rivista di statistica ufficiale

n. 1
2010

Temi trattati

Istat
Istituto nazionale
di Statistica

rivista di statistica ufficiale

n. 1
2010

Temi trattati

Istat

Per contattare la redazione o per inviare lavori scrivere a:
Segreteria del Comitato di redazione della Rivista di Statistica Ufficiale
All'attenzione di Gilda Sonetti
Istat - Via Cesare Balbo, 16 - 00184 Roma
e-mail: rivista@istat.it

rivista di statistica ufficiale

n. 1/2010

Periodico quadrimestrale
ISSN 1828-1982

Registrazione presso il Tribunale di Roma
n. 339 del 19 luglio 2007

Istituto nazionale di statistica
Servizio Editoria
Via Cesare Balbo, 16 - Roma

Stampato nel mese di aprile 2011
presso il Centro stampa dell'Istat
Via Tuscolana 1788 - Roma
Copie 300

Evaluating reliability of combined responses through latent class models

Marcello D'Orazio[1]

Abstract

The evaluation of the potential impact of the response errors on the final survey estimates requires ad hoc studies. Often these studies consist in additional reinterview surveys: a subsample of the respondent units at the main survey is interviewed again. In such cases, the evaluation can be done by means of the theory introduced by Hansen et al. (1964) and further investigated in Biemer and Forsman (1992). More recent studies (Biemer, 2004) present an approach based on the fitting of latent class models. These models allow for a more detailed analysis of the impact of response errors on the final survey estimates but, on the other hand, they require some additional assumptions to hold. In this paper, the usage of latent class models is extended to tackle the case of couples of survey questions involved in a questionnaire skip. An application of such models to the data of the control survey on the 2001 Population and Housing Census in presented.

Keywords: Response Errors, Simple Response Variance, Latent Class Models, Questionnaire Skip

1. Introduction

A measurement error consists in the differences between the value observed for a variable and the true value of this variable for the investigated unit. The measurement errors arising in the survey data collection are called *response errors*. When dealing with categorical variables these errors are also known as *misclassification errors*. In order to deal with the response errors, Hansen *et al.* (1964) introduced a simple model widely used at the US Bureau of the Census (1985). They showed that the measurement errors represent a source of additional variability and may introduce a bias (*response bias*) when estimating the population characteristics. The additional variability can be decomposed into the *simple response variance* (SRV) and the *correlated response variance* (CRV). The first term reflects the variability of the responses that can be collected for a single question in a series of repetitions of the data collection on the same unit. The second term reflects the correlation of the responses collected for the same question on different units in a given survey occasion. A well known source of CRV is represented by the interviewers. Usually, in a self-administered interview the CRV is assumed to be zero.

A common practice for estimating the bias and the variance due to response errors consists in carrying out a reinterview study. A reinterview survey with the purpose of obtaining error-free responses (*gold standard*) permits to estimate the response bias. On the

[1] Researcher (Istat), e-mail: madorazi@istat.it.

other hand, an independent survey consisting in a perfect replication of the main survey on a subset of the responding units (*test-retest reinterview*) has to be carried out to estimate of the SRV (for details see Biemer and Lyberg, 2003, pp. 291-301). In most of the cases, a reinterview survey consists in a combination of both (cf. Biemer and Forman, 1992).

In this work the focus is on the usage of the reinterview survey data in order to estimate the SRV for categorical variables. The paper is structured as follows: Section 2 introduces the approach used at the US Bureau of the Census in order to evaluate the impact of SRV on the final survey estimates. Some results obtained using this approach with the data of the quality control survey of the 2001 Population and Housing census in Italy are reported in Section 2.1. In Section 3 it is introduced the problem of couples of questions involved in a questionnaire skip. Section 4 provides a summary of the theory underlying the LC models for the analysis of repeated measurements. In Section 4.1 it is presented a relatively new approach, based on the application of Latent Class (LC) models, with the objective of evaluating the reliability of the couples of variables involved in a questionnaire skip. Section 4.2 reports the results obtained applying this new approach to the data of the quality control survey of the 2001 Census.

2. Evaluating reliability: the US Bureau approach

In the approach of the US Census Bureau (Hansen *et al.* 1964; US Census Bureau 1985; Biemer and Forsman, 1992; Biemer, 2004) a key role is played by the *index of inconsistency*:

$$I = \frac{\text{SRV}}{\text{SV} + \text{SRV}} \tag{2.1}$$

It represents the proportion of total variance (SV is the sampling variance) due to response errors (cf. Biemer, 2004), hence $0 \le I \le 1$. It is worth noting that the quantity $R = 1 - I$ is known as *reliability ratio*. Once estimated *I*, the following rule of thumb can be considered (US Census Bureau, 1985, p. 95):

$0 \le \hat{I} \le 0.20$	high reliability (low inconsistency)
$0.20 < \hat{I} < 0.50$	moderate reliability (moderate inconsistency)
$0.5 \le \hat{I} \le 1$	low reliability (high inconsistency)

In order to estimate *I*, let consider a categorical variable Y with J ($J \ge 2$) response categories and let assume $y_k^{(1)}$ to be the value observed for Y on the *k*th unit in the main survey ($t = 1$) and $y_k^{(2)}$ the value observed on the same unit in the reinterview survey ($t = 2$). Usually, the reinterview is carried out on a subsample of the respondents to the main survey, selected according to a given probabilistic sampling design $p(s^{(2)})$, being $w_k^{(2)}$ the survey weight of the *k*th unit (inverse of inclusion probability, maybe corrected for unit nonresponse in reinterview survey). In this framework I is estimated by (US Bureau of the Census, 1985, p. 88):

$$\hat{I} = \frac{g}{1 - \sum_{j=1}^{J} \hat{P}_{+j}\hat{P}_{j+}} \tag{2.2}$$

with

$$g = 1 - \sum_{j=1}^{J} \hat{P}_{jj} \tag{2.3}$$

Note that $\hat{P}_{ij} = \hat{N}_{ij} / \hat{N}$ ($i, j = 1,\ldots,J$) are the relative frequencies of the cells of the contingency table $Y^{(2)} \times Y^{(1)}$. In particular, $\hat{N} = \sum_{s^{(2)}} w_k^{(2)}$ and $\hat{N}_{ij} = \sum_{s^{(2)}} w_k^{(2)} C\left(y_k^{(2)} = i, y_k^{(1)} = j\right)$, being $C(\cdot) = 1$ if the condition within parenthesis is satisfied and 0 otherwise. Note that these frequencies are estimated before the editing and imputation phase and all units with missing values (at the main survey, at the reinterview survey or at both) are discarded from the computation.

The quantity g is known as the *gross difference rate* (GDR) or *disagreement rate*. It can be shown that $g/2$ provides an unbiased estimate of the SRV under the assumption of: (A1) equal probabilities of misclassification of the responses at original interview and at reinterview and, (A2) independence between the responses at the main survey and those at the reinterview (absence of *between-trial correlation*) (for details see Biemer and Forsman, 1992; Biemer 2004). They are assumed to hold when the reinterview survey is a perfect independent replication of the main survey.

When the assumption (A1) does not hold, then $g/2$ does not provide an unbiased estimate of the SRV at $t = 1$ and, consequently, the estimated I is unreliable. When both the assumptions do not hold, the negative bias introduced by the presence of positive between-trial correlation (failure of A2) tends to determine an underestimation of the SVR. On the contrary, when only (A2) holds, the sign of the bias is determined by the relationship between misclassification probabilities at the two survey occasions. If misclassification probabilities at the control survey ($t = 2$) are lower than those at the original survey ($t = 1$), then an underestimation of SRV at $t = 1$ it is expected. Overestimation occurs in the opposite situation (cf. Biemer and Forsman, 1992, pp. 919-920).

Note that in case of binary variables ($J = 2$), the assumption (A1) (equal error distribution) corresponds to assuming $H_0 : P^{(t=1)} = P^{(t=2)}$ (cf. Biemer, 2004, p. 430). In the more general case ($J > 2$) it corresponds to assuming the *marginal homogeneity* (cf. D'Orazio, 2008).

Finally, it is worth noting that $\hat{I} = 1 - \kappa$, being κ the Cohen's *kappa* measure of reliability (cf. Biemer, 2004, p. 423).

2.1 Reliability in 2001 Population and Housing Census in Italy

The quality evaluation program of the 2001 Population and Households Census (CEN) consisted in a single control survey, referred as Post Enumeration Survey (PES), aimed at estimating the impact on Census estimates of both coverage and measurement errors.

The PES was based on a stratified two stage sample of approximately 1,100 census Enumeration Areas (EAs), located into 98 sample municipalities (the Primary Sampling Units). In each sample area a new complete enumeration of the households was carried out. At the end of the data collection, approximately 68,000 households and about 180,000 individuals were surveyed. A complex procedure of record linkage allowed the units covered by both the CEN and the PES to be identified. In particular, in order to evaluate the impact of response errors, $n = 172{,}620$ linked individuals were considered, out of the 182,519 people found in the sample EAs at the CEN. The linkage procedure was very successful given that linkage rate was quite high (CEN linkage rate was 95%) and some consistency checks led to the conclusion that the false links (couples of units erroneously linked) had a very low chance of occurring. This is an important result because the presence of false links may negatively affect the estimation of the response variance as shown in Brancato *et al.* (2004).

In order to compare the responses provided by the same individuals at the two survey occasions (CEN and PES) the questionnaire of the PES was enlarged with a subset of questions (about fifteen) selected from those in the census form. The PES data were collected by means of a self-administered paper questionnaire - as for the CEN - in the period November-December 2001, about a month after the CEN data collection (CEN reference date is 21st of October 2001). The overlapping with the CEN field operations was accurately avoided. In practice, the PES data collection was designed as a perfect replication of the census one (test-retest reinterview) in order to fulfill assumption (A1). Given the time lag between the two surveys, the between-trial correlation, due to the situation of respondents that at the PES recall the answers provided at the CEN and repeat them, can be considered negligible (assumption A2 holds). Biemer and Lyberg (2003, pp. 298-299) report 5-10 days to be a sufficient time lag for carrying out an independent reinterview.

Table 1 reports the GDRs and the estimated values for I for some of the most important variables (cf. Istat, 2009, pp. 121-122).

Table 1 - GDR and estimates of *I* for some of the variables.

Census Questions	Number of response categories	GDR (x 100)	$\hat{I}$ (x 100)
Relationship with the household head	16	5.25	7.40
Gender	2	0.81	1.62
Age *(in classes)*	16	2.19	2.34
Marital status	6	1.80	3.04
Education level *(Age > 6)*	16	12.72	15.71
Labour status *(Age > 14)*	10	12.09	16.24
Full-time/part-time occupation (*for who declared to have an occupation*)	2	3.56	20.97
Professional status (*for who declared to have an occupation*)	6	7.79	18.33
Limited/unlimited duration of contract (*for occupied as employees*)	2	5.76	28.95

Reliability is quite high when dealing with the first four variables. As expected, Gender is the most reliable variable. The estimates of I increase starting from the question concerning the education level. According to previously mentioned rule of thumb, reliability is always "high" ($\hat{I} \leq 0.20$) with the exception of the question "full time/part time" occupation and for the one concerning the duration of the contract (for those who declared to be occupied as employees).

3. Evaluating the reliability for questions involved in a questionnaire skip

The census form had several *filter questions*. A filter question is crucial because the answer to it determines the response path, i.e. which section of the questionnaire has to be filled in and which one has to be skipped. For instance, the question related to the professional status had to be filled in only by those individuals who responded to be occupied at the question concerning the labour status. A complex questionnaire can contain several filter questions; hence, assessing the reliability of these key questions is crucial in order to assess the overall reliability. On the other hand, evaluating the reliability of the questions which depends on the responses to the filter questions may give rise to some difficulties. For simplicity, let consider the case of two binary variables, X and Y, such that: (i) an answer to Y (i.e. $y=1$ or $y=2$) is due if and only if $x=1$ and, (ii) Y should be skipped (i.e. $y=\text{"Not Applicable"}$) if $x=2$. In the ideal situation of absence of errors in the skip from one question to another one (e.g. when the data collection is assisted by PC and the questionnaire skips are managed by the software) the following combinations of responses are admitted:

X	Y
1	1
1	2
2	NA

Unfortunately, when dealing with self-administered paper questionnaire, the errors in the questionnaire skip have to be taken into account. In the previous example, when evaluating the reliability of the Y variable there are two possibilities: (a) consider all the cases without caring of the values of the X variable and, (b) limit the attention to the subset of the cases with $x^{(2)}=x^{(1)}=1$. In the case (a) the disagreements between responses to the filter question (X) are implicitly considered when evaluating the disagreement for Y variable. On the other hand, in the case (b), by considering only the disagreements on Y conditioning to an agreement on the X variable, the response errors in the X are discarded. In this latter case the GDR is computed using:

$$g_{Y|X1=X2=j} = \frac{\sum_{k=1}^{n} w_k^{(2)} C(y_k^{(1)} \neq y_k^{(2)} | x_k^{(1)} = x_k^{(2)} = j)}{\sum_{k=1}^{n} w_k^{(2)} C(x_k^{(1)} = x_k^{(2)} = j)} \tag{3.1}$$

and can be referred as a *"net" disagreement rate* on the *Y* variable. In practical situations, the two options (a) and (b) lead to similar GDR values, unless there are many errors in the questionnaire skips (in absence of such errors they provide the same estimate).

In the PES survey, when computing the GDRs (see Table 1) for the variables depending on a filter question, the formula (3.1) was used. In any case, by comparing the GDRs under option (a) (i.e. all the units are considered) and (b) the "net" case, similar values come out as shown in Table 2, thus denoting the presence of few errors in the skip among the questionnaire sections.

Table 2 - GDR and estimates of *I* for some variables in the Census form involved in questionnaire skips.

Variables	Number of categories	GDR (x100) All units	GDR (x100) "net"	$\hat{I}$ (x 100) All units	$\hat{I}$ (x 100) "net"
Education level *(Age > 6)*	16	12.79	12.72	15.79	15.71
Labour status *(Age > 14)*	10	12.01	12.09	16.01	16.24
Full-time/part-time occupation (*for who declared to have an occupation*)	2	3.69	3.56	21.03	20.97
Professional status (*for who declared to have an occupation*)	6	8.00	7.79	18.65	18.33
Limited/unlimited duration of contract (*for occupied as employees*)	2	6.04	5.76	28.77	28.95

A more detailed investigation on the relationship between the GDRs of two variables, *X* and *Y*, involved in a questionnaire skip can be found in D'Orazio (2008).

4. Latent Class Models in presence of repeated measurements

Biemer (2004) shows how to use LC models to investigate the reliability of responses when dealing with a categorical variable. These models allow the assumption of equal probabilities of misclassification in the two survey occasions to be relaxed but, on the other hand, their application relies on the following assumptions:

(B1) the misclassification probabilities at both survey occasions do not vary among individuals:

$$\phi_{kij}^{(t)} = \Pr\left(y_k^{(t)} = i \middle| y_k = j\right) = \Pr\left(y^{(t)} = i \middle| y = j\right) = \phi_{ij}^{(t)} ,$$

for $t = 1,2$, $k = 1,\dots,n$, $i, j = 1,\dots,J$, being y_k the true unobserved (*latent*) *Y* classification of the *k*th unit;

(B2) the *local independence* holds:

$$\Pr\left(y^{(1)} = i,\ y^{(2)} = j,\ y = h\right) = \\ = \Pr\left(y^{(1)} = i \middle| y = h\right) \times \Pr\left(y^{(2)} = j \middle| y = h\right) \times \Pr\left(y = h\right); \quad i, j, h = 1,\dots,J$$

In the traditional approach to LC models, if the model is identifiable, i.e. the maximum value of the likelihood exists and is unique (for details see Goodman, 1974; Huang, 2005) the ML estimates of the cell probabilities involved in (B2) are derived using well known iterative algorithms such as Newton-Raphson or EM. Then, these estimates can be combined to get an estimate of SRV and I in correspondence of each survey occasion (cf. Biemer, 2004, p. 427). As far as identifiability of the model is concerned, it is worth noting that the factorization in (B2) involves $(J-1)\times(2\times J+1)$ parameters but, given that the observed table of $Y^{(2)}\times Y^{(1)}$ has only $(J\times J-1)$ degrees of freedom, the model is not identifiable (the number of parameters is greater than the degrees of freedom), unless a new additional auxiliary variable G is introduced, with the constraint that the error probabilities $\phi_{ij}^{(t)}$ are equal across the groups identified by the categories of G (cf. Biemer, 2004, pp. 425).

Given that the application of LC models does not require the assumption (A1) of equal misclassification probabilities at both survey occasions, they can be applied to the general situation of reinterview studies carried out in different conditions with respect to the main survey. Moreover, fitting a LC model provides an estimate of the true unobserved $\Pr(y=h)$, hence it is possible to derive an estimate of the response bias, without having to carry out a reinterview study aimed at ascertaining the gold standard. For instance, an estimate of the response bias at the main survey ($t=1$) is:

$$\widetilde{B}_P^{(t=1)}=\hat{\Pr}\left(y^{(1)}=h\right)-\hat{\Pr}_{\mathrm{LC}}\left(y=h\right)=\frac{\hat{N}_h^{(t=1)}}{\hat{N}}-\hat{\Pr}_{\mathrm{LC}}\left(y=h\right). \tag{4.1}$$

Unfortunately, the usage of a LC model poses some problems. The crucial assumption of the local independence (B2) may not hold; in this case it has to be relaxed by resorting to one of the approaches proposed in literature (see e.g. Hagenaars, 1988; Vermunt, 1997). Assumption (B1) can also be relaxed by introducing the dependence of the distributions of both latent and observed variables on a set of individual covariates (see e.g. Huang *et al.*, 2004). In both cases the risk is that of increasing too much the complexity of the model, thus affecting its identifiability. Another known problem when fitting LC models is represented by local maxima: the iterative algorithm may stop at local maximum and therefore the resulting estimates for the parameters would not be the ML ones. Usually, the chance of having local maxima increases with increasing number of categories of the latent variable. Finally, the ordering of the estimated latent classes is arbitrary and may become difficult to identify the effective response category associated to each latent class (likely to happen with a high number of latent classes).

