i \in S} \ln\left(I_i^m\right)}{n}\right] \quad (9).$$

The relative convenience of these three formulas has been deeply debated in the literature. Each approach entails specific assumptions on consumers' elasticity to price changes, which are reflected on the implicit weighting of transactions (Ilo 2004, chap.9; Leifer 2002, 2008; Viglino 2003; Balk 2003, 2008, chap. 5; Silver and Heravi 2006). Notice also that the systematic use of the Carli formula has been banned for the estimates of the HICP (Commission Regulation (EC) No 1749/96, Art.7; EUROSTAT 2001, p.129, 155-156). Roughly speaking, the Carli formula implies equal value weights for the product-offers and, for each product-offer, a constant expenditure in the price reference and in the reporting month. The Dutot formula implies equal and time-invariant quantities for each product-offer. The Jevons formula assumes that the expenditure shares of the price reference month do not change when relative prices change: a substitution due to the change in relative prices is therefore implied. With respect to the Carli formula, the Dutot approach gives a higher weight to the product-offers with a higher price in the price reference month and the Jevons approach gives a higher weight to the product-offers with a lower price dynamics. Without entering the issue of the choice of the "right" formula, we want to investigate here some of their statistical properties in terms of precision within different sampling designs (Fenwick 2008; Baskin and Leaver 1996). Formula (7) provides an unbiased estimator of (4)-(5) only if the probability of selection is proportional to the weight of each product (Adelman 1958). The same applies to (9) with respect to the target set by expressions (4) and (6). For the Dutot formula (8) to be unbiased with respect to (4)-(5) it is necessary to add the condition that the price relatives are independent from the price levels in the price reference month (Balk 2008, chap.5).

Given a confidence level α and an relative error expressed as a share of the sample mean ($r = \beta I^{m,q}$, where $q \in \{C, D, J\}$), the adoption of these three methods implies some differences in the sample sizes which are needed to produce an error that is inferior to the $\beta\%$ of the true value of the parameter with a probability of $\alpha\%$. These differences depend on the standard errors of the three types of sample means and on the form of their distributions. By adopting standard simple random sampling theory (Cochran 1977, chap. 4-6) separately for the three aggregation formulas, in the case of the Carli formula the necessary sample size in month *m* can be expressed as follows:

$$n_{SRS}^{m,C}(\alpha,\beta) \geq \left[\frac{t_\alpha}{\beta}\frac{s_{\hat{I}}^{m,C}}{\hat{I}^{m,C}}\right]^2 = \left[\frac{t_\alpha}{\beta}\right]^2 C_{I^{m,C}}^2 \qquad (10),$$

where t is the corresponding value of the t-Student distribution, $s_{\hat{I}}^{m,C}$ is the standard error of the sample Carli mean and $C_{I^{m,C}}$ is the coefficient of variation of the Carli index (Cochran 1977, sec. 4.6; Ilo 2004, chap.5; hereafter, the sample fraction correction is not considered).

For the Dutot index we obtain:

$$n_{SRS}^{m,D}(\alpha,\beta) \geq \left[\frac{t_\alpha}{\beta}\right]^2 \left[C_{p^m}^2 + C_{p^0}^2 - 2\rho_{0,m} C_{p^0} C_{p^m}\right] \qquad (11),$$

where $\rho_{0,m}$ is the Pearson's correlation coefficient of the price levels in the price reference and in the comparison month, while C_{p^m} and C_{p^0} are the coefficients of variation of the price series in the two months (Cochran 1977, sec. 6.3-6.5; Ilo 2004, chap.5).

For the Jevons index, the following expression is derived:

$$n_{SRS}^{m,J}(\alpha,\beta) \geq \left[\frac{t_\alpha}{2\log(x_\beta)}\right]^2 \left(s_{\log(I)}^m\right)^2 =$$

$$= \left[\frac{t_\alpha}{2\log(x_\beta)}\right]^2 C_{I^{m,J}}^2 \qquad (12),$$

where $x_\beta = \frac{+\beta + \sqrt{\beta^2 + 4}}{2}$, $s_{\log(I)}^m$ is the standard deviation of the logarithms of the individual indices I_i and it is equal to the coefficient of variation of the Jevons mean $C_{I^{m,J}}$ (Cochran 1977, sec. 4.6; Ilo 2004, chap.5). The expression for x_β can be derived by applying the SRS formula for confidence interval to the log transformed variable and then transforming back and resolving by n. Following Norris (1940) $\hat{I}^{m,J} s_{\log(I)}^m$ corresponds to an estimate of the standard deviation of the geometric mean.