It is worth noting that in the traditional LC framework, the ML estimates of the parameters are derived by assuming i.i.d. observations, therefore when the reinterview study is based on a complex sample (with stratification and clustering) the ML estimates can be considered valid under the further assumption that the sampling design plays no role in the inference (*model-based inference*, cf. Särndal *et al.*, 1992, pp. 513-520). If the sampling design can not be ignored, it has to be taken into account jointly with the sampling weights. This can be done by resorting to one of the approaches available in literature such as the *Pseudo-ML* estimation (cf. Patterson *et al.* 2002) or to the two-step approach suggested by Vermunt (2002). Vermunt and Magidson (2007) compared these and other approaches to fit LC models when dealing with complex samples.

4.1 Latent class models to evaluate reliability for coupled variables involved in questionnaire skips

Let consider the simple case of two binary variables, *X* and *Y*, such that: (i) an answer to *Y* (i.e. $y=1$ or $y=2$) is due if and only if $x=1$ and, (ii) *Y* should be skipped (i.e. $y=\text{"Not Applicable"}$) if $x=2$. In the census questionnaire there were a number of situations situation that could be summarised in such a simple situation. Two examples are reported in Tables 3a and 3b.

Table 3a - Relationship between the labour status and the type of occupation.

X="Labour Status"	*Y="Type of occupation"*
1 = "With occupation"	1 = "Full time"
1 = "With occupation"	2 = "Part time"
2 = "Without occupation"	NA = "Not Applicable"

Table 3b - Relationship between the professional status and the duration of contract (for those with an occupation).

X="Professional Status"	*Y="Duration of contract"*
1 =""Employee"	1 = "Unlimited duration"
1 = "Employee"	2 = "Limited duration"
2 = NOT "employee" (self-employed, family worker,..)	NA = "Not Applicable"

Let *Z* be the variable obtained by combining the response categories of *Y* and *X*. Following (i) and (ii), only three categories are admitted for *Z*, as can be seen in Table 4 (last column).

Table 4 - Response categories obtained by combining *X* and *Y*.

$X^{(t)}$	$Y^{(t)}$	$Z^{(t)}$	Z
1	1	1	1
1	2	2	2
1	"NA"	4	-
2	1	5	-
2	2	6	-
2	"NA"	3	3

If response errors in the skip pattern are considered, then the variable $Z^{(t)}$, obtained by combining responses to $X^{(t)}$ and to $Y^{(t)}$, can admit up to 6 categories (third column in the Table 4). Therefore, the contingency table $Z^{(2)}\times Z^{(1)}$, built before the editing and imputation phase, may show more than the expected 9 ($=3\times3$) cells. This is the case, for instance, of the variables in Table 3a observed at the CEN ($t=1$) and at the PES ($t=2$).

Table 5 - Occupation and type of occupation (estimated relative frequencies x 100).

Reinterview		Census						
		With occupation = "Yes"			With occupation = "No"			
With occupation	Full/part time	1	2	NA	1	2	NA	Total
1 = "Yes"	1= "Full-time"	33.465	0.925	1.293	0.103	0.034	1.025	36.844
	2= "Part-time"	0.429	2.810	0.104	0.008	0.081	0.287	3.720
	NA	0.249	0.030	0.027	0.002	0.002	0.043	0.352
2 = "No"	1= "Full-time"	0.048	0.003	0.004	0.010	0.002	0.113	0.182
	2= "Part-time"	0.006	0.013	0.001	0.001	0.017	0.037	0.076
	NA	0.794	0.334	0.108	0.121	0.231	57.238	58.827
	Total	34.991	4.116	1.537	0.245	0.367	58.744	100.000

Note: NA corresponds to "Not Applicable" or "Not Answer"

Table 5, due to response errors and to errors in the skip pattern, has 36 non-empty cells instead of the expected 9. In any case, the frequencies estimated for the cells relating erroneous skips are very close to 0.

This undesired situation may turn out useful in fitting LC models. In fact let consider the variable *Z*, and assume that (B1) and (B2) hold for *Z* too. Then, the factorization

$$\Pr\left(z^{(1)} = i,\, z^{(2)} = j,\, z = h\right) = \\ = \Pr\left(z^{(1)} = i \middle| z = h\right) \times \Pr\left(z^{(2)} = j \middle| z = h\right) \times \Pr\left(z = h\right) \tag{4.2}$$

($i, j = 1,2,\ldots,6$, and $h = 1,2,3$) involves 32 ($= 5\times 3 + 5\times 3 + 2$) free parameters (probabilities) while the starting contingency table (e.g. Table 5) has 35 ($= 6\times 6 - 1$) degrees of freedom. In practice, the degrees of freedom are enough to make the model identifiable without having to introduce a grouping variable as in the case fitting a LC to *X* or to *Y*. This condition is necessary but not sufficient to ensure identifiability; to this purpose it is required that the model admits unique ML estimates of the unknown parameters too.

A further advantage in fitting a LC model to $Z^{(t=2)} \times Z^{(t=1)}$ consists in using the estimates of the probabilities in (4.2), to derive the estimates of *P*, the true population proportion of a given characteristics, and I for both the variables, *X* and *Y*, involved in the questionnaire skip. In other words, a single LC model is fitted instead of two separate models. It is worth noting that similar results can be obtained by fitting *path model* with latent variables to the repeated observations for the starting variables *X* and *Y*.

4.2 An application of the LC model to couples of variables from the census control survey

The theory presented in the previous Section has been applied to the data from the 2001 census control survey. In particular, repeated measurements for the couples of variables in the Tables 3a and 3b have been considered. Note that, due to the complex sampling design underlying the PES survey, the LC models has been applied starting from the contingency table $Z^{(t=2)} \times Z^{(t=1)}$ estimated by summing up the sample weights (scaled in order to sum up to the total sample size) of the cases falling in each cell of the observed table. This approach corresponds to *pseudo-Maximum Likelihood estimation* (cf. Patterson *et al.*, 2002; Vermunt and Magidson, 2007). Note that the identifiability of the model has been verified empirically by checking that the same ML estimates for the parameters are obtained by running the iterative estimation procedure with different starting values.

Table 6 shows the estimates obtained for P and I .

Table 6 - Results obtained applying LC models to the data of examples 3a and 3b.

	P estimates (x 100)					*I* estimates (x 100)		
Variables	CEN $t = 1$	PES $t = 2$	NDR	LCM	Bias at CEN	Standard formula	LCM CEN	LCM PES
X=1 (“With occupation”)	41.99	42.34	-0.35	42.95	-0.96	5.81	6.85	4.76
Y=1 (“Full-time”)	89.92	91.34	-1.42	88.19	1.73	20.97	13.82	27.18
X=1 (“Employee”)	73.98	75.00	-1.02	74.58	-0.60	8.34	8.97	7.63
Y=1 (“Unlimited contract”)	87.68	89.95	-2.27	85.01	2.67	28.95	20.28	36.20

As far as *P* estimates are concerned, it is interesting to observe that the *net difference rate* ($\mathrm{NDR} = \hat{P}_j^{(t=1)} - \hat{P}_j^{(t=2)}$,) is always negative; it is larger for the variables denoted as *Y*. Due to response errors, a negative response bias comes out when estimating $\Pr(x=1)$. The estimated bias, on the contrary, is positive when $\Pr(y=1)$ is concerned. The bias for *Y* variables is larger, in absolute terms, than that associated to *X*.

Finally, all the estimates of *I* concerning the *X* variables tend to be close, while marked differences emerge for the *Y* variables. The estimates of *I* obtained using the LC models, suggest a higher reliability of *Y* at CEN. This result is in contrast with the evidence from various studies that found lower misclassification probabilities in control surveys than in the main survey (cf. Biemer and Forsman, 1992, p. 920). As far as *Y* variables are considered, the discrepancies found in the estimates of *I*, jointly with the differences found for the NDR and the estimated response bias, lead to conclude that the assumption of equal misclassification probabilities (B1) may not hold in this case and, as a consequence, the estimates of SRV and *I* obtained with standard methods (formulas (2.2) and (2.3)) are misleading.

5. Conclusions

The study of methodologies involving the usage of LC models is important in order to evaluate the reliability when dealing with complex questionnaires with many skip patterns. The approach based on the LC models has the advantage of considering all the available data, including also not coherent answers according to the skip pattern in the questionnaire. Moreover, the estimates of the response bias, SRV and *I*, for each of the variables involved in a questionnaire skip are derived just by applying a single model to a new variable obtained by combining the answers to couples of questions. In addition, there is no need to resort to an additional grouping variable to make the model identifiable. Finally, LC models, providing an estimate of the true probabilities of an event, permit the estimation of the response bias, usually not allowed in the classic approach based on "test-retest" reinterview.

On the other hand, the usage of LC models presents some well known drawbacks. A crucial assumption is the local independence; it can be relaxed at the cost of increasing the complexity of the model. The same happens as far as homogeneity (assumption B1) is concerned. Moreover, when dealing with complex survey data, different approaches are available. The pseudo-ML provides good estimates of the parameters but it does not permit to evaluate the goodness of fit of the model using the standard tools (cf. Vermunt and Magidson, 2007). All these problems require further investigation in order to widely apply the LC models to evaluate responses reliability in complex surveys that collect data using questionnaires with many filter questions.

References

Biemer P.P. (2004), "The Twelfth Morris Hansen Lecture. Simple response Variance: Then and Now", *Journal of Official Statistics*, 3, pp. 417-439.

Biemer P.P., Forsman G. (1992), "On the quality of reinterview data with application to the current population survey", *Journal of the American Statistical Association*, 87, pp. 915-923.

Biemer P.P, Lyberg L.E. (2003), *Introduction to Survey Quality*, Wiley, New York.

Brancato G., D'Orazio M., Fortini M. (2004), "Response Error Estimation in Presence of Record Linkage Errors: The Case of the Italian Population Census". *Proceeding of the European Conference on Quality and Methodology in Official Statistics* (Q2004), Mainz, Germany, 24-26 May 2004.

D'Orazio M. (2008), "Reliability of responses at the 14th Italian Population and Housing Census in 2001". *Proceedings of Q2008 European Conference on Quality in Official Statistics*, Rome, 8-11 July 2008.

Goodman L.A. (1974), "Exploratory Latent Structure Analysis Using Both Identifiable and Unidentifiable Models", *Biometrika*, 61, pp. 215-231.

Hagenaars J.A. (1988), "Latent Structure Models with Direct Effects Between Indicators: Local Dependence Models", *Sociological Methods & Research*, 16, pp. 379-405

Hansen M., Hurwitz W.N., Pritzker L. (1964), "The Estimation and the Interpretation of Gross Differences and the Simple Response Variance". In C.R. Rao (ed.) *Contributions to Statistics*, Statistical Publishing Society, Calcutta, pp. 111-136.

Huang G.H. (2005), "Model Identifiability". In B.S. Everitt and D.C. Howell (eds.) *Encyclopedia of Statistics in Behavioral Science*, Wiley, Chichester, Society, Calcutta, Volume 3, pp. 1249-1251.

Huang G.H., Bandeen-Roche K. (2004), "Building an Identifiable Latent Class Model with Covariate Effects on Underlying and Measured Variables", *Psychometrika*, 66, pp. 109-132.

Istat (2009), "La qualità dei dati", *Conoscere il censimento, 14° Censimento Generale della popolazione e delle abitazioni*. Istat, Roma.

http://www.istat.it/dati/catalogo/20091217_00/la_qualita_dei_dati_conoscere_il_censimento.pdf

Patterson B.H., Dayton C.M., Graubard B.I. (2002), "Latent Class Analysis of Complex Sample Survey Data: Application to Dietary Data", *Journal of the American Statistical Association*, 97, pp. 721-729.

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

US Bureau of the Census (1985), "Evaluating Censuses of Population and Housing", *Statistical Training Document,* ISP-TR-5, Washington DC.

Vermunt J.K. (2002), "Comment on Latent Class Analysis of Complex Sample Survey Data: Application to Dietary Data", *Journal of the American Statistical Association*, 97, pp. 736-737."

Vermunt J.K., Magidson J. (2007), "Latent Class Analysis with Sampling Weights: A Maximum Likelihood Approach", *Sociological Methods & Research*, 36, pp. 87-111.

The decomposition of the chained price index rate of change: generalization and interpretative effectiveness

Alessandro Brunetti [1]

Abstract

The paper deals with the method of decomposition of the rate of change of a chained price index into the sum of the effects of the item groups covered by the index, as suggested by M. Ribe (1999). The aim is twofold: firstly, to provide a generalization of the envisaged method to the case of the decomposition of the average rates of change of the aggregate index over different time intervals (such as, for example, annual or quarterly average rates of change); secondly, to investigate the formal properties of the decomposition in order to evaluate its interpretative effectiveness as a measure of the impact of the evolution of the prices of different components on the overall inflation.

Keywords: Chained price index, decomposition of the price index rate of change, contribution to inflation.

1. Introduction

In analyzing inflation, it is usually important to evaluate to what extent the development of the overall index is influenced by the price changes of one commodity or a group of commodities. However, the estimation of the single sub-indices contribution to inflation is not straightforward when the overall index is calculated as a chained index. To this aim, a method of decomposition of the monthly and annual rates of change of a chained price index into the sum of the effects of its sub-indices has been suggested by M. Ribe [1999]. In the present paper, we argue that the envisaged method can be usefully generalized in order to cover the case of the decomposition of the average rates of change of the overall index over different time intervals (such as, for example, annual or quarterly average rates of change). Moreover, some formal properties of the decomposition are discussed in order to evaluate its interpretative effectiveness as a measure of the impact of distinct sub-indices on the overall inflation.

The paper is structured as follows: after a brief reference to the methodological underpinnings of the chained price index (section 2), Ribe's method is presented (section 3), where the decomposition of the rates of change is calculated both between two different months of the same year and between the same month of two consecutive years. In section 4, we show how Ribe's formulas can be generalized to the case of average rates of change over time periods of different length. The formal properties of this generalised method are considered in section 5, where we investigate the relationship between the price development of a component on a given time period and its effect on the rate of change of

[1] Researcher (Istat), e-mail: albrunet@istat.it. The author is grateful to Giuseppe Antonio Certomà, Carlo De Gregorio, Roberto Monducci, Mauro Politi and to an anonymous referee for useful comments to a preliminary draft of the paper. Usual disclaimer applies.

the overall index. As a result we argue that, generally, a zero yearly rate of change of a sub-index (or even a zero yearly average rate of change) is neither a necessary nor a sufficient condition for its effect on the rate of change of the overall index to be null in the same time interval. In the concluding section, an application of the decomposition method to the Italian Harmonized Index of Consumer Price is presented.

2. Chain system for the construction of the price index

A fixed based index provides a measure of the average changes in the price levels by means of the series of binary comparisons between a base reference period and other reporting periods in a specified time interval. On the contrary, in the chained index approach, binary indices are used only to assess the change of price going from one period to the following. The evolution of the price levels over time is then obtained by linking the rates of change of the single-periods in a sequence[2].

The main advantage of the chain system for the construction of a price index consists in the possibility of updating its base at each step of the sequence. Therefore, from the interpretative point of view, the chained price index (of the Laspeyres type) represents the changing cost of a basket of products which is fixed at the beginning of each period and renewed from period to period. More precisely, with reference to chained price indices computed monthly by most of the National Statistical Institute, the basket and its weighting structure are fixed at December of every year and kept fixed in the following twelve months[3].

Formally, let $I_{0,y}^{n,y}$ be the price index of month *n* of year *y* expressed in base December *y-1* = 1 (conventionally indicated hereafter as month 0 of year *y*). The index is computed as:

$$I_{0,y}^{n,y} = \sum_{k=1}^{K} \pi^{0,y}(k) \cdot I_{0,y}^{n,y}(k) \qquad n = 1, 2, \ldots, 12$$

where the weights $\pi^{0,y}(k)$ and the price indices $I_{0,y}^{n,y}(k)$ of the K components of the overall index are given by:

$$\pi^{0,y}(k) = \frac{q^{0,y}(k) \cdot p^{0,y}(k)}{\sum_{k=1}^{K} q^{0,y}(k) \cdot p^{0,y}(k)}; \qquad I_{0,a}^{n,y}(k) = \frac{p^{n,y}(k)}{p^{0,y}(k)} \qquad k = 1, 2, \ldots, K$$

and $q^{0,y}(k)$ and $p^{0,y}(k)$ are respectively the quantity and the price of the component *k* in month 0 of year *y*.

[2] The chained price index can be considered as an approximation in discrete time of the continuous time Divisia index (see. among others R.D.G. Allen (1975), B.M. Balk (2008), F.G. Forsyth and R.F. Fowler (1981), ILO (2004)).

[3] For more information about the adoption of the chained price index by the Italian Statistical Institute, see Quaranta V., Di Iorio F. (1997).

Assuming that year $(y-j)$ is arbitrarily chosen as the base year (henceforth BY), the series of chained price indices are calculated as follows:

$$I_{BY}^{n,y-j} = \frac{I_{0,y-j}^{n,y-j}}{\overline{I}_{0,y-j}^{y-j}} \quad \text{where} \quad \overline{I}_{0,y-j}^{y-j} = \frac{1}{12}\sum_{n=1}^{12} I_{0,y-j}^{n,y-j}$$

and

$$I_{BY}^{n,y-j+1} = I_{BY}^{12,y-j} \cdot I_{0,y-j+1}^{n,y-j+1} \qquad n = 1,2,\ldots,12$$

$$I_{BY}^{n,y-j+2} = \left(I_{BY}^{12,y-j} \cdot I_{0,y-j+1}^{12,y-j+1}\right) \cdot I_{0,y-j+2}^{n,y-j+2} = I_{BY}^{12,y-j+1} \cdot I_{0,y-j+2}^{n,y-j+2} \qquad n = 1,2,\ldots,12$$

$$\ldots\ldots \quad \ldots\ldots \quad \ldots\ldots$$

$$I_{BY}^{n,y} = \left(I_{BY}^{12,y-j} \cdot I_{0,y-j+1}^{12,y-j+1} \cdot I_{0,y-j+2}^{12,y-j+2} \cdot \ldots \cdot I_{0,y-1}^{12,y-1}\right) \cdot I_{0,y}^{n,y} = I_{BY}^{12,y-1} \cdot I_{0,y}^{n,y} \qquad n = 1,2,\ldots,12$$

It is useful to draw attention to the well-known fact that indices expressed in their reference base do not satisfy the additive property. That is[4]:

$$I_{BY}^{n,y} = I_{BY}^{12,y-1} \cdot I_{0,y}^{n,y} \neq \sum_{k=1}^{K} \pi^{y}(k) \cdot I_{BY}^{n,y}(k) = \sum_{k=1}^{K} I_{BY}^{12,y-1}(k) \cdot \pi^{y}(k) \cdot I_{0,y}^{n,y}(k) \,.$$

3. Ribe's decomposition of the monthly and annual rates of change of a chained price index

We start by considering the decomposition of the rates of change calculated between two months of the same year, and then the attention will be drawn to the case of price changes between the same month of two adjacent years.

Let ${}_{m,y}\Delta_{n,y}$ be the rate of change of the overall index between months *m* and *n* of the same year *y*, or alternatively in the period [(*m,y*); (*n,y*)]:

$${}_{m,y}\Delta_{n,y} = \frac{I_{BY}^{n,y}}{I_{BY}^{m,y}} - 1 = \frac{\sum_{k=1}^{K} \pi^{y}(k) \cdot \left[I_{0,y}^{n,y}(k) - I_{0,y}^{m,y}(k)\right]}{I_{0,y}^{m,y}}$$

where $0 \leq m < n \leq 12$.

[4] The equality holds only in the special case in which the links of the *K* components of the aggregate index, as well as the link of the aggregate index, are all at the same level. Notably, in order to minimize the notation, concerning the weights, the reference to the base 0 will be omitted from now on.

It is then clearly apparent that:

$$ {}_{m,y}\Delta_{n,y} = \sum_{k=1}^{K} \varepsilon^{m,y}(k) \cdot {}_{m,y}\Delta_{n,y}(k) \tag{1} $$

in which $\varepsilon^{m,y}(k) \equiv \pi^{y}(k) \cdot \dfrac{I_{0,y}^{m,y}(k)}{I_{0,y}^{m,y}}$.

The interpretation of (1) is straightforward: the rate of change of the overall index measured between months *m* and *n* can be expressed as the weighted arithmetic mean of the rates of change of its sub-indices, with weights given by $\varepsilon^{m,y}(k)$. Accordingly, it is possible to define a measure of the effect of the change of the price of the component *k* on the overall index[5]:

$$ {}_{m,y}C_{n,y}(k) = \pi^{y}(k) \cdot \frac{I_{0,y}^{m,y}(k)}{I_{0,y}^{m,y}} \cdot {}_{m,y}\Delta_{n,y}(k) \tag{2} $$

In other words, ${}_{m,y}C_{n,y}(k)$ provides a measure of the contribution of the sub-index *k* to the dynamics of the overall index, in the period considered.

It is worth noting that if, and only if, $I_{0,y}^{m,y}(k) = I_{0,y}^{m,y}$, the effect of the component *k* can be calculated by multiplying its corresponding rate of change and weight. Specifically, this case occurs when $m = 0$: when the lower bound of the time interval is placed on the month representing the base for the computation of the index of a given year, the equality $I_{0,y}^{m,y}(k) = I_{0,y}^{m,y} = 1$ holds, by definition, for every *k*.

Consider now the yearly rate of change of the overall index:

$$ {}_{n,y-1}\Delta_{n,y} = \frac{I_{BY}^{n,y}}{I_{BY}^{n,y-1}} - 1 = \frac{I_{0,y}^{n,y} \cdot I_{0,y-1}^{12,y-1}}{I_{0,y-1}^{n,y-1}} - 1 $$

It is, then, possible to write:

[5] By putting m = n-1, expression (2) is equivalent to the monthly effect as defined in Ribe (1999) (see page 5 and 6).

$$ {}_{n,y-1}\Delta_{n,y} = \frac{I_{0,y}^{n,y} \cdot I_{0,y-1}^{12,y-1} - I_{0,y-1}^{12,y-1} + I_{0,y-1}^{12,y-1} - I_{0,y-1}^{n,y-1}}{I_{0,y-1}^{n,y-1}} = $$

$$ = \frac{I_{0,y-1}^{12,y-1}}{I_{0,y-1}^{n,y-1}} \cdot {}_{0,y}\Delta_{n,y} + {}_{n,y-1}\Delta_{12,y-1} $$

The decomposition (1) and definition (2) can now be applied to the expression of ${}_{n,y-1}\Delta_{n,y}$:

$$ {}_{n,y-1}\Delta_{n,y} = \frac{I_{0,y-1}^{12,y-1}}{I_{0,y-1}^{n,y-1}} \cdot \sum_{k=1}^{K} \pi^{y}(k) \cdot {}_{0,y}\Delta_{n,y}(k) + \sum_{k=1}^{K} \varepsilon^{n,y-1}(k) \cdot {}_{n,y-1}\Delta_{12,y-1}(k) \quad \Rightarrow $$

$$ {}_{n,y-1}\Delta_{n,y} = \sum_{k=1}^{K} \left(\frac{I_{0,y-1}^{12,y-1}}{I_{0,y-1}^{n,y-1}} \cdot {}_{0,y}C_{n,y}(k) + {}_{n,y-1}C_{12,y-1}(k) \right) \qquad (3) $$

Finally, the effect of the sub-index k on the yearly rate of change of the overall index is defined according to[6]:

$$ {}_{n,y-1}C_{n,y}(k) = \frac{I_{0,y-1}^{12,y-1}}{I_{0,y-1}^{n,y-1}} \cdot {}_{0,y}C_{n,y}(k) + {}_{n,y-1}C_{12,y-1}(k) \qquad (4) $$

4. The generalized decomposition method

In this section, we show how the envisaged decomposition formulas can be generalized to the case of the average rates of change of the overall index over different time intervals. To this end, it can be useful to look firstly at the decomposition of the annual average rate of change, and then consider the more general case.

The rate of change of the aggregate index between year *y-1* and year *y* is given by:

[6] Expression (4) corresponds to the "twelve-month effect" defined in Ribe (1999) (see pages 6-8).

$$ {}_{y-1}\Delta_y = \frac{\overline{I}_{BY}^{\,y}}{\overline{I}_{BY}^{\,y-1}} - 1 = \sum_{n=1}^{12}\left[\frac{I_{BY}^{n,y-1}}{\sum_{n=1}^{12} I_{BY}^{n,y-1}} \cdot \left(\frac{I_{BY}^{n,y} - I_{BY}^{n,y-1}}{I_{BY}^{n,y-1}}\right)\right] = $$

$$ = \sum_{n=1}^{12}\left[\frac{I_{0,y-1}^{n,y-1}}{\sum_{n=1}^{12} I_{0,y-1}^{n,y-1}} \cdot {}_{n,y-1}\Delta_{n,y}\right] \tag{5} $$

By substituting (3) and (4) in (5), we have:

$$ {}_{y-1}\Delta_y = \sum_{n=1}^{12}\left[\frac{I_{0,y-1}^{n,y-1}}{\sum_{n=1}^{12} I_{0,y-1}^{n,y-1}} \cdot \sum_{k=1}^{K} {}_{n,y-1}C_{n,y}(k)\right] = \sum_{k=1}^{K}\left\{\sum_{n=1}^{12}\left[\frac{I_{0,y-1}^{n,y-1}}{\sum_{n=1}^{12} I_{0,y-1}^{n,y-1}} \cdot {}_{n,y-1}C_{n,y}(k)\right]\right\} $$

The effect of component k on ${}_{y-1}\Delta_y$ can thus be defined in the following way:

$$ {}_{y-1}C_y(k) = \sum_{n=1}^{12}\left[\frac{I_{0,y-1}^{n,y-1}}{\sum_{n=1}^{12} I_{0,y-1}^{n,y-1}} \cdot {}_{n,y-1}C_{n,y}(k)\right] \tag{6} $$

According to (6), ${}_{y-1}C_y(k)$ corresponds to the weighted arithmetic mean of the contributions of the sub-index k to the yearly rates of change of the overall index, in the twelve months of the current year.

In the more general case, let ${}_{y-1}\Delta_y^{T,\tau}$ be the average rate of change of the overall index calculated in the time interval $(T;\, T+\tau)$, that is:

$$ {}_{y-1}\Delta_y^{T,\tau} = \frac{\sum_{n=T}^{T+\tau} I_{BY}^{n,y}}{\sum_{n=T}^{T+\tau} I_{BY}^{n,y-1}} - 1 $$

where $T = 1, 2, \ldots, 12$ and $0 \leq \tau \leq 12 - T$.

It is easy to show that the following formula holds:

$$_{y-1}\Delta_y^{T,\tau} = \sum_{k=1}^{K} {}_{y-1}C_y^{T,\tau}(k)$$

in which

$$_{y-1}C_y^{T,\tau}(k) = \sum_{n=T}^{T+\tau} \left[\frac{I_{0,y-1}^{n,y-1}}{\sum_{n=T}^{T+\tau} I_{0,y-1}^{n,y-1}} \cdot {}_{n,y-1}C_{n,y}(k) \right] \tag{7}$$

As a concluding remark, it is worth noting that expression (4) can be obtained, as a special case, by setting $\tau = 0$ in (7).

5. The effect of sub-indices as a measure of their contribute to overall inflation

The rest of the paper focuses on the formal properties of the decomposition formulas hitherto presented. The principal aim is to investigate the relationship between the rate of change of the index of a component *k* and the measure of its effect on the overall index. Specifically, it will be shown that the two magnitudes not necessarily have the same sign. More precisely, under well defined conditions, the rate of change of a sub-index and its contribution to the overall rate of change have opposite signs. However, this possibility never occurs when the effect on the monthly rate of change of the all-items index is considered. This case will be discussed at first.

5.1 The monthly rate of change of component *k* and its effect on the aggregate index

According to expression (2), the effect of the sub-index *k,* measured by ${}_{m,y}C_{n,y}(k)$, is a number representing how much of the monthly rate of change of the overall index can be imputed to the development of the price of the component *k,* in the same period.

As the definition of ${}_{m,y}C_{n,y}(k)$ makes clear, the size of the effect depends on three different factors:

- the rate of change of the sub-index *k* between month *m* and *n* of the current year;
- the relative weight of the sub-index *k*;
- the ratio between the level of the sub-index and the level of the overall index, in the lower bound of the time interval considered.

Regarding the first two factors, the direct relationship tying them to the effect on the overall index is intuitive enough: the higher is ${}_{m,y}\Delta_{n,y}(k)$ and the higher is $\pi^{y}(k)$, the higher is ${}_{m,y}C_{n,y}(k)$. As for the third element, it represents a "scale factor" in the computation of the contribution of the sub-index *k*. That is, assuming that two different sub-indices k_1 and k_2 have the same weight and exhibit the same rate of change in [(*m,y*);(*n,y*)], if $I_{0,y}^{m,y}(k_1)$ is twice as much as $I_{0,y}^{m,y}(k_2)$, their respective contributions to the rate of change of the overall index, in the considered period, will be in the same proportion.

Formally, with reference to expression (2), the following propositions can be immediately verified:

a.1 *${}_{m,y}C_{n,y}(k)$ is a linear (strictly) increasing function of the monthly rate of change of the index of component k in* [(*m,y*);(*n,y*)].

b.1 *The contribution ${}_{m,y}C_{n,y}(k)$ is null if, and only if, ${}_{m,y}\Delta_{n,y}(k)=0$.*

Notably, proposition ***a.1*** and ***b.1*** imply that ${}_{m,y}C_{n,y}(k)$ is a sign conservative function of ${}_{m,y}\Delta_{n,y}(k)$. Moreover, as a corollary of ***b.1***, the rate of change of the aggregate index, that would be measured under the hypothesis of constancy of the sub-index *k* in [(*m,y*);(*n,y*)], is given by the sum of the contributions of the other sub-indices, different from *k*.

Finally, let $K_H=\{k_1,\, k_2,\ldots,k_H\}$ be a subset of *H* components of the overall index and $I_{0,y}^{n,y}(K_H)$ the corresponding synthetic index. The following proposition can also be easily derived from expression (2)[7]:

c.1 *The contribution of K_H to the rate of change of the overall index in [(m,y);(n,y)] is given by the sum of the contributions of the k_H components. That is:* ${}_{m,y}C_{n,y}(K_H)=\sum_{k\in K_H}{}_{m,y}C_{n,y}(k)$.

In the next subsection, we will turn the attention to the decomposition of the rates of change of the overall index calculated on two adjacent years. We will argue that, in such a case, the failure to satisfy the additive property by the chained price index has relevant implications on the formal properties of the contribution function.

[7] In practical terms, the proposition ***c.1*** states that the effect of an index corresponding, for example, to a COICOP division on the monthly inflation rate is equal to the sum of the effects of the elementary indices belonging to the same COICOP division.

5.2 The "year on year" rate of change of component *k* and its effect on the aggregate index

Similarly to the previous case, the decomposition (3) allows a measure of the effect of the single sub-index *k* on the yearly rate of change of the all-items index. In other words, it provides the basis for the measurement of the contribution of *k* to the dynamics of the overall index in a given year. However, as shown by expression (4), the effect of the sub-index *k* on the overall index in $[(n, y-1);(n, y)]$ is not generally reducible to the sum of the contributions of *k* on the two subintervals of the year, $\theta_1 = [(n, y-1);(12, y-1)]$ and $\theta_2 = [(0, y);(n, y)]$. Moreover, the effect of the sub-index *k* on the yearly overall rate of change is not independent from the dynamics of the other components in θ_1.

More in detail, taking into account the expressions of ${}_{0,y}C_{n,y}(k)$ and ${}_{n,y-1}C_{12,y-1}(k)$, the contribution function (4) can be written as:

$${}_{n,y-1}C_{n,y}(k) = \frac{I_{0,y-1}^{12,y-1}}{I_{0,y-1}^{n,y-1}} \cdot \pi^{y}(k) \cdot {}_{0,y}\Delta_{n,y}(k) + \pi^{y-1}(k) \cdot \frac{I_{0,y-1}^{n,y-1}(k)}{I_{0,y-1}^{n,y-1}} \cdot {}_{n,y-1}\Delta_{12,y-1}(k) \quad (8)$$

The following propositions can then be stated:

a.2 *${}_{n,y-1}C_{n,y}(k)$ is an increasing function of the rate of change of the index of the component k in both θ_1 and θ_2;*

b.2 *${}_{n,y-1}C_{n,y}(k) = 0$ if, and only if:*

$$I_{0,y-1}^{12,y-1} \cdot {}_{0,y}C_{n,y}(k) = -\, I_{0,y-1}^{n,y-1} \cdot {}_{n,y-1}C_{12,y-1}(k) \quad (9)$$

or, in terms of (8):

$${}_{0,y}\Delta_{n,y}(k) = -\frac{\pi^{y-1}(k)}{\pi^{y}(k)} \cdot \frac{I_{0,y-1}^{n,y-1}(k)}{I_{0,y-1}^{12,y-1}} \cdot {}_{n,y-1}\Delta_{12,y-1}(k) \quad (10)$$

The interpretation of (9) is straightforward: the expression on the left hand side of the equality, $I_{0,y-1}^{12,y-1} \cdot {}_{0,y}C_{n,y}(k)$, is the amount of the change of the level of the overall index, measured in θ_2 (i.e. $I_{0,y-1}^{n,y} - I_{0,y-1}^{12,y-1}$), that can be imputed to the sub-index *k*. Similarly, the expression at the right hand side, $I_{0,y-1}^{n,y-1} \cdot {}_{n,y-1}C_{12,y-1}(k)$ represents how much of the difference $I_{0,y-1}^{12,y-1} - I_{0,y-1}^{n,y-1}$ can be attributed to the same component. Accordingly, ${}_{n,y-1}C_{n,y}(k)=0$ if the first magnitude exactly offsets the second one.

It should be noted - see (10) - that a trivial case, in which the contribution ${}_{n,y-1}C_{n,y}(k)$ vanishes, occurs when ${}_{0,y}\Delta_{n,y}(k)$ and ${}_{n,y-1}\Delta_{12,y-1}(k)$ are both equal to zero. In this case, the yearly rate of change of the index of the sub-index *k*, ${}_{n,y-1}\Delta_{n,y}(k)$, is zero as well. However, it is important to stress that, generally, the condition ${}_{n,y-1}\Delta_{n,y}(k)=0$ is neither necessary nor sufficient for ${}_{n,y-1}C_{n,y}(k)=0$.

It order to investigate this issue more in depth, it is useful to write the condition (10) as follows:

$$ {}_{0,y}\Delta_{n,y}(k) = -\frac{\pi^{y-1}(k)}{\pi^{y}(k)} \cdot \frac{I_{0,y-1}^{12,y-1}(k)}{I_{0,y-1}^{12,y-1}} \cdot \frac{{}_{n,y-1}\Delta_{12,y-1}(k)}{1+{}_{n,y-1}\Delta_{12,y-1}(k)} \tag{11} $$

Let ${}_{0,y}\widetilde{\Delta}_{n,y}(k)$ indicate the rate of change of the sub-index *k,* in the subinterval θ_2 that satisfies (11).

Now, the annual rate of change of the sub-index *k,* ${}_{n,y-1}\Delta_{n,y}(k)$ is null if, and only if :

$$ {}_{0,y}\Delta_{n,y}(k) = -\frac{{}_{n,y-1}\Delta_{12,y-1}(k)}{1+{}_{n,y-1}\Delta_{12,y-1}(k)} \equiv -\,{}_{12,y-1}\Delta_{n,y-1}(k) \tag{12} $$

In a similar way, let ${}_{0,y}\hat{\Delta}_{n,y}(k)$ be the rate of change of the sub-index *k,* in θ_2 that satisfies (12).