Expressions (10)-(12) derive sample size from the product between two elements: one dependent on α and β, and the other one based on the coefficients of variation of indices and – in the case of Dutot - price levels. For reasonably low values of β (e.g., inferior to 10%), the comparison among these formulas can be limited to the latter. In general, when the variability of prices and indices is very small, the three approaches lead to very similar sample sizes. On the contrary, when the variability of the distribution of indices and price levels is relatively large some important differences might appear. The Dutot index needs a higher sample size when there is a strong heterogeneity in price levels with negative or low positive correlation between price levels, and in particular when the largest price changes are associated with the goods with a higher price level in the price reference month. In the case of the Jevons formula, the sample size tends to be relatively higher if the distribution of the price changes is negatively skewed while the opposite happens with a positive skew. It is reasonable to expect that a partition in consumption segments can potentially isolate homogeneous product-offers and reduce consequently the heterogeneities among aggregation formulas.

3.2. Stratified Random Sampling (STRS)

We assume now that more information is available concerning the consumption expenditure in the weight reference year: the weighting structure (w_h) of a partition in consumption segments (Γ_M) is known, although no other information is available on the weighting of the product-offers within each segment. If a stratified random design is adopted, the estimate of the target parameter in fact may be obtained as a value-weighted arithmetic mean of the indices of each segment (stratum):

$$\hat{I}_{STRS}^{m,q} = \sum \hat{I}_h^{m,q} w_h \ .$$

The standard deviations within each stratum, for the three alternative formulas, will be given by:

$$s_h^{m,C} = \frac{\sum_{i \in h} \left(I_i - \hat{I}_h^{m,C} \right)^2}{n_h} ,$$

$$s_h^{m,D} = \frac{1}{n_h \left(\bar{p}_h^0 \right)^2} \left[s_{p_h^m}^2 + \left(\hat{I}_h^{m,D} s_{p_h^0} \right)^2 - 2 \hat{I}_h^{m,D} s_{p_h^0 p_h^m} \right] ,$$

$$s_h^{m,J} = \hat{I}_h^{m,J} s_{\log(I_h)}^m \ .$$

Independently of the type of elementary aggregation, total sample size with optimal allocation can be expressed as follows:

$$n_{STRS}^{m,q}(\alpha,\beta) \geq \left[\frac{t_\alpha}{\beta} \right]^2 \left[\frac{\sum_h s_h^{m,q} w_h}{\sum_h \hat{I}_h^{m,q} w_h} \right]^2 \qquad (13),$$

where $s_h^{m,C}$ is the standard deviation within stratum *h* (Cochran 1977, chap. 5.4-5.9; as in the case of SRS, the sampling fraction correction has not been skipped). Following the optimal allocation per strata (i.e. proportional to the standard deviation), sample size in each stratum can be expressed as follows:

$$n_{STRS,h}^{m,q}(\alpha,\beta) = n_{STRS}^{m,q}(\alpha,\beta) \frac{s_h^{m,q} w_h}{\sum_h s_h^{m,q} w_h} \qquad (14).$$

Expression (13) suggests that stratified designs can reduce the source of discrepancies among aggregation formulas, depending on their ability to compress the variance within strata through a good clustering of market segments, able to isolate the criteria used to define pricing policies.

3.3. Alternative Temporal Targets

The sample sizes derived in expressions (10)-(12) and (13)-(14) all concern the target price index referred to a generic reporting month *m*. Nevertheless, the objective is to produce the complete series of twelve monthly estimates. This fact bears a number of consequences. The sample size for month *m* is in general different from the one needed to arrange the same precision for another month *m'*. In fact, the nature of price dynamics possibly changes from month to month and in a way which depends on the specific demand and supply characteristics of each consumer market (for a classification of price index dynamics within the HICP see De Gregorio (2010)). A first consequence is that monthly sizes can differ substantially, and this implies that the sample in the price reference month has to take into account the size needs in all the twelve following months. In particular, in SRS designs sample size in the price reference month must be equal to the maximum size needed in the twelve months:

$$n_{SRS}(monthly) = \max_m (n^m_{SRS}(\alpha, \beta)) \tag{15}$$

Anyway, the use of the whole sample size for the price reference month derived from expression (15) is not necessarily needed in all the following months, since price collection can be modulated according to monthly needs: while the sample size in the base month should be determined on the basis of expression (15), the successive monthly price collections can be based on expressions (10)-(12). As a matter of fact this fact suggests a modular approach, implying a stronger collection effort in the price reference month: there is no reason to use a sample of constant size every month if the target is the whole set of the monthly estimates (for an application of this modular approach to seasonal products see De Gregorio, Munzi and Zavagnini (2008)).