Finally, the necessary and sufficient condition for ${}_{n,y-1}C_{n,y}(k)=0$ can be written as:

$$ {}_{0,y}\Delta_{n,y}(k) = {}_{0,y}\widetilde{\Delta}_{n,y}(k) = \frac{\pi^{y-1}(k)}{\pi^{y}(k)} \cdot \frac{I_{0,y-1}^{12,y-1}(k)}{I_{0,y-1}^{12,y-1}} \cdot {}_{0,y}\hat{\Delta}_{n,y}(k) \tag{13} $$

Equation (13) states that the difference between ${}_{0,y}\widetilde{\Delta}_{n,y}(k)$ and ${}_{0,y}\hat{\Delta}_{n,y}(k)$ depends on the change in the weight of component *k* between the two years and on the *links ratio*, that is on the size of the link of the sub-index *k* with respect to the (weighted arithmetic) mean of the links of all the components. Formally:

$$\left|{}_{0,y}\tilde{\Delta}_{n,y}(k)\right| \begin{matrix} > \\ = \\ < \end{matrix} \left|{}_{0,y}\hat{\Delta}_{n,y}(k)\right| \quad \Leftrightarrow \quad \frac{\pi^{y-1}(k)}{\pi^{y}(k)} \cdot \frac{I_{0,y-1}^{12,y-1}(k)}{I_{0,y-1}^{12,y-1}} \begin{matrix} > \\ = \\ < \end{matrix} 1 \qquad (14)$$

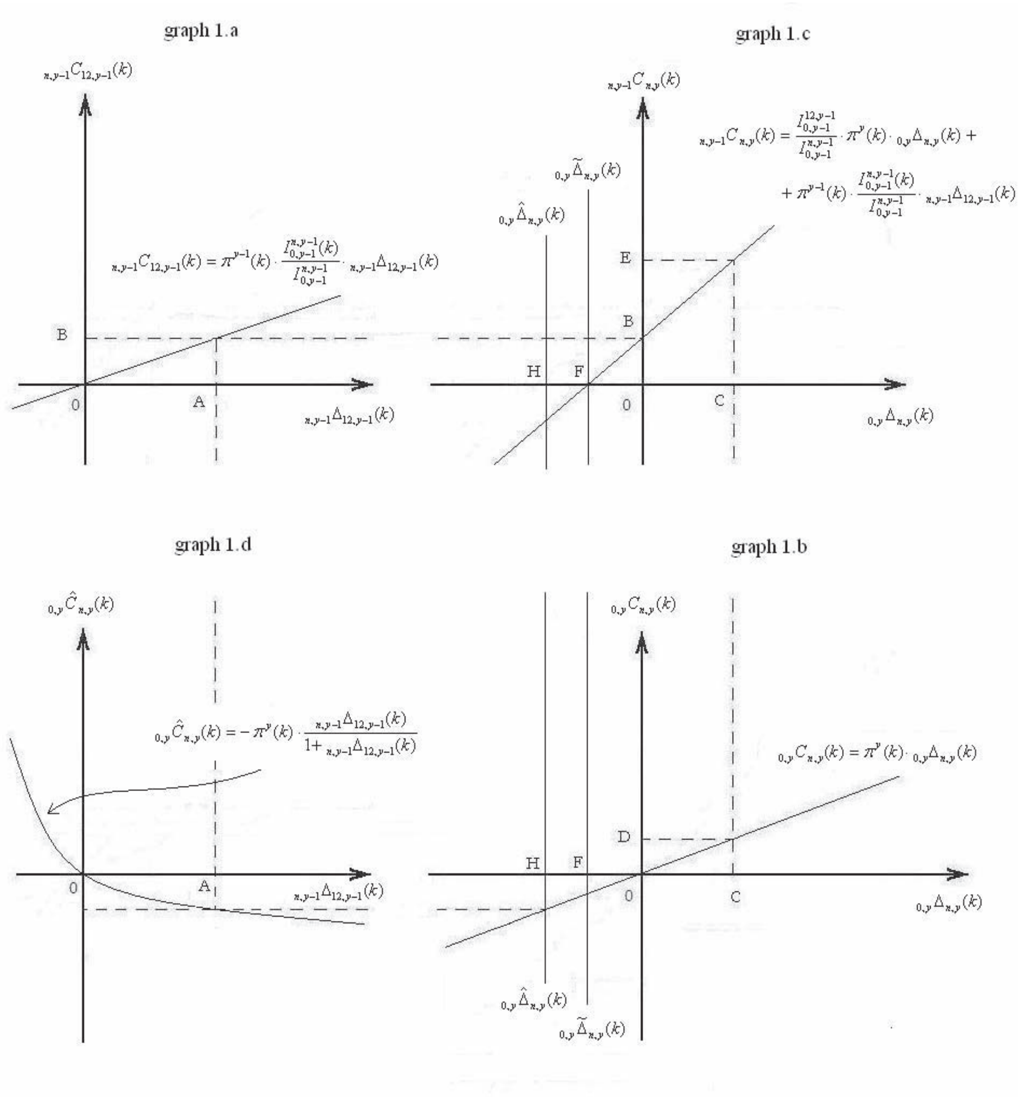
As a crucial implication of the condition (13), the sign of the yearly rate of change of the sub-index *k* and the sign of its contribution to the overall yearly rate of change can differ. In order to clarify this point, it could be useful to show graphically the relation between ${}_{0,y}\tilde{\Delta}_{n,y}(k)$ and ${}_{0,y}\hat{\Delta}_{n,y}(k)$. To this aim, in Figure 1, the contribution function ${}_{n,y-1}C_{12,y-1}(k)$ has been plotted (see graph 1.a). If we assume, for example, that the rate of growth of the sub-index *k*, between month *n* of year *y*-1 and December of the same year, is given by the segment OA, the corresponding effect on the aggregate index, in θ_1, would be equal to OB. Similarly, in graph 1.b, the contribution function ${}_{0,y}C_{n,y}(k)$ is drawn: should the rate of change of the sub-index, in the first *n* months of the following year, be OC, the contribution of *k* to the rate of change of the overall index in θ_2 would be OD, while the effect on the inflation rate would be $OE > OB + OD$ (see graph 1.c).

The point of intersection F between the function ${}_{n,y-1}C_{n,y}(k)$ and the horizontal axis in graph 1.c measures the rate of change of the sub-index *k* in θ_2 that satisfies (10). Moreover, having assumed ${}_{n,y-1}\Delta_{12,y-1}(k) = OA > 0$, the yearly rate of change of the sub-index *k* is zero if (and only if) its rate of decline, in θ_2, is equal to OH (graph 1.b). The position of the point *H* in graph 1.b can be easily set through the curve drawn in graph 1.d. For any given ${}_{n,y-1}\Delta_{12,y-1}(k)$, this curve defines the contribution ${}_{0,y}\hat{C}_{n,y}(k)$ which is associated with the rate of change of *k*, in θ_2, that satisfies condition (12).

In the example shown in the picture, the point H lies at the left of F: in this case, there is a set of different values of ${}_{0,y}\Delta_{n,y}(k)$, in the interval $(H;F)$, for which the yearly rate of change of the sub-index *k* is positive, even though its contribution to the all-items index is negative. However, it is also possible to consider the opposite case, in which the point H is at the right of F. As it has already been noticed, the occurrence of the first or of the second eventuality strictly depends on the condition (14).

It is also important to stress that, as an implication of proposition ***b.2***, the sum of the contributions of the other sub-indices different from *k* cannot be considered as the measure of the yearly rate of change that the overall index would have if it is assumed that ${}_{n,y-1}\Delta_{n,y}(k) = 0$.

Figure 1 - The change of the index of the component k and its effect on the aggregate index.



Lastly, with reference to the subset of *H* components $K_H = \{k_1, k_2, \ldots, k_H\}$ of the overall index, the following proposition can be proved[8]:

[8] The proof can be obtained by decomposition (3), taking into account that the proposition ***c.1*** applies to $_{0,y}C_{n,y}(k)$ and to $_{n,y-1}C_{12,y-1}(k)$.

c.2 *The contribution of* K_H *to the rate of change of the overall index in [(n,y-1);(n,y)] is given by the sum of the contributions of the* k_H *components. That is:* ${}_{n,y-1}C_{n,y}(K_H) = \sum_{k \in K_H} {}_{n,y-1}C_{n,y}(k)$.

5.3 The annual average rate of change of *k* and its effect on the overall index

In this subsection we will limit the analysis to the decomposition of the overall yearly average rate of change, given by (6)[9]. Putting expression (4) in (6) and rearranging, we have:

$$ {}_{y-1}C_y(k) = \frac{I_{0,y-1}^{12,y-1}}{\sum_{n=1}^{12} I_{0,y-1}^{n,y-1}} \cdot \sum_{n=1}^{12} {}_{0,y}C_{n,y}(k) + \sum_{n=1}^{12} \frac{I_{0,y-1}^{n,y-1}}{\sum_{n=1}^{12} I_{0,y-1}^{n,y-1}} \cdot {}_{n,y-1}C_{12,y-1}(k) $$

If we take into account the expressions of ${}_{0,y}C_{n,y}(k)$ and ${}_{n,y-1}C_{12,y-1}(k)$, with some algebra, it is possible to write the contribution function ${}_{y-1}C_y(k)$ as follows:

$$ {}_{y-1}C_y(k) = \frac{I_{0,y-1}^{12,y-1}}{\bar{I}_{0,y-1}^{y-1}} \cdot \pi^y(k) \cdot {}_{0,y}\Delta_y(k) + \pi^{y-1}(k) \cdot \frac{\bar{I}_{0,y-1}^{y-1}(k)}{\bar{I}_{0,y-1}^{y-1}} \cdot {}_{y-1}\Delta_{12,y-1}(k) \quad (15) $$

where :

- $\bar{I}_{0,y-1}^{y-1} = \dfrac{\sum_{n=1}^{12} I_{0,y-1}^{n,y-1}}{12}$ and $\bar{I}_{0,y-1}^{y-1}(k) = \dfrac{\sum_{n=1}^{12} I_{0,y-1}^{n,y-1}(k)}{12}$;
- ${}_{0,y}\Delta_y(k) \equiv \{\bar{I}_{0,y}^{y}(k) - 1\}$ is the measure of the rate of change of the yearly average sub-index *k* with respect to its base, $I_{0,y}^{0,y}(k)$, which is by definition equal to one (conventionally, ${}_{0,y}\Delta_y(k)$ will be referred to as the rate of change on the period $[(0,y);y]$);

[9] However, the results can be easily extended to the more general case.

- ${}_{y-1}\Delta_{12,y-1}(k) \equiv \left\{ \dfrac{I_{0,y-1}^{12,y-1}(k) - \bar{I}_{0,y-1}^{y-1}(k)}{\bar{I}_{0,y-1}^{y-1}(k)} \right\}$ is the measure of the relative change of the December (*y-1*) index of the sub-index *k* with respect to the yearly average index of the same component (similarly, ${}_{y-1}\Delta_{12,y-1}(k)$ will be referred to as the rate of change on period $\left[y-1;(12,y-1)\right]$)[10].

Regarding (15), the following propositions hold:

a.3 *${}_{y-1}C_y(k)$ is an increasing function of ${}_{0,y}\Delta_y(k)$ and ${}_{y-1}\Delta_{12,y-1}(k)$;*

b.3 *${}_{y-1}C_y(k)=0$ if, and only if:*

$$ {}_{0,y}\Delta_y(k) = -\frac{\pi^{y-1}(k)}{\pi^y(k)} \cdot \frac{\bar{I}_{0,y-1}^{y-1}(k)}{I_{0,y-1}^{12,y-1}} \cdot {}_{y-1}\Delta_{12,y-1}(k) $$

or alternatively:

$$ {}_{0,y}\Delta_y(k) = -\frac{\pi^{y-1}(k)}{\pi^y(k)} \cdot \frac{I_{0,y-1}^{12,y-1}(k)}{I_{0,y-1}^{12,y-1}} \cdot \frac{{}_{y-1}\Delta_{12,y-1}(k)}{1+{}_{y-1}\Delta_{12,y-1}(k)} \tag{16} $$

Let ${}_{0,y}\tilde{\Delta}_y(k)$ be the rate of change of the sub-index *k,* in $\left[(0,y);y\right]$, satisfying (16). Moreover, let ${}_{0,y}\hat{\Delta}_y(k)$ be the rate of change of the sub-index *k,* in $\left[(0,y);y\right]$, such that ${}_{y-1}\Delta_y(k)=0$. It is easy to show that:

$$ {}_{0,y}\hat{\Delta}_y(k) = -\frac{{}_{y-1}\Delta_{12,y-1}(k)}{1+{}_{y-1}\Delta_{12,y-1}(k)} \tag{17} $$

[10] It is worth noting that, generally speaking, ${}_{y-1}\Delta_{12,y-1}$ and ${}_{0,y}\Delta_y$ correspond (approximately) to a decomposition of the annual average rate of change ${}_{y-1}\Delta_y$ in two parts: the first one measures the amount of ${}_{y-1}\Delta_y$ which is due to the development of prices in the final months of the year *y-1*, while the second one corresponds to that part of ${}_{y-1}\Delta_y$ which depends on the price changes in the current year. For this reason, sometimes they have been respectively considered the "*inflation rate inherited*" by the year *y* from the year *y-1* and the "*proper inflation rate*" of the year *y*. See Predetti A. (1994), page. 111 and the followings, for more details.

Since generally ${}_{0,y}\widetilde{\Delta}_y(k) \neq {}_{0,y}\hat{\Delta}_y(k)$, concerning the decomposition of the overall yearly average rate of change, the condition ${}_{y-1}\Delta_y(k) = 0$ is neither necessary nor sufficient in order to have ${}_{y-1}C_y(k) = 0$.

As in the previous case, proposition ***b.3*** implies that the sum of the contributions of the other components different by *k* cannot be considered as the measure of the annual average rate of change that the overall index would exhibit under the hypothesis of ${}_{y-1}\Delta_y(k) = 0$.

Considering the subset of *H* components $K_H = \{k_1, k_2, \ldots, k_H\}$ of the overall index, it is possible to prove that[11]:

c.3 *The contribution of* K_H *to the rate of change of the overall index in [y-1;y] is given by the sum of the contributions of the* k_H *components. That is:* ${}_{y-1}C_y(K_H) = \sum_{k \in K_H} {}_{y-1}C_y(k)$.

6. Conclusion

In the present paper the method of decomposition of the monthly and yearly rates of change of a chained price index into the sum of the effects deriving from price changes of its sub-indices has been generalized to the case of the average rates of change over different time intervals (such as yearly or quarterly average rates of change). Moreover, the formal properties of the decomposition have been discussed in order to allow an evaluation of its interpretative effectiveness as a measure of the impact of the evolution of the prices of different components on overall inflation. We showed that, in the more general case, the rate of change of a sub-index and its effect on the overall rate of change do not necessarily have the same algebraic sign, that is: the contribution to inflation of the component, in a given time interval, may be positive even though the rate of change of the price index of the same component is negative (and vice versa). As a complement to the previous analysis, it might be of some utility to provide some evidence derived from the Italian Harmonized Index of Consumer Prices (HICP). Specifically, we consider the yearly average rates of change of the HICP sub-indices, corresponding to the three digits of the COICOP-HICP classification and their contribution to the rate of change of the all-items index for year 2009.

[11] Proposition ***c.3*** can be proved from definition (6) and proposition ***c.2***.

Table. 1 - Distribution of the HICP three-digit COICP sub-indices according to the sign of their annual average rate of change and to the sign of their contribution to the yearly average rate of change of the all-items index - Year 2009

		sign of the contribution to the annual average rate of change of the all-items index		total
		+	-	
sign of the annual average rates of change of the HICP components	+	31	1	32
	-	0	7	7
total		31	8	39

Table 1 represents the distribution of the sub-indices according to the sign of their rate of change and to the sign of contribution to rate of change of the overall index (see Table 2 in the Appendix for more details).

One sub-index (C.03.1 – "*clothing*"), out of 39, proves to have a (slight) deflationary effect in the concerned year, even though the corresponding price sub-index exhibits a positive rate of growth.

This result can be easily explained by verifying that: ${}_{0,2009}\hat{\Delta}_{2009}(C.03.1) < {}_{0,2009}\Delta_{2009}(C.03.1) < {}_{0,2009}\tilde{\Delta}_{2009}(C.03.1)$. It is in fact possible to show that, in the case at hand:

$$\begin{cases} {}_{0,2009}\hat{\Delta}_{2009}(C.03.1) = -\,0.0528; \\ {}_{0,2009}\Delta_{2009}(C.03.1) = -\,0.0509; \\ {}_{0,2009}\tilde{\Delta}_{2009}(C.03.1) = -\,0.0505 \end{cases}$$

Counterintuitive as it may appear, this event depends on specific conditions that have been formally investigated in the previous sections of the paper.

Appendix

Table. 2 - Yearly average rates of change and contribution to the yearly inflation of the sub-indices of the Italian HICP - Year 2009

Coicop-hicp	annual average rates of change	contribution to annual inflation
C.00 - All-items HICP	**0.8**	**-**
C.01.1 - Food	1.6	0.271
C.01.2 - Non-alcoholic beverages	1.4	0.018
C.02.1 - Alcoholic beverages	2.8	0.023
C.02.2 - Tobacco	4.1	0.092
C.03.1 - Clothing	0.2	-0.003
C.03.2 - Footwear including repair	1.5	0.032
C.04.1 - Actual rentals for housing	3.3	0.077
C.04.3 - Maintenance and repair of the dwelling	2.9	0.038
C.04.4 - Water supply and miscellaneous services relating to the dwelling	4.7	0.107
C.04.5 - Electricity, gas and other fuels	-5.0	-0.221
C.05.1 - Furniture and furnishings, carpets and other floor coverings	1.6	0.056
C.05.2 - Household textiles	1.3	0.006
C.05.3 - Household appliances	0.4	0.005
C.05.4 - Glassware, tableware and household utensils	2.8	0.023
C.05.5 - Tools and equipment for house and garden	1.6	0.005
C.05.6 - Goods and services for routine household maintenance	1.9	0.056
C.06.1 - Medical products, appliances and equipment	4.5	0.077
C.06.2 - Out-patient services	2.2	0.029
C.06.3 - Hospital services	1.7	0.010
C.07.1 - Purchase of vehicles	1.2	0.057
C.07.2 - Operation of personal transport equipment	-3.6	-0.346
C.07.3 - Transport services	-3.0	-0.069
C.08.1 - Postal services	5.6	0.009
C.08.2 - Telephone and telefax equipment and services	-0.6	-0.015
C.09.1 - Audio-visual, photographic and information processing equipment	-5.9	-0.055
C.09.2 - Other major durables for recreation and culture	1.2	0.004
C.09.3 - Other recreational items and equipment, gardens and pets	0.9	0.012
C.09.4 - Recreational and cultural services	2.6	0.047
C.09.5 - Newspapers, books and stationery	2.4	0.043
C.09.6 - Package holidays	-0.5	-0.002
C.10 - Education	2.7	0.029
C.11.01 - Catering services	2.3	0.205
C.11.02 - Accommodation services	-2.4	-0.072
C.12.01 - Personal care	1.7	0.049
C.12.03 - Personal effects n.e.c.	4.4	0.053
C.12.04 - Social protection	2.4	0.020
C.12.05 - Insurance	2.7	0.038
C.12.06 - Financial services n.e.c.	3.1	0.028
C.12.07 - Other services n.e.c.	2.0	0.025

References

Allen R. D.G. (1975), *Index number in Theory and Practice*, Mac Millan Press, London.

Balk B.M. (2008), *Price and Quantity Index Numbers. Models for Measuring Aggregate Change and Difference*, Cambridge University Press, New York.

Forsyth F.G., Fowler R.F. (1981), "The Theory and Practise of Chain Price Index Numbers", *Journal of the Royal Statistical Society*. Series A (General), Vol. 144, N° 2, pp. 224 – 246.

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

Predetti A. (1994), *I Numeri Indici. Teoria e Pratica*, Giuffrè Editore, Milano.

Quaranta V., Di Iorio F. (1997), "Frequenza dell'aggiornamento dei pesi degli indici dei prezzi al consumo ed indici a catena", *Contributi Istat*, n. 14.

Ribe M. (1999), "Effects of subcomponents on a chained price indices like the HICPs and the MUICP", *mimeo*, Statistics Sweden.

List prices vs. bargain prices: which solution to estimate consumer price indices?[1]

Carlo De Gregorio[2]

Abstract

Alternative approaches to CPI surveys are here evaluated, in markets where final prices are based on some sort of price listing. Three types of surveys are compared: local surveys (LOC), with small samples and a local price collection; list price surveys (LIS), with huge samples and centralised collection; mixed surveys (MXD), in which LOC and LIS are jointly used. Based on a multiplicative pricing model, some conditions are derived to establish the relative efficiency of these approaches. The alternatives have also been tested on five different random populations. LIS surveys appear very efficient under very restrictive hypotheses on the regularity of discount policies. LOC surveys may be efficient only if the variability of list prices is reasonably low. MXD surveys appear the most promising solution if a correction parameter is introduced to account for the covariance between list prices and discount policies. MXD surveys appear better positioned for monitoring consumer market and the range of products available to the consumer.

Keywords: Consumer price index, Survey design, Sampling

Introduction

In some consumer markets the prices actually paid by customers are decided by retailers on the basis of some kind of official listing of prices[3]. This happens for instance in the case of new vehicles, package holidays, tourist services, pharmaceutical drugs, tobacco, ICT, housing: list prices affect with varying degrees a wide set consumer markets. Generally list prices may be regarded as a kind of rule for pricing policy, by means of which producers, institutions or retailers themselves influence the dynamics of actual bargain prices[4]. This conditioning usually does not imply a straight control: list prices may in fact generally diverge from transaction prices, especially if retailers have some margins to apply their own

---

[1] This paper has been prepared during a period of secondment at Eurostat, during which the author collaborated with the unit G6, in charge of the production of the HICP, and coordinated the works of a HICP Task Force on Sampling which stimulated the work on this subject. The author is obviously the only responsible for the views expressed in this paper, which do not necessarily reflect the views of Istat nor those of Eurostat.

[2] Senior researcher Istat, email: cadegreg@istat.it.

[3] We will use the term "list prices" to refer to what are also usually called "*the (manufacturer's) suggested retail price (MSRP or SRP), list price or recommended retail price (RRP)*". See Wikipedia, *Suggested retail price*. This source also specifies that a SRP is "*the price the manufacturer recommends that the retailer sell it for. The intention was to help to standardize prices among locations. While some stores always sell at, or below, the suggested retail price, others do so only when items are on sale or closeout*". We will use indifferently the terms "transaction price" or "bargain price" to mean the price effectively paid by the consumer.

[4] The connection between oligopoly or imperfect competition and the use of list prices is, for instance, mentioned by Sweezy (1939) in his milestone explanation of the concept of the "kinked demand curve". Sweezy notes moreover that " *(…) list prices become less trustworthy guides to real prices the longer bad times last*" (pag. 572). In completely different perspectives and contexts, but on the same subject, see, for example, Marn et al. (2003) and Thanassoulis (2005).

pricing and sales policy. Generally these margins imply the possibility to apply discounts to list prices: the latter can be thus interpreted as a sort of superior limit.

There are several market-specific ways to intend list prices, and it is not our intention to classify them here. In general, the markets where list prices are applied are usually characterised by some degree of oligopoly, at least at producer level, or by a regulatory framework: anyway, the intrinsic nature of pricing policies may differ to some extent from market to market. In the case of new cars, for instance, producers' list prices represent a medium term pricing policy tool. Effective prices diverge from list prices because of at least two distinct factors: the last minute discount policies decided by the producers are used to bring short term corrections to list prices; at the end of the chain, dealers apply their point-of-sale policies[5]. Some kinds of pharmaceutical drugs are sold at a retail price that discounts an official price[6]. In some countries, even tobacco can be priced by retailers on the basis of list prices, although more often this market is strictly constrained to official price lists. The case of package holidays looks very similar to the one of new cars: official list prices, published in the catalogues, are usually discounted for advance or last minute booking, with a pricing policy that reflects short term market related concerns by tour operators and travel agents. Other markets, where there is no distinction between producers and retailers, are also regulated by suggested list prices: this is the case, for example, of accommodation services and of many tourist services. Something different happens in the housing market, where the prices publicly asked for by the sellers play the same role as list prices[7].

In general, the use of list prices for CPI estimates is discouraged by standard practices, although it is tolerated in those cases where the collection of bargain prices is deemed too heavy. In general, when list and bargain prices coexist, in order to produce accurate estimates of price dynamics we have to face a trade off between alternative data collection techniques. List prices are in fact much easier to collect[8], they can be managed centrally and with more sophisticated sampling designs: but they may be not representative of transaction prices. The collection of bargain prices, on the other hand, implies a much higher burden, with the assistance of a local price collection network and a smaller sample size. This burden may happen to be very high: in the case of cars, for example, highly skilled price collectors are strongly needed (Eurostat, 2005).

These two approaches are generally conceived as mutually exclusive. In the case of new cars, for example, existing approaches can be divided in two broad classes: some use list prices and others collect actual bargain prices by means of local surveys. In both cases, large drawbacks can be identified. Even a good or almost-perfect measurement of the dynamics of list prices can be misleading, more probably in the short term. On the other side, the smaller samples used for the collection of bargain prices affects heavily the coverage of market segments and the management of the turnover in the product range.

This last point is having a growing importance. The regulatory framework of the HICP has explicitly introduced the concept of consumption segment, and has identified it as the

[5] There is an appreciable economic literature (mainly in the US) on the divergence between list and bargain prices in the market for new cars (Zettelmeyer et al., 2003;Goldberg, 1996; Jung, 1959). This literature mainly focuses on price discrimination. Retail policies are tailored on the various types of customers (e.g.: age, social conditions, sex) (Scott-Morton et al., 2003). They are also based on how much information the customer shows to have upon the market. For a similar approach, concerning the European car market, see Lutz (2004).

[6] This is for example the case of non prescription drugs in Italy.

[7] There is some literature on this subject: see for example Genesove at al. (2001), Knight et al. (1998), Horowitz (1992).

[8] They are very often available on the internet.

cornerstone in the definition of the population parameter in the estimation of the CPI (European Commission, 2007). This mere point has given a major role to the analysis of consumer markets in the estimates of a CPI. In this context the mixed approaches, which try to combine the price collection techniques, might be encouraged: nevertheless, at the moment they are not explicitly applied nor it seems that they have been systematically considered. Given this context, we address in particular the following questions: Can it be useful to split the sampling issue in two components, one aimed at estimating list prices dynamics and the other dedicated to the estimate of the dynamics of retailers' policies? Is it efficient to estimate a list price index by means of a centralised and almost exhaustive survey, and to estimate separately an average retailers' discount index by means of a much smaller local survey? Is it possible to obtain an unbiased estimate and a good gain in precision? At which conditions? And what about sample size? In other words, is it possible to exploit the large centralised samples – and the high quality of the available information on the behaviour of list prices - to support the information collected by means of smaller local surveys?

Quite obviously, the answer to these questions depends crucially on the behaviour of each component of the variability of final prices. In section 1 it is proposed a model to analyse this variability. In section 2 some conditions are derived in order to evaluate the relative efficiency of different approaches. In section 3 some sampling simulations are made on randomly generated populations. In section 4 some generalisations are examined. Finally, some conclusions are extracted from the results of this analysis.

1. Two stage pricing and variance decomposition

Let's assume that we are dealing with a market where transaction prices are based on suggested list prices. In particular, if we use the subscript *i* to identify each "model"[9] and the suffix *ij* to identify the single transaction involving model *i*, we could express the effective price p_{ij} of product *i* in transaction *ij* as:

$$p_{ij} = l_i r_{ij} \qquad (1.1)$$

where *l* identifies the list price and *r* a multiplicative discount factor[10]. Each transaction price is thus viewed as a list price multiplied by a correction factor reflecting the pricing policy of the retailer. Suppose also that it is possible to have a one-to-one correspondence between each transactions in the base and its replication in the current period[11]. Using

[9] For explanatory purposes, the terms we use are implicitly referred to the car market, but they should be regarded as having more general validity. It is also implicitly assumed, in order not to complicate too much the formulas, that the models present in this market have all similar market shares (in value terms).

[10] We can easily assume that $r_{ij} \leq 1$, so that $p_{ij} \leq l_{ij}$, that is retail prices are discounted list prices. In other words, for each model *i* there exists one only list price in each moment and a plurality of product offers *ij*, each one characterised by an effective price. Alternative descriptions might be used to express the same concept, for example using an additive discount component: moreover, multiplicative factors can be readily translated in additive ones by means of logarithmic transformations.

[11] This is a major hypothesis, since it gives the possibility to simplify very much the computation of the index and the sample design. See Ribe (2000). It is equivalent to assume that no major change in the range of products available for consumption has taken place between the base and the current month: in other terms, there is no need to take account of replacements. See in section 4 some thoughts on the consequences connected to the relaxation of this hypothesis.

capital letters to express the corresponding indices we have[12]:

$$Y_{ij} = X_i D_{ij} \qquad (1.2)$$

where the suffix for the current month *m* and the reference month *0* are omitted[13], while *i=1, ..., Q* and *j=1, ..., n*$_i$. Let $N = \sum_i n_i$ be the total number of transactions. Assume for simplicity that the true population average of the index *Y* is:

$$\overline{Y} = \frac{\sum Y_{ij}}{N} = E[Y] = E[X]E[D] + Cov(X,D) = \overline{X} * \overline{D} + \sigma_{XD} \qquad (1.3)^{14}.$$

With some algebra, and assuming that X and D are normally distributed, the variance of *Y* can be decomposed as follows:

$$\sigma_Y^2 = E\left[\left(Y - \overline{Y}\right)^2\right] = \overline{X}^2\sigma_D^2 + \overline{D}^2\sigma_X^2 + \sigma_X^2\sigma_D^2 + 2\overline{XD}\sigma_{XD} + \sigma_{XD}^2 =$$
$$\overline{X}^2\sigma_D^2 + \overline{D}^2\sigma_X^2 + \sigma_X^2\sigma_D^2 + 2\overline{XD}\sigma_D\sigma_X\rho_{XD} + \left[\sigma_D\sigma_X\rho_{XD}\right]^2 \qquad (1.4)^{15},$$

where ρ_{XD} is the correlation coefficient between the indices of list prices and of the discount policies. The last two terms of expression (1.4) depend on the linear relationship between list prices and discount policies.

If the index D_{ij} is constant for every couple *ij* then:

$$\sigma_Y^2 = \overline{D}^2\sigma_X^2 \qquad (1.5),$$

that is the variance of transaction price indices (*Y*) depends only on the variance of list price indices (*X*) while the index of the discount policy represents merely a scale factor. They coincide if no retail policy is applied (D_{ij}=1).

If *X* and *D* are only linearly independent, their covariance and correlation coefficient are both null and expression (1.4) can be simplified as follows:

$$\sigma_Y^2 = \overline{X}^2\sigma_D^2 + \overline{D}^2\sigma_X^2 + \sigma_X^2\sigma_D^2 \qquad (1.6).$$

---

[12] That is: $Y^m = \frac{p^m}{p^0}$, $X^m = \frac{l^m}{l^0}$, and $D^m = \frac{r^m}{r^0}$.

[13] According to the current annual chaining procedures, which nowadays characterises the HICP project, month 0 can be identified with the month of December of the preceding year. Due to annual chaining, and following Ribe (2000), the objects of CPI estimates (i.e.: the population parameters to be estimated) is the set of 12 monthly indices of the current year based December of the preceding year. This same approach has been followed in Istat centralised CPI surveys (Istat, 2007; De Gregorio, 2006; De Gregorio, Fatello et al., 2008; De Gregorio, Munzi et al., 2008).

[14] In order to have a value weighted Laspeyres index we should have included weights. We omit them in order not to complicate the notation: these results would not have change substantially if we had used weighted means of the indices or any other elementary aggregation technique (e.g.: Jevons, Dutot, etc.). The simplification that is adopted here is useful for a straightforward application of simple random sampling results.

[15] The general formula, with no restrictions on the distributions of X and D, is:

$$\sigma_Y^2 = \overline{X}^2\sigma_D^2 + \overline{D}^2\sigma_X^2 + \sigma_X^2\sigma_D^2 - 2\overline{XD}\sigma_D\sigma_X\rho_{XD} - \left[\sigma_D\sigma_X\rho_{XD}\right]^2 + \sigma_{X^2Y^2} \qquad (1.4.n).$$

2. Estimators and biases

A comparison can be made by observing the characteristics of the estimates obtained with three different types of surveys: local surveys (*LOC*), in which a sample of n_{loc} bargain prices is observed by means of a network of price collectors visiting a sample of outlets; centralised surveys of list prices (*LIS*), by means of which only list prices are collected, presumably centrally, with a sample size n_{lis} which it is reasonable to expect much larger than n_{loc} and eventually very near to be exhaustive[16]; mixed surveys (*MXD*), in which list and transaction prices are observed with independent *LIS* and *LOC* surveys and distinct sample sizes[17].

2.1 Local survey of bargain prices (LOC)

Examining the first case, let's assume that a simple random sample of transactions is drawn from the population Y_{ij}. Given formula (1.3), an estimate of the price index is

$$\bar{y}_{loc} = \frac{\sum y_{ij}}{n_{loc}} \qquad (2.1.1)$$

where n_{loc} is the sample size, corresponding to a small sampling fraction. Obviously, this sample mean is an unbiased estimate of the population mean:

$$B_{loc} = E\left[\bar{y}_{loc} - \bar{Y}\right] = 0 \qquad (2.1.2)$$ [18]

The average square deviation (*S*) between the sample mean and the population mean coincides with the variance of the sample mean $V(\bar{y})$[19]:

$$S(\bar{y}_{loc}) = \sigma^2_{\bar{Y}} = \frac{\sigma^2_Y}{n_{loc}}(1 - f_{loc}) \qquad (2.1.3)$$

where f is the sampling fraction ($f_{loc} = \frac{n_{loc}}{N}$) and σ^2_Y is defined in (1.4). The sampling fraction will be probably very close to zero, so that the preceding expression can be modified as follows:

$$S(\bar{y}_{loc}) = \frac{\sigma^2_Y}{n_{loc}} \qquad (2.1.4)$$

---

[16] This type of survey is generally centralised, since it is more efficient to concentrate the know how and the information needed to perform it.

[17] From now on *LIS*, *LOC* and *MXD* will be used as shortcuts to identify each type of survey. Furthermore, capital letters (Y, D, X) will be used to indicate the population values, while small ones (y, d, x) will be correspondingly used for sample values.

[18] See Cochrane (1974), p. 22.

[19] See Cochrane (1974), p. 23.

2.2 Centralised survey of list prices (LIS)

List prices are generally collected with a centralised production process: the sample, and in particular its coverage, may be much larger if compared with *LOC* surveys and even nearly coincide with the entire population. An estimate of population mean is then obtained as follows:

$$\bar{y}_{lis} = \frac{\sum x_i}{n_{lis}} = \bar{x}_{lis} \quad (2.2.1).$$

The bias of this estimate is then derived by averaging the difference between sample and population mean (the latter being derived from (1.3)):

$$B_{lis} = E\left[\bar{y}_{lis} - \bar{Y}\right] = E\left[\bar{x}_{lis} - \overline{XD} - \sigma_{XD}\right] = \left(1 - \bar{D}\right)\bar{X} - \sigma_{XD} \quad (2.2.2).$$

The bias includes two effects. The first depends on the dynamic of the discount policy: if this policy varies on average (e.g.: an increase in the correction coefficient applied by retailers, that is $\bar{D} > 1$), then expression (2.2.1) will bring an underestimate of population mean since it is not able to account for this change of policy. The second is given by the covariance between X and D, which cannot be measured if only X is collected.