With stratified designs, some further complications may arise, since allocation is also relevant. The sample size in the price reference month derives, in fact, from the sum of the largest monthly size of each stratum:

$$n_{STRS}(monthly) = \sum_h \max_m (n^m_{STRS,h}(\alpha, \beta)) \tag{16}$$

This amount may be significantly larger than the maximum monthly size as derived from expression (14). This happens, in particular, where the peaks in the variability of price dynamics have distinct time patterns across strata. For instance, in those markets characterized by seasonal pricing - where peak months generally show higher variability- the timing of seasonal peaks might differ across strata and this mere fact induces the need of larger samples in the price reference month; something similar might likely happen in sectors characterized by highly variable and irregular patterns.

Actually, the sub-indices with a relatively large variability, or characterized by seasonal behaviors, are generally relatively a few. In De Gregorio (2010) it is estimated that within the euro zone, in the period 2004-2008 only slightly more that 25% of HICP four-digit sub-indices showed a relatively strong monthly dynamics and only 7.3% showed a clear seasonal pattern. Most of the indices are referred to markets where, at least in periods of low inflation, price changes are quite slow. In those cases, the first months of the year may reflect a remarkable inertia: this may imply that the distribution of price changes in the first months is positively skewed, since most observations are concentrated on the no-change zone. This type of asymmetry progressively loses ground as we move away from the price reference month and we proceed towards the final part of the year. Nevertheless, for all these cases the form of the distributions might results in a problem, since the hypothesis of normality is really far from being met.

As a consequence of this inertia, the last months of the year might easily result as those with the largest sample size. Also for this reason it makes sense to use as a target for the estimates the annual link of December, whose importance relies also in the fact that this index has a permanent effect on the chained index *H* (see expression (1); Fenwick (1999) examines the issue of the choice of the price reference month, emphasizing the problems that may arise in the choice of the aggregation formula in case of large variability of price dynamics). In this case, with the two types of design we obtain:

$$n(link) = n^{12}(\alpha, \beta) \tag{17}$$

This formula bears relevant gains in sample size with respect to expressions (15) and (16) only if the variability of price dynamics is diluted during the year and it is not concentrated in the final month (this result does not apply to the case of monthly chained indices while it can be easily generalized to link moths different from December). Alternative reasonable targets might be set on quarterly or yearly averages:

$$n(quarterly) = \max_Q (n^Q(\alpha, \beta)) \tag{19}$$

and

$$n(yearly) = n^Y(\alpha, \beta) \quad (20).$$

In particular, for quarterly targets in stratified designs, allocation effects must also be considered as in the case of monthly estimates. It is obviously possible to use a combined target, for instance to guarantee the precision level on quarterly and annual link estimates.

4. TWO CASE STUDIES

4.1. Artificial Populations

In order to test empirically the combined effect of sample design and aggregation formulas on the variance of the estimates we have generated artificial target populations starting from survey microdata and we have iterated the extraction of samples to estimate a target parameter as defined in Section 2. In particular, two case studies are here presented, based on the microdata from the surveys used by ISTAT (the Italian National Statistical Institute) to produce the monthly estimates of two sub-indices of the HICP in 2007 (these data have been treated here with different objectives with respect to the official purposes of the Institute, and as a consequence the results cannot be compared at any rate with those actually disseminated). They are referred to series, both characterized by high variability although with quite distinct heterogeneous behaviors. In the case of European air transports, the high volatility of prices can be only partly explained by seasonal patterns. Overlapping seasonal peaks, on the other side, strongly affect the price changes in the market of package holidays, with some inertia in the first months of the year. In De Gregorio (2010), passenger transports by air and package holidays are both identified as the sub-indices with the most heterogeneous behaviors across euro zone, possibly needing further harmonization (for a methodological overview of the methods actually adopted by ISTAT to estimate these indices, see ISTAT (2009) and De Gregorio, Fatello, Lo Conte, Mosca and Rossetti (2008)).