Based on (2.2.2), the average quadratic deviation from the population mean can be expressed as follows:

$$S\left(\bar{y}_{lis}\right) = \frac{\sigma_X^2}{n_{lis}}\left(1 - f_{lis}\right) + \left[\bar{X}\left(1 - \bar{D}\right) - \sigma_{XD}\right]^2 = \sigma^2_{\bar{x},lis} + B^2_{lis} \quad (2.2.3).$$

The first term of (2.2.3) corresponds to the variance of the sample mean of X. Since we expect a huge sample size, its value is very likely to be near to zero, especially because sampling fraction tends to unity. In this case the first term of the right hand side of expression (2.2.3) can be ignored. Hence:

$$S\left(\bar{y}_{lis}\right) \cong B^2_{lis} \quad (2.2.4).$$

2.3 Mixed independent estimates of discount and list prices

Let's assume in this case that list prices are observed for the entire population while effective prices are observed on a local sample only, whose size is n_{mix}. Suppose moreover that the two surveys are run independently, so that by means of list prices we measure $\bar{X}$ while with the sample we obtain an estimate of $\bar{D}$. In this case:

$$\bar{y}_{mix} = \bar{X}\frac{\sum d_{ij}}{n_{mix}} = \bar{X} * \bar{d}_{mix} \quad (2.3.1).$$

The bias in this case is given by:

$$B_{mix} = E\left[\bar{y}_{mix} - \bar{Y}\right] = E\left[\bar{X} * \bar{d}_{mix} - \overline{XD} - Cov\left(X, D\right)\right] = -\sigma_{XD} \quad (2.3.2).$$

Comparing this result with B_{lis} (2.3.2), only one component of the bias is present: the one depending on the interaction between X and D and which cannot be measured by means of the estimator (2.2.1). In this case, the average quadratic distance of (2.3.1) from the population mean is:

$$S(\bar{y}_{mix}) = \frac{\sigma_D^2}{n_{mix}}(1 - f_{mix}) + \sigma_{XD}^2 = \bar{X}^2 \sigma_{\bar{d},mix}^2 + B_{mix}^2 \qquad (2.3.3).$$

2.4 Mixed covariance-corrected estimates of discount and list prices (MXD)

Given the results derived just above in (2.3.2) and (2.3.3), and in particular given the bias that affects the mixed estimates, it seems possible to introduce an estimate of the correction factor needed to bring that bias to zero. We introduce, more specifically, a correction on (2.3.1). Moreover, we differentiate from the case treated in paragraph 2.3 by considering the case in which a sample is used also to estimate the dynamics of list prices. As a consequence, two samples are used for this type of survey: a large sample of size n_{mxd}^x for list prices and a smaller sample whose size is n_{mxd}^d for bargain prices: the corresponding sampling fractions will be presumably strongly different. The two samples, the one used for list prices and that used for bargain prices, are independently drawn. By means of the first we estimate $\bar{X}$ while with the latter we estimate $\bar{d}$ and *Cov(xd)*. In particular:

$$\bar{y}_{mxd} = \bar{x}_{mxd} * \bar{d}_{mxd} + \sigma_{XD} \qquad (2.4.1)$$

Consequently:

$$B_{mxd} = 0 \qquad (2.4.2).$$

The variance of the sample mean is then:

$$S(\bar{y}_{mxd}) = \bar{X}^2 \sigma_{\bar{d},mxd}^2 + \bar{D}^2 \sigma_{\bar{x},mxd}^2 + E\left[(\sigma_{xy} - \sigma_{XY})^2\right] \qquad (2.4.3),$$

where the last term (the variance of the sampling covariance) can be assumed close to zero[20].

3. An evaluation of the different approaches to sampling

3.1 Comparing formulas

A comparison between *LOC* and *LIS* surveys can be performed by confronting expressions (2.1.4) and (2.2.4). If we assume the absence of linear relations between X and D, and if we consider $\sigma_{\bar{x},lis}^2$ very close to zero, we can derive that:

$$S(\bar{y}_{lis}) > S(\bar{y}_{loc}) \Leftrightarrow (1 - \bar{D}^2)\bar{X}^2 < \frac{\bar{X}^2 \sigma_D^2 + \bar{D}^2 \sigma_X^2}{n} \Leftrightarrow \frac{(1 - \bar{D}^2)}{\bar{D}^2} < C_D^2 + C_X^2 \qquad (3.1.1)$$

[20] This result derives from the fact that the samples used in *LIS* and *MXD* are independent.

where $C^2_{\overline{D}}$ and $C^2_{\overline{X}}$ are the squared coefficients of variation of the corresponding sample means. In other words, *LIS* surveys deliver better results as long as the variability of retail policies is below a threshold whose value depends directly on the variability of *X* and *D*. The higher they are, the higher will be that threshold, since the results of *LOC* surveys becomes increasingly volatile. If a linear relation exists between the two pricing components, its presence tends to make *LOC* surveys more preferable since the *LIS* bias is, as a consequence, relatively higher.

The comparison between *LOC* (or *LIS*) and *MXD* approaches is instead relatively trivial, since *MXD* uses more information. Quite clearly, larger gains of precision will be obtained with respect to *LOC* surveys if the variance of list price indices is relatively high.

We already know that with *LOC* surveys estimates are unbiased with a confidence interval that depends on the variance of the sampling means of X and D, and on the sampling size. This last is presumably kept small in absolute and relative terms by the high burden of price collection: the sampling fraction is presumably very close to zero. In the case of *LIS* surveys confidence intervals for the estimates of the dynamics of list prices will be very small, mainly because the sampling fraction will be close to unity. Moreover it is reasonable to expect a small variability in list prices and, consequently, more precise estimates for this component. These are biased estimates of effective price changes if retailers modify their pricing policy: the bias can also be eventually influenced by any covariance between their policy and the actual dynamics of list prices. Quite obviously, the more $\overline{D} \neq 1$ and/or $\rho_{XD} \neq 0$, the more *LIS* surveys are not a good solution. Finally, *MXD* surveys with covariance correction deliver unbiased estimators with lower confidence intervals with respect to local surveys. The highest gain in efficiency can be obtained if the variability of list prices is relatively high.

3.2 Some simulations with random populations

Other elements can be inductively derived from the results of some simulations: the three approaches to sampling have in fact been tested taking into account some different hypotheses concerning the nature of the target population.

In particular a set of artificial reference populations have been created starting from a benchmark random population (Simulation 1 in Table 1) defined as follows: (a) it is composed by 250 models in the market and about 12 thousands transactions referred to these models in the base and in the current month; (b) the list prices of the models are increased on average by 2.5% from month *0* (base) to month *m*: the rate of increase of the price of each model is uniformly distributed, between a minimum of 0 and a maximum of 5%; (c) retailers apply in each period and in each transaction an average 5% discount on the corresponding list price: for each transaction the rate of discount is uniformly distributed between a minimum of 0% and a maximum of 10%; (d) a set of 300 independent samples are drawn from this population: each run of sampling produces a sample for *LOC* survey, a sample for *LIS* survey, while the *MXD* is derived from the joint use of the first two samples; (e) in particular, the *LOC* survey is based on a random sample extracted with a sampling fraction of 1% of all transactions, while the *LIS* survey uses a sample that includes about 80% of the models.

Other four random populations are drawn from this same structure by simply changing the characteristics of the pricing policy, concerning namely the dynamics of the list prices and that of the discount policies. The objective is to simulate the performance of each type of survey in all such populations.

Table 1. - Distinguishing features of the random populations

Type of survey	Simulation 1	Simulation 2	Simulation 3	Simulation 4	Simulation 5
Maximum increase in list prices	5.0%	**20.0%**	5.0%	5.0%	5.0%
Distribution of discounts in the reference month	uniform	uniform	uniform	uniform	uniform
Distribution of discounts in the current month	uniform	uniform	uniform	**Correlated with X^2**	**constant within model and uniform between models**
Maximum discount rate in the current month	10.0%	10.0%	**5.0%**	10.0%	10.0%

Table 2 and Chart 1 show some results obtained from these simulations. Simulation 1 will show on average no correlation between *X* and *D* and a nearly zero dynamics of the discount policy (that is $\overline{D} \cong 1$). Since the *LIS* bias is negligible, this type of survey is expected to perform quite well. The resulting 2.53% increase in Y is determined by a 2.47% increase in list prices X and a 0.06% increase due to the average change in retail policies. Covariance and correlation are close to zero. The variability of the discount policy *D* is almost ten times larger than that of list prices *X*.

The estimates produced by the 300 *LIS* samples (see also Chart 1) - although they are slightly biased - present a very small range in the sample means: the range between 97.5th and the 2.5th percentile in the distribution of the sample means is 0.207. On the contrary, while the estimates obtained with *LOC* surveys are unbiased, their range is much larger than *LIS* survey's, by a factor of more than 7. The performance of *MXD* surveys looks better than the *LOC* ones, but anyway very close to it: in particular, *MXD* shows lower extreme values in the distribution of the population mean.

This picture changes slightly if the variability of *X* is increased. In the second simulation the variability of list prices is much higher than in simulation 1. As a consequence, *MXD* surveys perform now much better than *LOC* surveys. In fact, in this simulation the largest gains derive now from the better coverage of *X* variability that is offered by *LIS* surveys. The differences between the corresponding percentiles of the distribution of the sample estimates derived from *MXD* surveys appear then much reduced with respect to *LOC* surveys (nearly 25% less).

In simulation 3 the hypothesis of zero dynamics in the discount policy has been removed. In particular, with respect to simulation 1 it has been assumed that the maximum rate of discount decreases from 10% in the reference month to 5% in the current one: this implies an average increase in the prices applied by retailers. Discounts keep on being applied randomly: so covariance and correlation are still negligible. The index *Y* showed a 5.24% increase, determined by a 2.45% increase in *X* and a decrease of 2.72% in *D*. This change in retail policy implies a strong increase in the bias deriving from the exclusive use of list prices. In particular the estimate based on price lists makes an average error of more than 2.65% and presumably a much higher mean square error with respect to the other survey techniques. *LIS* survey, in fact, only records the increase in list prices, and cannot measure the fact that effective prices increase faster because the average rate of discount has also decreased.

Table 2. - Main sampling indicators, by type of survey and simulation

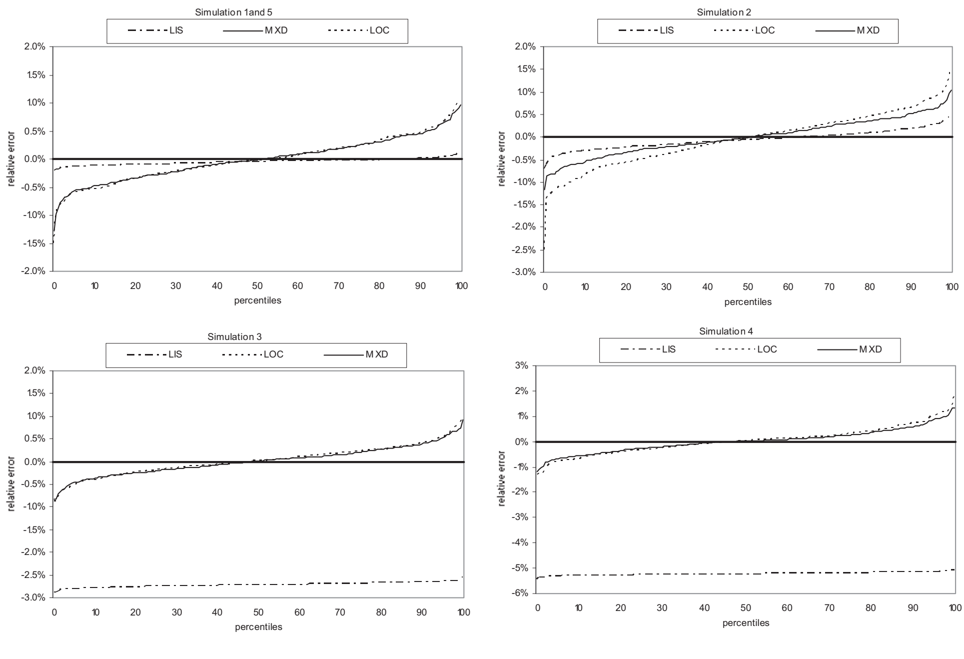
Type of survey	Bias (%)	95% interval of the mean	Mean(Y)	Mean(D)	Mean(X)	V(D)	V(X)	R^2	Sample size
				Simulation 1					
LIS	0.06	0.207	2.48%		2.48%		0.00021		9507
LOC		1.546	2.52%	0.05%	2.47%	0.00182	0.00021	-0.006	120
MXD		1.388	2.53%	0.05%	2.48%	0.00182	0.00021	-0.006	120
Population			2.53%	0.06%	2.47%	0.00185	0.00021	-0.011	11938
				Simulation 2					
LIS	0.07	0.749	9.63%		9.63%		0.00326		9465
LOC		2.172	9.65%	0.06%	9.58%	0.00180	0.00323	0.013	119
MXD		1.530	9.70%	0.06%	9.63%	0.00180	0.00326	0.013	119
Population			9.71%	0.07%	9.63%	0.00181	0.00327	0.002	11891
				Simulation 3					
LIS	2.65	0.193	2.46%		2.46%		0.00020		9622
LOC		1.310	5.26%	2.73%	2.46%	0.00118	0.00020	0.020	122
MXD		1.252	5.26%	2.73%	2.46%	0.00118	0.00020	0.020	122
Population			5.24%	2.72%	2.45%	0.00119	0.00020	0.015	12081
				Simulation 4					
LIS	4.89	0.216	2.50%		2.50%		0.00022		9964
LOC		2.167	7.89%	5.21%	2.50%	0.00293	0.00021	0.555	126
MXD		1.796	7.89%	5.21%	2.50%	0.00293	0.00022	0.551	126
Population			7.87%	5.18%	2.51%	0.00297	0.00022	0.562	12508
				Simulation 5					
LIS	-0.01	0.212	2.39%		2.39%		0.00021		9600
LOC		1.568	2.37%	-0.01%	2.38%	0.00191	0.00021	-0.046	120
MXD		1.419	2.38%	-0.01%	2.39%	0.00191	0.00021	-0.046	120
Population			2.37%	-0.01%	2.39%	0.00192	0.00021	-0.040	12065

Notes: (a) The interval is given by the difference between the 97.5^{th} and 2.5^{th} percentile of the distribution of the sample means.
(b) The means of Y, D and X have been expressed in terms of rates of change (between the reference and the current month) and not of index.
(c) The columns *V(...)* report the estimate of the variance of the variable between parentheses.

In simulation 4 some covariance between list prices and retail policies is introduced: in particular, it is assumed that discount corrections in the current period are correlated with the level of X^2. This implies that a change has taken place in the retail policy from the base and the current month (as it happened in simulation 3), and in particular that the effect of the correlation between list and retail policies must be taken into account. In particular, it has been assumed that maximum discount in period *0* is 10% and that in the current period the retail correction can be increased proportionally to the square of *X*: a random effect is also included. This correction implies a strong growth in the *D* index, that is a price increase stronger than the one registered by list prices. The table also shows a higher value for the correlation coefficient, which now surpasses 0.5. This effect mainly implies that the bias of *LIS* estimates is quite larger then in former simulations, while the comparative advantage of *LOC* and *MXD* remains substantially unaltered. *MXD* keeps on performing better as far as precision is concerned.

A last variant of simulation 1 can concern the case in which retail policies vary at model level. We assume that discount is uniformly distributed between 0 and 10% in period 0, while in the current period its pattern changes: it is fixed for each transaction involving a same model, and changes only from model to model according to as uniform distribution. In terms of consequences for the estimates, this case is very similar to that presented in simulation 1. In fact, whatever correlation between discount policies and any variable uncorrelated with *X* does not induce any bias in the estimates.

Chart 1. - Percentiles of the distribution of the relative error of the sample mean, by type of survey and simulation



4. Some generalisations

It seems that *MXD* surveys guarantee generally better performances, by allowing large improvements to the results of *LOC* surveys. *LIS* surveys are instead very weak since they need very strict hypotheses on retailers behaviour in order to be considered viable. As concerns *LOC* surveys, they need conditions of low variability of list prices to give appreciable results. Several hypotheses have been done anyway to reach these conclusions with some formalisation. It is worth the while to spend a few words on the consideration of possible generalisations of these results, trying in particular to draw some hints on the consequence of relaxing the hypothesis of one-to-one correspondence between the transactions of the reference and of the current month. Other aspects might interfere meaningfully with the production of such estimates, as for example the consideration of the desired time horizon of our estimates (monthly, quarterly or yearly indices) and the consideration of the possibility of adopting a more articulated sampling design. Furthermore, it must also be kept in due account the fact that many of the conditions that we have derived in sections 2 and 3 may be valid in a specific month but not in the generality of the twelve months

4.1 The correspondence between transactions

As summarised in expression (3.1.1), the comparative evaluation of the three survey methods (*LOC, LIS* and *MXD*) depends crucially on the relative variability of the indices of list prices and retail policies (*X* and *D*). In general, the advantages of adopting *LIS* surveys are higher the higher is *X* variability. If we consider the case of new cars, such results seem to imply that the CPI for cars ought to be estimated on the basis of a *LOC* survey, since list prices are very stable. But if we recognise that the variability of list prices depends not merely on the dynamics of the list price attached to each model but also on models' turnover than things might change a lot. The variability of list prices in presence of a high model turnover may increase strongly and may push in favour of the use of *LIS* surveys.

In the simulation described in section 3 only the first source has been considered. But the change in the product range may be very influential on *X* variability, independently of which quality adjustment technique is being implicitly or explicitly adopted. In fact the change in the product range impairs the one-to-one correspondence and this implies the necessity to operate replacements and to compare the prices of different models within a same segment. This source of higher variability of *X* implies relatively more advantages in the use of *LIS* surveys, especially in the *MXD* approach, while the defects of *LOC* surveys emerge more clearly. *LIS* surveys in fact give the possibility to exploit large samples and more articulated sample designs. It may guarantee, for example, a good segmentation through stratification: *LOC* surveys do not give this possibility. Surveying systematically list prices offers the possibility to deliver a better representation of the turnover of models and of the product range, and more generally guarantees a better picture of the consumer market, which may be difficult to obtain with the smaller sample size implied by *LOC* surveys[21].