Each set *Z* of the microdata has been interpreted as a sample of the product-offers available in year *y* ($Z \subset \Omega^y$). Each record is characterized by a product identifier (g_i) and by a vector of 13 price quotes - from month *0* (the price reference month, namely December 2006) to month *12* (December 2007). For each market, it is given a detailed and exhaustive partition of the goods in the market in M_0 disjoint sets of consumption segments:

$$\Gamma_{M_0} = \{G_h, h = 1, \ldots, M_0\}.$$

Alternative less detailed partitions Γ_{M_i} might be obtained by hierarchical aggregation of the subsets of Γ_{M_0}. For each partition it is accordingly defined a vector of normalized weights:

$$W_{M_i} = \left\{ w_h, h = 1, \ldots, M_i \mid \sum_h w_h = 1 \right\}.$$

The microdata in the sets *Z* derive from stratified samples which have not been selected with probabilistic rules (ISTAT 2009; De Gregorio, Fatello et al. 2008, p. 20, 28-32): nevertheless, they are treated here as if they were derived from random selections. Each set is expanded proportionally to a volume estimate of the weighting structure W_{M_0}. *K* simple random samples of *n* product-offers are drawn from these artificial populations. The yearly series of the average price indices $\hat{I}_k^{m,q}$ (*k=1,…,K*) are derived from each sample adopting alternatively the Jevons, Dutot or Carli aggregation (expressions (7)-(9)). An inductive estimate of the variance of the sample mean is produced and, consequently, it is provided an estimate of the sample size by means of the formulas derived in the preceding sections. An identical approach is used to estimate the sample size for stratified designs based on alternative partitions of the target population.

All the simulations for the three markets have been made by extracting iteratively 300 samples of 500 products. Given a level of error of 1% and a confidence level of 95% distinct temporal targets have been separately considered. The tables commented here below describe a relative measure of sample size, calculated as multiple of a benchmark size. Firstly, the sample size has been determined in order to obtain the desired precision level separately for each single month. Secondly, the cumulative target of precision in every month has been considered (using expression (15) and (16)). Thirdly, the target is moved to the quarterly and yearly averages (using expressions (18) and (19)). Given the chaining procedure, the desired precision has been finally set on the link month of December, which affects permanently the fixed base series (expression (17)).

4.2. European Air Transports

For the construction of the artificial population we have used data from the original sample of *N=328* product-offers, concerning as many European return flights from the country of origin (national) to the other countries (foreign). Each flight is defined by a national area of origin and a foreign area of destination, both subsets of the territory of the respective countries.

Four distinct partitions provide exhaustive segmentations of the target population. An elementary stratification Γ_{51} (51 strata) provides an exhaustive segmentation by national and foreign areas and by type of carrier (low cost vs. full service carriers). A less detailed partition collapses the areas within a same foreign country (Γ_{38}); a further aggregation of consumption segments skips the definition of national areas and uses only the country of destination and the type of carrier (Γ_{15}); the less detailed partition uses only the country of destination (Γ_{11}). An elementary consumption segment might identify, for example, the low cost flights from the area A1 in country A (national) to the area of B1 in the foreign country B; less detailed partitions identify, orderly, all the low cost flights from the area A1 to B, all the low cost flights from A to B and all the flights from A to B.

Table 1 reports the indicators of the sample sizes needed separately in each month and quarter, and for the yearly average: the sample size needed to estimate the yearly average with a sample random design using a Carli aggregation is used as a benchmark and has been set equal to 100. The table resumes the main characteristics of the monthly profile of the variability of price dynamics within each time period. It is possible to notice that with SRS, the monthly size is significantly higher as compared to the respective annual targets. Seasonal peaks are hit in May and August, where the distributions of prices and price changes appear positively skewed: the Jevons formula, in fact, is less demanding, especially in May where the Dutot formula delivers the worst result. Smaller samples are needed at the beginning and at the end of the year, and in June. The Carli formula needs the largest sample size in eleven months out of twelve: the median monthly size is more than 35% higher with respect to Jevons formula. Dutot aggregation generates lower sample sizes in most of the off-peaks months (first and fourth quarter), due to a more appreciable homogeneity in price levels. If we consider the most stratified design (Γ_{51}) sample size is strongly reduced to 25-30% of the values needed for SRS. Such decrease is particularly strong in the peak month, and more evidently in May. The effect of stratification is stronger with the Carli formula, whose performance in terms of sample size improves (10% higher than Jevons, in median). The Jevons formula, with stratified designs, brings systematically lower sample sizes.

Table 2 reports the sample size needed in the base month in order to meet the alternative targets outlined in section 3.3. For the whole set of monthly targets with a SRS approach, the use of a Carli aggregation would need nearly 5.5 times the benchmark size. The Dutot formula delivers an even worse result, due to the high heterogeneity in price levels. The Jevons approach needs half the sample size with respect to the monthly Carli. Such large samples derive from the high volatility of price levels and price dynamics in peak months. If we reduce the target to quarterly estimates, the sample sizes shrink drastically (between 1.3 to 1.8 times the benchmark) and the differences among methods also appear strongly reduced. The yearly estimates need nearly half the sample used for the quarterly target, while limiting the target to the link month of December (nearly a peak month) brings results situated between the quarterly and the yearly target. In the case of the yearly and the link targets, Dutot performs slightly better than Jevons.

Table 1. Sample size for European air transports by temporal target (Indices. Base: size for yearly target with Carli aggregation with SRS=100)

	SRS			STRS Γ_{51}		
Period	Carli	Jevons	Dutot	Carli	Jevons	Dutot
Jan	100	93	92	28	25	26
Feb	176	134	115	49	43	45
Mar	137	98	91	53	46	48
Apr	223	139	140	55	46	50
May	554	230	712	76	72	74
Jun	110	81	85	30	27	28
Jul	278	160	144	37	35	35
Aug	438	279	354	138	117	138
Sep	197	151	137	44	39	40
Oct	147	116	104	44	39	40
Nov	228	211	176	47	43	43
Dec	138	118	104	40	34	37
Year	100	74	74	25	22	23
Q1	97	74	70	29	25	26
Q2	153	107	149	41	36	38
Q3	181	130	134	45	41	43
Q4	123	102	88	29	25	26

Table 2. Sample size for European air transports (Indices. Base: size for yearly target with Carli aggregation in SRS=100)

		Temporal target			
Design	Aggregation	Monthly	Quarterly	Yearly	Annual link
SRS	Carli	554	181	100	138
	Jevons	279	130	74	118
	Dutot	712	149	74	104
STRS Γ_{11}	Carli	329	149	81	105
	Jevons	239	105	59	81
	Dutot	284	116	63	88
STRS Γ_{15}	Carli	280	106	61	90
	Jevons	225	90	50	74
	Dutot	267	98	55	81
STRS Γ_{38}	Carli	191	65	32	55
	Jevons	171	59	28	48
	Dutot	188	61	29	50
STRS Γ_{51}	Carli	169	55	25	40
	Jevons	151	49	22	34
	Dutot	167	52	23	37

Stratification induces a sharp reduction in sample size and to more homogeneous results across the three aggregation approaches. The introduction of the first two levels of stratification brings large improvements, in particular for the Carli formula. The partition in 38 strata is extremely fruitful for all the types of formulas, while the most detailed partition brings a comparatively minor reduction in sample size. In the passage from SRS to the most detailed stratified design brings a reduction of almost 70% of the sample size.

The temporal patterns of variability within strata are quite differentiated: consequently, the allocation effect induces appreciable differences between the sample size needed to target the whole set of monthly prices and the maximum size for separate monthly targets. If we consider the whole set of monthly targets (table 2 and formula (16)), the sample size for Carli and Dutot aggregation is nearly 20% higher than the maximum size reported in table 1; Jevons formula, although it is in general more efficient, needs a sample nearly 30% higher than the respective maximum (151 vs. 117). For the quarterly indices, the size increase needed to meet all the monthly targets is slightly above 20%. The estimate of the annual link requires a larger sample as compared to the yearly target. The latter, independently of the design, requires about 15-20% of the sample size needed for the monthly targets, while the link requires nearly 30% of it.

In general the Jevons aggregation performs better, with a couple of exceptions where Dutot appears less demanding. Carli generally implies larger samples, although the differences collapse as stratification runs deeper. The heterogeneity of price levels damages the performance of the Dutot formula, especially where stratification is absent or limited, while the sample size derived from the Jevons formula appears less influenced by the presence of larger prices.

4.3. Package Holidays

For the construction of the artificial population we have used data from the original sample of $N=246$ records, concerning as many packages. Each package is defined by an area of destination. Two distinct partitions provide exhaustive segmentations of the target population. An elementary stratification Γ_{43}, with 43 market segments, gives a partition by country and type of vacation, and a less detailed segmentation Γ_{12} by group of countries. As an example, an elementary segment could be the market for package holidays for area A1 in country A; a less detailed partition would concern all the packages for holidays for country A.

Differently from air transport, after the summer peaks (July and August) the required sample size remains high with respect to the first months of the year (Table 3). This is the effect of the inertia of price dynamics, since price levels in the first months tend to range near to their reference level. The annual link, in particular, still needs a large sample. The Dutot formula, which works relatively well at the beginning of the year, becomes the less appealing (in terms of sample size) from May. This is due, probably, to the high heterogeneity of the packages, whose prices may vary considerably across markets and whose level in the base month might be directly related to the entity of their change. The use of the deepest level of stratification brings large improvements especially in the first part of the year and with respect to the Dutot formula, whose performance becomes comparable with the other two. It is worth to notice that the adoption of the stratification defined by Γ_{43} is particularly effective in the seasonal peaks, while December is the month needing a larger sample.