[21] A better adaptation to HICP Regulation 1334 (see European Commission, 2007) and to the consumption segments approach is also guaranteed by LIS or MXD surveys.

4.2 Time horizon of the estimates

LIS surveys are clearly affected by a bias when retail policies change or when a linear correlation exists between these policies and the dynamics of list prices. This mere fact makes *LIS* surveys a weak short term price indicator. But what can we say about their longer term performance?

The answer to this question depends on the nature of retailer policies. Since it is difficult to imagine systematic trends (increasing or decreasing) in the behaviour of the index *D*, the issue is to know how it varies, how regular or cyclical is its behaviour, how it affects year-on-year price changes.

What we can expect is that a yearly average of the 12 monthly price indices may be less affected by changes in discount policies and by correlation with list prices. The choice of the survey technique depends then also on the target of our estimates: the choice to be made to make a good quarterly index may be quite different from the case in which the target is a monthly index. The longest the time horizon the less *LOC* surveys are useful.

There is a further complication: any regularity in the behaviour of prices, must be measured with an instrument – the price index – whose base is defined with reference to a single month (December t-1), at least in the HICP context. As a consequence, any weakness in the method that we apply may affect the base month and all the estimates referred to the current year. This problem is part of the categories of problems connected for example to seasonal products. If the month of December represents a particular month in the seasonal evolution of indices, it may happen that in this base month estimates error and confidence intervals are larger[22]. The use of *LIS* or *LOC* surveys may, for opposite reasons, suffer the same drawback, while *MXD* surveys represent a safer solution.

4.3 Sampling design

Discount policies may depend on short term producers' policies or on point of sale policies. There exists a huge literature on this subject[23]. As we have seen before (see simulation 5), should there exist any correlation between these policies and other variables different from list prices indices[24], these would not imply a bias on *LIS* estimates. Nevertheless, the existence of these correlations may be helpful while deciding the sample design to be adopted, since it may help to define an efficient stratification especially for *MXD* surveys.

In this type of approach, two distinct samples are drawn: a relatively small *LOC* sample and a huge *LIS* sample. The sample designs of these two surveys may be drafted in order to guarantee an efficient usage of formula (2.4.1) at stratum level. A deeper study of the behaviour of retailers can give the possibility to study an efficient stratification in order to pass rapidly from a *LIS* survey to a *MXD* survey. It could be useful to test which structural variables mainly influence bargaining. If these variables have to do with the characteristics of customers (and much of the literature stresses this point) then the use of local surveys looks less promising since the characteristics of price collectors may induce a bias.

[22] See De Gregorio, Munzi et al. (2007).

[23] From different perspectives, see for example Sweezy (1939) , Jung (1959) , Goldberg (1996), Zettelmayer (2003), Scott Morton (2003), Marn et al. (2003).

[24] Retailers' policies may be related, for example, to the customer, to point-of-sale strategies, to the kind of vehicle, to the brand, to the age of the model, etc.

4.4 Monthly estimates

The relative advantage of *LIS*, *LOC* and *MXD* surveys may not be constant during the whole year. We have implicitly adopted Martin Ribe's approach to CPI estimates and sampling design[25], in which the unchained index is used as the parameter to be estimated. In a HICP context, this means that the parameter is the population monthly index based on December of the previous year. In this theoretical context, the problem of sampling deals with the definition of at least 12 distinct sample designs, one for each month.

Also in the cases where no seasonal effect is present, indices' variability is generally not constant across months, and changes systematically the more we get far from the base month. In the case of list prices, their variability depends in fact crucially on the time horizon of the index. It is reasonable to assume that, for each model, list prices tend to vary not very frequently during the year. The more the current month is far from the base the higher the probability that the list price of a given model has changed and thus the higher might be the variability of *X*, especially if we consider the effects of replacements. Taking December year t-1 as the base month, there will be a huge diversity between the profile of *X* variability that is registered in January and in December year t. In January most of the units will probably show no price change and the variability will be very low although concentrated in a few models. In December the price changes which have taken place in the eleven months before will have cumulated and will affect reasonably most of the models in the population. These differences do not only imply the heterogeneity of the monthly patterns of variability: what may be different is also the distribution of the observations, which is probably very far from approximating well the normal or the log-normal distributions in the first months of the year.

[25] Ribe (2000).

Concluding remarks

Explicit formalisations of the sampling design to be adopted for CPI estimates are quite rare in specialised literature[26]. In the preceding sections we have proposed some formalisations based on probability sampling in the simplified context of one-to-one correspondence between transactions in the reference and in the current month. This in order to deliver a base for analysing a dichotomy which is quite frequent in price statistics and which can be expressed as follows: Is it better to use a centralised survey of list prices (that is, a survey made in the statistical office by collecting and storing official listings) or to opt for a local price collection of bargain prices? Is it better a huge sample of models and a collection of prices which are not the real ones, or a small sample of actual transactions?

One first suggestion is that the dichotomy between adopting *LIS* or *LOC* surveys appears to some extent a too simplistic way to face the issue of producing a reliable estimate of the CPI. The use of mixed surveys (*MXD*), where *LIS* and *LOC* surveys are both used each month, appears to be a more complete tool to estimate price indices in markets where list prices are part of the pricing policy. But, as we have seen in paragraph 2.4, in order to produce reliable estimates and to avoid any systematic bias, *MXD* estimates must include a covariance correction, to be estimated by means of local surveys (see expression (2.4.1)). The conditions by which *LIS* or *LOC* surveys become more efficient are very likely to be changing with months and with circumstances due to sudden changes in the variability of list prices and discount policies.

There are not only strictly technical reasons behind the superiority of *MXD* surveys: it is legitimate to expect that the relaxing of the strict hypotheses that we have adopted to simulate the sample design bring further elements in support to *MXD* solutions. The main advantage that they convey – which is proper, although to a minor extent, also of *LIS* surveys - lays in the possibility to offer a more complete monitoring of consumer markets, since they give the possibility to trace the main changes in the range of product available to consumers and, as a consequence, to manage more properly replacements. Under this point of perspective, it seems that *MXD* surveys comply in a more satisfying way with the recent trends in applied CPI theory, and in particular with new concepts such as the consumption segments, which have been introduced recently within the HICP project with the objective to pave the way for a proper definition of the population parameter and, as a consequence, for more coherent sampling designs[27].

LIS surveys are more frequently adopted, typically in the case of new cars. Nevertheless, the adoption of *LIS* surveys is in general risky, and in particular for short term estimates. In fact, very strict conditions have to be fulfilled in order to guarantee the production of unbiased estimates: these assumptions concern in particular the dynamics of retailers correction on list prices. *LOC* surveys loose their efficiency if list price variability is relatively high, that is if a medium term horizon (generally some months, often one year or more) is considered.

Market analysis, market-specific and country-specific evaluations should then be made concerning the components of price and price index variability, and to support the choice of

[26] The work of Ribe (2000) gives an important contribution on this subject and, although the solution is still incomplete, it paves the way for further improvements. The most tricky aspect has to do, in particular, with the definition of the reference population and, consequently, of the parameter to be estimated.

[27] European Comission (2007).

a method. Such analyses could in particular help to model the variability of discount policies in order to provide a framework for improving sampling designs and to set the *LOC* survey in the context of a *MXD* approach, and thus to reduce its burden.

Further efforts should also be dedicated to compare the alternative approaches to CPI surveys, by means of more complex sample designs, based on stratified or multi-stage samples. The comparative advantage of *MXD* solutions might in fact be further qualified and quantified.

References

Cochrane W.G. (1977), *Sampling techniques*. 3rd edition. John Wiley & sons, New York, p.428.

De Gregorio C. (2006), "Il nuovo impianto della rilevazione centralizzata del prezzo dei medicinali con obbligo di prescrizione", *Contributi Istat*, n.16.

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.

European Commission (2007), "Commission Regulation (EC) no 1334/2007 of 14 November 2007 amending Regulation (EC) No 1749/96 on initial implementing measures for Council Regulation (EC) No 2494/95 concerning Harmonised Indices of Consumer Prices". *Official Journal of the European Union*, 15 Nov., L296/22-26.

Eurostat (2005), "HICP standards for cars and other vehicles". *Paper by Eurostat and the Task Force "Implementation Group on Quality Adjustment and Sampling"*. HCPI 05/520 Rev.1 Final.

Genesove D., Mayer C., (2001), "Loss aversion and shelter behaviour: evidence from the housing market". *The Quarterly Journal of Economics* (Nov.), Vol.116, No. 4, 1233–1260

Goldberg P. K., (1996), "Dealer price discrimination in new car purchase: evidence from the consumer expenditure survey". *The Journal of Political Economy*, Vol. 104, Issue 3 (June), 622-654.

Horowitz J.L., (1992), "The role of the list price in housing markets: theory and an econometric model", *Journal of Applied Econometrics*, vol. 7, 115-12.

Istat (2007), *The characteristics of Istat CPI centralised surveys*. Mimeo, Istat.

Jung A.F., (1959), "Price variations among automobile dealers in Chicago, Illinois", *The Journal of Business*, Vol.32, No.4, 315-326.

Knight J.R., Sirmans C.F., Turnbull C.K., (1998), "List price information in residential appraisal and underwriting", *Journal of Real Estate Research*, Vol.15, Nos.1/2, 59-76.

Lutz M. (2004), "Pricing in Segmented Markets, Arbitrage Barriers, and the Law of One Price: Evidence from the European Car Market". *Review of International Economics*, 12(3), 456–475.

Marn M.V., Roegner E.V., Zawada C.C., (2003), "The power of pricing", *The McKinsey Quarterly*, No.1, 27-36.

Ribe M. (2000), "Price index, statistical inference and harmonisation". *Proceedings of the 4th International Conference on Methodological Issues in Official Statistics*. 12-13 October.

Scott Morton F., Zettelmeyer F., Silva-Risso J. (2003), "Consumer Information and Discrimination: Does the Internet Affect the Pricing of new cars for woman and minorities?" *Quantitative Marketing and Economics*, March, Vol.1, No.1, 65-92.

Sweezy P., (1939), "Demand under conditions of oligopoly", *The Journal of Political Economy*, Vol. 47 No.4 (Aug), pp.568-573.

Thanassoulis J., (2005), "List prices, bargaining and resultant productivity diffusion delay". *Discussion paper,* series n.220, Department of economics, Oxford university.

Zettelmeyer F., Scott Morton F., Silva-Risso J. (2003), "Inventory Fluctuations and Price Discrimination: The Determinants of Price Variation in Car Retailing", Yale School of Management, manuscript.

Uno studio della non autosufficienza a partire dai dati dell'Indagine Multiscopo: il caso dell'Umbria

Giorgio E. Montanari[1], *M. Giovanna Ranalli*[2]

Sommario

In questo lavoro si propone una metodologia per affrontare il problema della stima del numero delle persone non autosufficienti a partire dai dati dell'Indagine Multiscopo dell'ISTAT sulla salute dei cittadini ed il ricorso ai servizi sanitari, con una applicazione alla regione dell'Umbria. Qui si assume che la non autosufficienza sia descritta da un tratto latente sottostante al set di indicatori utilizzati dall'indagine che rilevano, fra le altre, le difficoltà di svolgere le attività elementari della vita quotidiana (Activities of daily living). A tal fine si utilizza un modello a classi latenti per la classificazione della popolazione secondo diversi gradi di non autosufficienza. L'analisi fornisce una classificazione in quattro classi. Sulla base delle probabilità a posteriori, gli individui appartenenti a ciascuna classe sono definiti Autosufficienti, Disabili lievi, Non autosufficienti iniziali e Non autosufficienti. L'indagine Multiscopo consente di ottenere stime affidabili fino al dettaglio regionale. Per la stima della dimensione delle classi a livello sub-regionale, che si configura, quindi, come un problema di stima per piccole aree, sono stati impiegati modelli multinomiali in cui la probabilità individuale di appartenere a ciascuna delle quattro classi dipende da covariate disponibili per tutta la popolazione.

Abstract

This paper proposes a methodology for the estimation of the number of people that show a severe disability and are dependent, using data coming from the Italian National Survey on Health conditions and Appeal to Medicare. Dependency is treated as a latent trait hidden behind a set of items that survey difficulties in movements and in accomplishing everyday tasks (Activities of daily living). Latent class models are used to classify the population according to different levels of disability. The analysis provides a good classification using four classes. Looking at posterior probabilities, people belonging to each class may be labelled as being without disability, with light disability, with some dependence, with severe disability (dependent). The survey provides reliable estimates at regional – NUTS 2 – level. Estimating the amount of population within each latent class at sub-regional level, e.g. sanitary districts, requires small area estimation techniques. To this end, a multinomial unit level model is used with individual level covariates.

Parole chiave: Variabili latenti, modelli a classi latenti, stima per piccole aree.

[1] Dipartimento di Economia, Finanza e Statistica, Università degli Studi di Perugia, giorgio@stat.unipg.it
[2] Dipartimento di Economia, Finanza e Statistica, Università degli Studi di Perugia, giovanna@stat.unipg.it

1. Introduzione

Nel presente lavoro si propone una metodologia per affrontare il problema della quantificazione del numero delle persone non autosufficienti con riferimento alla regione dell'Umbria e a partire dai dati dell'indagine Multiscopo condotta dall'ISTAT su *Condizioni di salute e ricorso ai servizi sanitari*. I dati di questa indagine costituiscono l'archivio più strutturato ed omogeneo al momento disponibile sul fenomeno della disabilità in Italia e in Umbria. Com'è noto, nell'impostare la rilevazione, l'ISTAT ha fatto riferimento alla Classificazione internazionale ICIDH (*International Classification of Impairments, Disabilities and Handicaps*) elaborata dall'Organizzazione Mondiale della Sanità (WHO, 1980). Secondo tale impostazione, con riferimento alle condizioni di salute di una persona, per *menomazione* si intende qualsiasi perturbazione mentale o fisica del funzionamento del corpo; con il termine *disabilità* si denota, invece, qualsiasi limitazione o perdita, conseguente a menomazione, della capacità di compiere un'attività nel modo o nell'ampiezza considerati normali per un essere umano; infine, con *handicap* si vuol indicare quella situazione di svantaggio nella partecipazione sociale vissuto a causa della presenza di menomazione o disabilità. Nell'indagine Multiscopo si considera disabile la persona che, escludendo limitazioni temporanee, dichiara il massimo grado di difficoltà in almeno una delle funzioni rilevate con specifiche domande (*items*), pur con l'ausilio di apparecchi sanitari (protesi, bastoni, occhiali, ecc; cfr. ISTAT, 2002; p. 142).

La disabilità così come viene rilevata dall'indagine Multiscopo non è, tuttavia, sinonimo di non autosufficienza nel senso più grave che qui interessa per le implicazioni in termini di servizi sociosanitari da organizzare sul territorio per farvi fronte. La disabilità, infatti, assume forme diverse per tipo di attività compromessa e per gravità. La sua misurazione viene di solito effettuata sulla base di scale multidimensionali e di indicatori complessivi, che, con grande approssimazione, tentano di ricondurre ad una dimensione semplice un fenomeno estremamente complesso. In questa sede si cercherà di identificare, dunque, i casi di disabilità considerati dall'indagine Multiscopo che possono essere considerati rappresentativi di situazioni di elevata gravità e, quindi, di non autosufficienza.

Molte sono le definizioni di non autosufficienza ed una rassegna sintetica, con particolare riferimento all'Italia, si può trovare in ISTAT (2004). Il tratto comune ai vari tentativi di definizione fa riferimento alla difficoltà a svolgere le "ricorrenti azioni quotidiane" (alzarsi da un letto o da una sedia, lavarsi, vestirsi, ecc.) ed alla necessità di assistenza da parte di altre persone per svolgere tali azioni. Anche in ISTAT (2004) si utilizzano i dati dell'indagine Multiscopo per tentare di stimare il numero delle persone in condizioni di non autosufficienza. In particolare vengono esaminati due metodi. Il primo è basato sull'analisi delle abilità che si perdono per ultimo e quindi proxy di una condizione di non autosufficienza, senza però dare un criterio su quante considerarne per definire una persona non autosufficiente. Il secondo metodo utilizza l'analisi delle corrispondenze multiple sugli item rilevati, che, tuttavia, non sembra fornire un'adeguata spiegazione della variabilità dei dati: i primi cinque fattori spiegano il 59% dell'inerzia totale; inoltre anche in questo caso viene fissata una soglia arbitraria lungo il primo asse fattoriale, rappresentativo della non autosufficienza, per classificare una persona come non autosufficiente.

In questo lavoro si propone un metodo di classificazione secondo il livello di gravità della disabilità rilevata, impiegando la modellistica già usata per la popolazione anziana dell'Umbria in Montanari (2005) e Montanari *et al.* (2006). A questo scopo si utilizzano i dati dell'indagine Multiscopo 1999-2000, di cui si è detto, relativi però all'intera popolazione e si perviene alla stima dei contingenti di popolazione in condizioni di progressiva non autosufficienza nella Regione dell'Umbria. Inoltre, si riportano alcuni risultati dello studio sulle variabili connesse ad una condizione di non autosufficienza e si propone una metodologia di stima della non autosufficienza per piccole aree.

Il lavoro è così organizzato. Il paragrafo 2 presenta i dati utilizzati e riprende l'analisi in classi latenti impiegata da Montanari (2005) e Montanari *et al.* (2006); questa viene quindi utilizzata per stimare il numero delle persone non autosufficienti nella popolazione umbra. Nel paragrafo 3 sono riportate le stime della diffusione della non autosufficienza nelle Aziende Sanitarie Locali dell'Umbria (ASL) e per classi di età. Il paragrafo 4 contiene alcune considerazioni finali.