Table 3. Sample size for package holidays by temporal target (Indices. Base: size for yearly target with Carli aggregation in SRS=100)

	SRS			STRS Γ_{43}		
Period	Carli	Jevons	Dutot	Carli	Jevons	Dutot
Jan	69	64	76	12	12	11
Feb	85	79	90	14	14	13
Mar	86	84	80	17	17	16
Apr	174	168	142	47	47	47
May	165	162	210	82	84	84
Jun	245	229	389	100	100	105
Jul	352	296	416	97	97	102
Aug	465	437	646	109	111	113
Sep	222	218	438	114	113	117
Oct	216	219	447	112	111	115
Nov	265	244	403	136	128	134
Dec	307	271	476	173	169	170
Year	100	97	179	53	52	53
Q1	71	68	73	13	12	11
Q2	139	133	180	66	67	68
Q3	307	282	458	95	95	99
Q4	220	209	390	118	113	117

Resuming the results by type of target, differently from the case of air transports, the quarterly indices with stratified designs need smaller samples as compared to the case in which the target is set on the annual link. With monthly targets, the sample size with SRS is between 6.4 to 9.5 thousands while it goes down to 1.5-2.6 thousands with the yearly target. The target on the annual link with SRS brings similar results as in the case of the quarterly target, mainly due to the fact that the last months of the last two quarters are relatively more volatile. Due to the effect of sample allocation, the Dutot aggregation with stratified designs works slightly better than the Jevons'.

The irregularity in the monthly variability within strata is less pronounced as compared to air transports. The sample size needed to target the whole set of twelve months is nearly 5% higher than the maximum size needed to meet separately the monthly targets (see respectively Tables 3 and 4). As for air transports, also in this case, the Jevons aggregation brings the largest increase.

Table 4. Sample size for package holidays (Indices. Base: size for yearly target with Carli aggregation in SRS=100)

		Temporal target			
Design	Aggregation	Monthly	Quarterly	Yearly	Annual link
SRS	Carli	465	307	100	307
	Jevons	437	282	97	271
	Dutot	646	458	179	476
STRS Γ_{12}	Carli	258	183	78	245
	Jevons	257	182	77	234
	Dutot	254	179	77	241
STRS Γ_{43}	Carli	183	123	53	173
	Jevons	184	123	52	169
	Dutot	182	123	53	170

CONCLUDING REMARKS

This work investigates several aspects of CPI sample design based on a simplified definition of the target universe derived from HICP legal basis: several conclusions can be consequently listed. Potentially heterogeneous results for optimal sample size might derive from different aggregation approaches, depending on the type of variability of price dynamics: the case studies reported in section 4 confirm this heterogeneity. Stratification on one side can strongly reduce samples and, on the other, can also reduce such heterogeneity, especially if it is based on a segmentation of the markets able to reproduce the outstanding pricing policies. The monthly changes in sample size make it possible a modular approach to sampling, concentrating larger price collections in the months where the variability of price dynamics hits a peak. The consideration of alternative temporal targets also appears as a strategic issue, in order to save resources and optimize their use.

In particular, stratified designs contribute significantly in controlling sample size, especially in consumer markets characterized by large variability in price dynamics. The empirical evidence provided in section 4 shows that stratification may produce a reduction from 50% to 70% of sample size with respect to not stratified designs. The logic which lays behind the choice of the strata obviously matters: here it has been roughly based on marketing criteria, trying to isolate possibly homogeneous consumption segments and clusters of pricing policies.

The issue of how much to stratify is also very important. The introduction of a first layer with a few strata brings immediately large gains in sample size. More complex stratifications usually - but not necessarily - produce comparable gains with respect to more elementary ones. This depends obviously on the relative efficiency of a deeper stratification to compress the variance within strata. In the case of air transports, for example, adding the type of carrier to the country of destination increases by nearly 40% the number of strata but does not seem to generate very large gains, at least with the Jevons or Dutot aggregation. The consideration of more detailed areas of destination produces instead quite important gains, since it probably better reflects pricing criteria.

With no or only a few strata, Jevons aggregation performs significantly better in terms of sample size, while Carli and Dutot approaches are more affected by heterogeneity. Deeper stratifications tend to reduce the differences between the alternative aggregation criteria. In the case of air transports, the coefficient of variation of the sample sizes necessary for monthly precision drops from 35% with SRS to nearly 5% with 51 strata. For package holidays, analogous results are produced. These effects depend on the higher homogeneity of price level and price changes within strata in more stratified designs. The issue of the choice of the aggregation method, which has been deeply discussed in literature, loses importance as stratification is considered, especially with highly stratified samples. In particular, empirical evidence suggests that the Jevons aggregation loses part of its advantages when the target is set on the whole series of monthly indices, due to a less favorable allocation of the units across strata (sections 4.2 and 4.3).

The monthly modulation of sample size according to the expected variability of price dynamics appears a very important tool in order to reduce the survey costs and the burden on respondents. Strong efforts for price collection can be concentrated in the price reference month, while monthly samples can be drastically reduced. If we pass from monthly targets to quarterly targets considerable gains in sample size are obviously obtained, although large differences among aggregation methods persist unless we consider highly stratified designs. They persist also if the target is moved on the yearly average or on the annual link. Considering the latter as main target may imply a large gain in sample size, as it happens for air transports; but if the link month is among those showing a higher variability (as in the case of package holidays) this objective may not produce large enough gains.

More general conclusions can finally be drawn regarding HICP concepts and methodology. The results derived in this work are based on several restrictive hypotheses on the dynamics of the set of the available product-offers (time invariant, with no changes in the range of the products and in the retail network). Such hypotheses were essential in order to provide a reliable definition of the statistical target and a one-to-one mapping of the set of the transactions in the weight reference period into the sets of product-offers available in the price reference and in the reporting month. Relaxing these hypotheses implies in fact a huge modeling of consumers' choices in order to produce more sophisticated mapping functions. Further developments on these issues might be obtained both on the theoretical and empirical grounds. Concerning the first, the pioneering work of Ribe (2000) deserves more analysis on the form and nature of the mapping functions and their implications, especially with reference to the structural characteristics of consumer markets. It could be fruitful to consider different classes of mapping functions to be used in particular clusters of consumer markets. Empirical studies might help in this work, by examining other sectors and by providing deeper insights on the relative efficiency of alternative stratification criteria.

REFERENCES

Adelman, I. (1958), "A New Approach to the Construction of Index Numbers," *Review of Economics and Statistics*, 40, 240-249.

Balk, B. M. (2003), "Price Indexes for Elementary Aggregates: the Sampling Approach," *paper presented at the Seventh Meeting of the International Working Group on Price Indices*, Paris, 27-29 May.

--------------- (2008), *Price and Quantity Index Numbers. Models for Measuring Aggregate Change and Difference*. Cambridge university press, New York, p.283.

Banerjee, K. S. (1956), "A Note on the Optimal Allocation of Consumption Items in the Construction of a Cost of Living Index," *Econometrica*, Vol. 24, No. 3 (Jul., 1956), p. 294-295.

Baskin, R. M., Leaver, S. G. (1996), "Estimating the Sampling Variance for Alternative Forms of the U.S. Consumer Price Index," *proceedings of the Survey Research Methods Section, American Statistical Association*, p. 192-197.

Biggeri, L., Falorsi, P. D. (2006), "A Probability Sample Strategy for Improving the Quality of the Consumer Price Index Survey Using the Information of the Business Register," *paper presented at the Joint ECE/ILO meeting of the of Group of Experts on Consumer Price Indices*, Working paper n.12, Geneva 10-12 May

Cochran, W. G. (1977), *Sampling Techniques*. 3rd edition. John Wiley and sons, New York, p. 428.

Dalèn, J. (1992), "Computing Elementary Aggregates in the Swedish Consumer Price Index," *Journal of Official Statistics*, Vol. 8, No.2, 129-147.

---------- (1998) “On the Statistical Objective of a Laspeyres’ Price Index,” *paper presented at the Fourth Meeting of the International Working Group on Price Indices*, Washington D.C., 22-24 April.

---------- (2001), “Statistical Targets for Price Indexes in Dynamic Universes,” *paper presented at the Sixth Meeting of the International Working Group on Price Indices*, Canberra, Australia, 2-6 April.

Dalèn, J., Ohlsson, E. (1995), “Variance Estimation in the Swedish Consumer Price Index,” *Journal of Business & Economic Statistics,* July 1995, Vol.13, No. 3.

De Gregorio, C. (2010), “The Variability of HICP Sub-Indices and Harmonisation Needs,” paper submitted to the Rivista di statistica Ufficiale.

De Gregorio, C., Fatello, S., Lo Conte, R., Mosca, S., Rossetti, F. (2008), “Sampling Design and Treatment of Products in ISTAT Centralised CPI Surveys,” *Contributi ISTAT,* n. 1.

De Gregorio, C., Munzi, C., Zavagnini, P. (2008), “Problemi di Stima, Effetti Stagionali e Politiche di Prezzo in Alcuni Servizi di Alloggio Complementari: Alcune Evidenze dalle Rilevazioni Centralizzate dei Prezzi al Consumo,” *Contributi ISTAT*, n.6.

De Haan, J., Opperdoes, E., Schut, C. (1997), “Item Sampling in the Consumer Price Index: A Case Study Using Scanner Data,” *paper presented at the Joint ECE/ILO meeting of the Group of Experts on Consumer Price Indices*, Working paper n.1, Geneva 24-27 November.

Dippo, C. S., Wolter, K. M. (1983), “A Comparison of Variance Estimators Using the Taylor Series Approximation,” *proceedings of the Survey Research Methods Section, American Statistical Association*, p. 113-119.

Dorfman, A. H., Lent, J., Leaver, S. G., Wegman, E. (2006), “On Sample Survey Designs for Consumer Price Indexes,” *Survey Methodology*, December, Vol. 32, No. 2, p. 197-216.

EUROSTAT (2001), *Compendium of HICP Reference Documents*. European communities, p. 354.

------------- (2003), “Outline of a Possible Strategy for Harmonizing HICP Lower Level Aggregation,” *paper presented at the Meeting of the Working Party "Harmonization of Consumer Price Indices"*, EUROSTAT, Luxembourg, 9-10 December.

Fenwick, D. (1998), “The Impact of Sample Design on the Performance of the Sample Geometric Mean and Related Issues,” *paper presented at the Fourth Meeting of the International Working Group on Price Indices*, Washington D.C., 22-24 April.

-------------- (1999), “The Impact of Choice of Base Month on the Relative Performance of Different Formulae Used for Aggregation of Consumer Price Index Data at an Elementary Aggregate,” *paper presented at the Fifth Meeting of the International Working Group on Price Indices*, Reykjavik, 25-27 August.

-------------- (2001), “Sampling in Consumer Price Indices: what role for scanner data?,” *paper presented at the Sixth Meeting of the International Working Group on Price Indices*, Canberra, 2-6 April.

Hansen, C. B. (2006), “Price Updating of Weights in the CPI,” *paper presented at the Ninth Meeting of the International Working Group on Price Indices*, London, 14-16 May.

Koskimäki, T., Ylä-Jarkko, M. (2003), “Segmented Markets and CPI Elementary Classifications,” *paper presented at the Seventh Meeting of the International Working Group on Price Indices*, Paris, 27-29 May.

Kott, P. S. (1984), “A Superpopulation Theory Approach to the Design of Price Index Estimators With Small Sampling Biases,” *Journal of Business & Economic Statistics*, Vol.2, No.1, January.

ILO (ed.) (2004), *Consumer Price Index Manual: Theory and Practice*. Geneva

ISTAT (2009), *Come si Rilevano i Prezzi al Consumo*. www.istat.it

Leaver, S. G., Johnstone, J. E., Archer, K. P. (1991), "Estimating Unconditional Variances for the U.S. Consumer Price Index for 1978-1986," *proceedings of the Survey Research Methods Section, American Statistical Association*, p. 614-619.

Leifer, H.-A. (2002), "The 'Elementary Level' of a Consumer Price Index and Cost-of-Living Index: Similarities and Differences," *paper presented at the Meeting of the Working Party "Harmonization of Consumer Price Indices"*, EUROSTAT, Luxembourg 30 September-1 October.

--------------- (2008), "Are the Carli, Dutot and Jevons Indices Time-Path Dependent?," *paper presented at the Fifth Meeting of the HICP Task Force Sampling*, EUROSTAT, Luxembourg, 19 February.

Norberg, A. (2004), "Comparison of Variance Estimators for the Consumer Price Index. *paper presented at the Eight Meeting of the International Working Group on Price Indices*, Helsinki, 23-25 April.

Norris, N. (1940), "The Standard Errors of the Geometric and Harmonic Means and Their Application to Index Numbers," *The Annals of Mathematical Statistics*, Vol. 11, No. 4 (Dec.), p. 445-448.

Ribe, M. (2000), "Price Index, Statistical Inference and Harmonisation," *proceedings of the 4th International Conference on Methodological Issues in Official Statistics*. 12-13 October.

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

Viglino, L. (2003), "Harmonisation of HICP lower level aggregation. The use of the geometric mean for EEG and EA aggregation," *paper presented at the Meeting of the Working Party "Harmonization of Consumer Price Indices"*, EUROSTAT, Luxembourg, 9-10 December.

Wilkerson, M. (1967), "Sampling Error in the Consumer Price Index," *Journal of the American Statistical Association*, Vol. 62, No. 319 (Sep., 1967), p. 899-914