2. Individuazione delle persone non autosufficienti

2.1. La rilevazione della disabilità nell'indagine Multiscopo

I dati utilizzati in questo lavoro provengono dall'indagine Multiscopo dell'ISTAT sulle condizioni di salute e il ricorso ai servizi sanitari della popolazione italiana nel biennio 1999-2000. La scelta di questo biennio è dettata unicamente dal fatto che si tratta dei soli dati messici a disposizione con il dettaglio comunale necessario per l'analisi condotta nei paragrafi seguenti. Si tratta di un'indagine campionaria che in Umbria ha coinvolto circa 1.800 famiglie per un totale di 4.900 individui distribuiti in 51 comuni campione. L'indagine in questione considera disabile la persona che, escludendo le condizioni riferite a limitazioni temporanee, dichiara il massimo grado di difficoltà in almeno una tra una serie di attività, incluse quelle della vita quotidiana (ADL), pur tenendo conto dell'eventuale ausilio di apparecchi sanitari come protesi, bastoni, occhiali, ecc. Le ADL (*Activities of Daily Living*) sono state proposte inizialmente da Katz *et al*. (1963) e consistono in un insieme di quesiti relativi alla capacità della persona di espletare azioni quali lavarsi, vestirsi, mangiare da sola, ecc. Riferimenti più aggiornati sono Branch e Meyers (1987) e Wiener (1990). La batteria di quesiti utilizzata da ISTAT segue le direttive fornite dal Consiglio d'Europa e dall'Organizzazione Mondiale della Sanità (ISTAT, 2002; pag. 63). A seconda della sfera di autonomia funzionale compromessa, sono state costruite le seguenti quattro tipologie di disabilità:

- confinamento,
- difficoltà nel movimento,
- difficoltà nelle funzioni (della vita quotidiana),
- difficoltà sensoriali.

Per confinamento si intende una costrizione permanente a letto, oppure su una sedia, o nella propria abitazione per motivi fisici o psichici. Le persone con difficoltà nel movimento hanno problemi nel camminare (riescono solo a fare qualche passo), non sono in grado di salire e scendere una rampa di scale, non riescono a chinarsi per raccogliere oggetti da terra. Le difficoltà nelle funzioni della vita quotidiana riguardano l'assenza di autonomia nello svolgimento delle essenziali attività quotidiane o di cura della persona, quali mettersi a letto o sedersi, vestirsi, lavarsi o farsi il bagno o la doccia, mangiare anche tagliando il cibo. Le difficoltà sensoriali includono le limitazioni nel sentire (non riuscire a seguire una trasmissione televisiva anche alzando il volume e nonostante l'uso di apparecchi acustici), le limitazioni nel vedere (non riconoscere un amico ad un metro di distanza) e nel parlare (non essere in grado di esprimersi).

Nella Tavola 1 sono riportati gli item sottoposti agli intervistati, per rilevare l'eventuale esistenza di difficoltà nelle funzioni considerate, e le possibili modalità di risposta proposte agli stessi. Per maggiori dettagli si veda ISTAT (2002). È definita disabile la persona che in almeno uno dei 16 quesiti della tavola opta per la modalità di risposta con il codice numerico più alto. In questo lavoro considereremo le persone che hanno compiuto i 6 anni di età; l'indagine Multiscopo, infatti, non rileva le ADL sui bambini di età inferiore. Dal momento che il campo di indagine è quello delle famiglie, è da tenere presente che nelle considerazioni che verranno svolte non sono inclusi i cittadini istituzionalizzati, quelli cioè che vivono permanentemente in un istituto. Inoltre, occorre tenere presente che non vengono classificate come disabili le persone che soffrendo di una qualche forma di menomazione mentale, anche grave, sono tuttavia in condizioni di svolgere le attività della vita quotidiana.

Tavola 1 - Indicatori impiegati per la rilevazione della disabilità nella indagine Multiscopo ISTAT 1999-2000 e relativa categorizzazione

Sfera di autonomia	Quesito	Modalità di risposta
Confinamento	A1 = Costretto a letto	0 = No 1 = Si
	A2 = Costretto seduto	come per A1
	A3 = Costretto in casa	come per A1
Difficoltà nel movimento	B1 = Distanza più lunga percorribile	0 = Più di 200 m. 1 = Meno di 200 m. 2 = Solo qualche passo
	B2 = Scendere o salire rampa di scale	0 = Si 1 = Con qualche difficoltà 2 = Con molta difficoltà 3 = No
	B3 = Chinarsi a terra	come per B2
Difficoltà nelle funzioni (della vita quotidiana)	C1 = Mettersi e alzarsi dal letto	0 = Senza difficoltà 1 = Con qualche difficoltà 2 = Solo con l'aiuto di qualcuno
	C2 = Sedersi e alzarsi dalla sedia	come per C1
	C3 = Vestirsi e spogliarsi	come per C1
	C4 = Fare il bagno o la doccia	come per C1
	C5 = Lavarsi le mani e il viso	come per C1
	C6 = Mangiare tagliando il cibo	come per C1
	C7 = Masticare	0 = Si 1 = Con qualche difficoltà 2 = Con molta difficoltà 3 = No
Difficoltà sensoriali	D1 = Sentire una trasmissione TV	0 = Si 1 = Solo a volume alto 2 = No
	D2 = Vedere e riconoscere un amico	0 = Si 1 = Solo a un metro di distanza 2 = No
	D3 = Parlare	0 = Si 1 = Con qualche difficoltà 2 = Con molta difficoltà 3 = No

Le unità campionarie residenti in Umbria sono complessivamente 4.879 (inclusi 208 bambini con meno di 6 anni). A ciascuna di esse è associato un peso di riporto all'universo ottenuto tramite stimatori a ponderazione vincolata (Deville e Särndal, 1992). Per maggiori dettagli sulle modalità di determinazione dei pesi campionari si veda ISTAT (2002).

Mediante i pesi campionari è possibile stimare qualsiasi indice statistico di interesse riferito alla popolazione campionata. A questo punto, utilizzando i dati dell'indagine, è immediato calcolare il numero (stimato) delle persone disabili per tipologia di disabilità (sfera di autonomia compromessa) e/o ambito territoriale in Umbria. La Tavola 2 riporta il numero assoluto degli individui di almeno 6 anni classificati come disabili e la prevalenza delle diverse forme di disabilità. Da essa si desume ad esempio che gli individui confinati sono stimati in 16.289, pari al 2,1% degli individui di almeno 6 anni residenti in Umbria (che alla data del 1.1.2000 risultavano essere 788.826). Gli individui disabili in totale, invece, qualunque sia la tipologia della disabilità, sono in tutto 40.206, pari al 5,1% della popolazione con 6 anni e più.

Si osservi che i disabili in totale non sono pari alla somma del numero dei disabili per ciascuna tipologia in quanto queste sono spesso compresenti (in media sono presenti circa due tipologie per disabile). Nella Tavola 3 è riportata la distribuzione dei disabili per numero di tipologie di disabilità presenti. Ad esempio, i disabili con almeno tre tipologie sono 14.419, pari al 1,8% della popolazione di riferimento.

Tavola 2 - Individui di almeno 6 anni disabili per tipologia, in totale e per cento individui di almeno 6 anni residenti in Umbria al 1.1.2000 (totale individui di almeno 6 anni: 788.826; totale popolazione: 827.674)

Tipo disabilità	Numero di unità	% su popolazione con 6 anni e più
Confinamento	16.289	2,06
Difficoltà nel movimento	25.964	3,29
Difficoltà nelle funzioni	27.870	3,53
Difficoltà sensoriali	11.425	1,45
Disabili in totale in Umbria	40.206	5,10

Fonte: Indagine Multiscopo "Condizioni di salute e ricorso ai servizi sanitari", anni 1999–2000

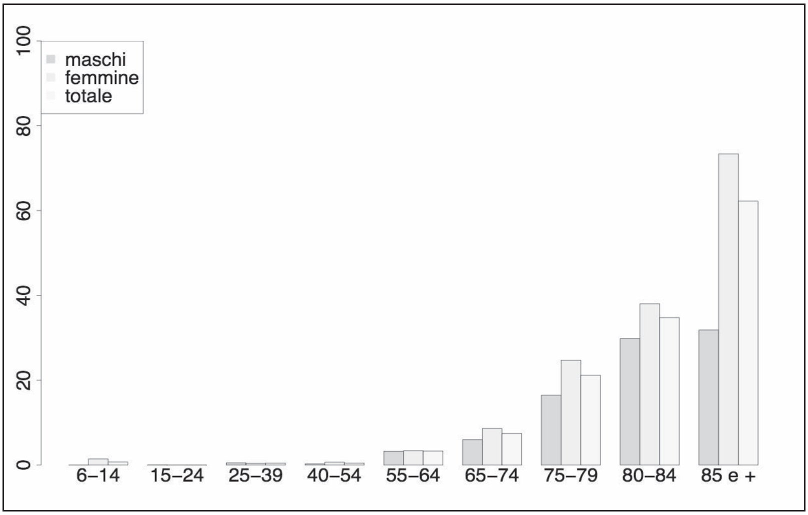
Tavola 3 - Individui con almeno 6 anni disabili per numero di tipologie presenti, in totale e per cento individui di almeno 6 anni d'età residenti in Umbria al 1.1. 2000

Compresenza di tipologie di disabilità	Numero di persone	% su popolazione con 6 anni e più
Una sola tipologia	17.249	2,19
Due tipologie	8.538	1,08
Tre tipologie	10.455	1,33
Quattro tipologie	3.964	0,50
Disabili in totale	40.206	5,10

Fonte: Indagine Multiscopo "Condizioni di salute e ricorso ai servizi sanitari", anni 1999 – 2000

Se si analizza la distribuzione dei 40.206 disabili in Umbria per sesso ed età, si osservano delle notevoli differenze. La Figura 1 mostra l'andamento dei tassi di disabilità per sesso e classi di età. Come si può osservare, i tassi crescono decisamente con l'età e sono generalmente superiori per le donne; in particolare, nell'ultima classe d'età la disabilità è presente nel 74% circa delle donne contro il 32% circa degli uomini. Il dato è spiegabile con la maggiore longevità delle donne che si traduce in un maggior numero di anni in condizioni di disabilità.

Figura 1. - Tassi percentuali di disabilità in Umbria per sesso e classi di età.



2.2. Dalla disabilità alla non autosufficienza attraverso i modelli a classi latenti

In questo lavoro si segue l'approccio sviluppato in Montanari (2005) e Montanari *et al.* (2006) facendo ricorso ai Modelli a Classi Latenti (MCL) individuati mediante l'Analisi delle Classi Latenti (ACL). Questo tipo di modellistica è stato riscoperto più volte in letteratura. Riferimenti base sono Lazarsfeld e Henry (1968) e Goodman (1974); una breve ma chiara introduzione a tale metodologia si trova in McCutcheon (1987). Sviluppi più recenti si possono trovare in Lindsay, Clogg e Grego (1991), Uebersax (1993), Heinen (1996) e Vermunt e Magidson (2002). In particolare, quest'ultimo lavoro tratta delle applicazioni dell'ACL all'analisi dei gruppi (*Cluster Analysis*) dove, a differenza di questa, si postula un modello statistico per la popolazione da cui il campione di osservazioni proviene. Ciò consente di costruire gruppi di unità della popolazione omogenee rispetto alle modalità di risposta delle variabili rilevate e di assegnare ad ogni unità statistica la sua probabilità di appartenere a ciascuno dei gruppi individuati. L'ACL è infatti un metodo statistico atto a trovare sottogruppi omogenei (classi latenti) da dati categorici multivariati.

Tra i risultati dell'analisi c'è anche l'assegnazione delle unità alla classe di appartenenza più probabile. Il vantaggio di questo tipo di analisi consiste nel fatto che le variabili categoriche utilizzate non devono avere necessariamente lo stesso numero di modalità di risposta e ad esse non vengono attribuiti pesi a priori nel determinare l'esito della classificazione. Inoltre, essendo un modello probabilistico, è in grado di gestire anche la variabilità delle risposte dovuta alla valutazione soggettiva connessa alla autovalutazione delle proprie condizioni di salute.

Il MCL è un modello probabilistico in cui la struttura di correlazione tra le modalità di risposta delle variabili rilevate viene modellizzata ipotizzando che condizionatamente al gruppo di appartenenza le risposte date da un soggetto alle diverse variabili categoriche sono indipendenti. Un MCL definisce quindi delle classi latenti in base al criterio della indipendenza condizionata. Ad esempio, all'interno di una classe corrispondente ad una data sindrome, la presenza o assenza di un sintomo è visto come indipendente dalla presenza/assenza degli altri. In altre parole, se si rimuove dai dati l'effetto dell'appartenenza alle diverse classi latenti, ciò che rimane è pura casualità (nel senso di completa indipendenza tra le misure). Secondo Lazarsfeld e Henry (1968) questo criterio conduce a gruppi più naturali ed utili ai fini della classificazione. L'ACL ha inoltre il pregio di sfruttare, qualora esista, l'ordinamento naturale delle modalità di risposta.

Si ricordi che le unità campionarie disponibili di almeno 6 anni d'età con riferimento all'Umbria sono 4.671. Per aumentare la base campionaria impiegata per lo studio e l'analisi del fenomeno della non autosufficienza ed ottenere così risultati più affidabili dal punto di vista della variabilità campionaria (si consideri anche che la disabilità è un fenomeno piuttosto raro nella popolazione non anziana), contenendo al contempo l'eterogeneità delle situazioni considerate, a tali unità sono state aggiunte quelle relative alle regioni confinanti con l'Umbria facenti parte della ripartizione territoriale "Centro", così come definita dall'ISTAT. In particolare sono state considerate le seguenti regioni (e relative numerosità campionarie): Toscana (6.712), Marche (5.241), Lazio (6.728). Così facendo il campione complessivo impiegato è risultato composto da 23.352 individui. L'ACL è stata applicata alle 23.352 unità così individuate e considerando per ciascun soggetto rilevato, nell'ambito dell'indagine Multiscopo, le variabili di cui alla Tavola 1, tranne quelle relative al confinamento. Queste ultime sono state ricondotte ad un'unica variabile con quattro modalità di risposta secondo la gravità: 0 = nessun confinamento; 1 = costretto a casa; 2 = costretto seduto; 3 = costretto a letto. In definitiva l'ACL è stata condotta sulle 14 variabili risultanti dalla Tavola 1 dopo aver ridotto ad una quelle relative al confinamento. Nel seguito la condizione di confinamento verrà sempre riassunta da questa nuova variabile. Inoltre, con la dizione "Centro" si intendono l'Umbria e le tre regioni confinanti di cui si è detto sopra.

L'analisi delle classi latenti è stata condotta utilizzando il software WINMIRA (Von Davier, 2001). I parametri del modello sono stimati con il metodo della massima verosimiglianza nell'ambito di un approccio *model-based*, senza cioè utilizzare i pesi campionari delle unità statistiche e questo per due motivi. Il primo è che ci sono dubbi sulla utilità di usare i pesi campionari nella ACL, a fronte delle complicazioni che comportano (si veda ad esempio Vermut, 2002); il secondo motivo è che un'analisi su dati replicati secondo il peso di riporto all'universo, condotta per valutare l'influenza dei pesi campionari, ha dato risultati del tutto simili a quelli ricavati con i dati non pesati, per quanto attiene alla classificazione delle unità.

Il programma di calcolo, oltre alla stima dei parametri del modello, fornisce le probabilità delle modalità di risposta delle diverse variabili rilevate condizionatamente all'appartenenza a ciascuna delle classi latenti individuate. Inoltre, per ciascuna unità campionaria viene indicata la probabilità di appartenere a ciascuna delle classi individuate e la classe latente a cui corrisponde la probabilità più alta di appartenervi.

Sono stati interpolati diversi MCL, uno per ciascun valore intero tra 2 e 8 del numero delle classi latenti ammesse. Il valore delle statistiche di adattamento del modello ai dati CAIC e BIC è sostanzialmente costante dopo quattro classi latenti. Tramite l'esame delle probabilità stimate di osservare le diverse modalità di risposta condizionatamente all'appartenenza a ciascuna delle 4 classi è possibile caratterizzare il contenuto delle stesse. Alla luce di tale analisi le 4 classi sono state così denominate (con riferimento alle persone che vi appartengono): *Autosufficienti, Disabili lievi, Non autosufficienti iniziali e Non autosufficienti*. La Tavola 4 riporta le suddette probabilità condizionate sulla base delle quali è stata attribuita la denominazione alla classe.

Come si può osservare, nella classe *Autosufficienti*, per tutte le variabili si assegna una probabilità pari ad 1 o quasi alla modalità più bassa (contraddistinta dallo 0). Si tratta indubitabilmente di persone che non lamentano alcuna limitazione nelle sfere di autonomia considerate. Nella classe dei *Disabili lievi* troviamo invece elevate probabilità di avere persone che lamentano difficoltà nel camminare, nello scendere o salire le scale e nel chinarsi a terra; altre difficoltà, come fare il bagno o la doccia, sono presenti più sporadicamente. Nel caso dei *Non autosufficienti iniziali*, quasi tutti lamentano qualche o molta difficoltà nello scendere o salire le scale o nel chinarsi a terra, nel fare il bagno o la doccia e nel mettersi e alzarsi dal letto. Infine, nell'ultima classe, quella dei *Non autosufficienti*, troviamo le probabilità più alte in corrispondenza dei livelli massimi di difficoltà, ad eccezione delle variabili "Masticare", "Parlare", "Sentire" e "Vedere". Vale la pena osservare che, per verificare la bontà della soluzione trovata, l'ACL qui descritta è stata applicata anche alle altre ripartizioni territoriali italiane ed ha prodotto risultati del tutto analoghi per quanto riguarda numero e contenuto delle classi latenti e in particolare la struttura delle probabilità condizionate.

Una volta individuate e caratterizzate le classi latenti, sulla base delle probabilità di appartenere alle diverse classi, ciascuna unità campionaria è stata attribuita alla classe latente per la quale ha la massima probabilità di farne parte. Mediante i pesi campionari è quindi ora possibile stimare il numero delle persone assegnate alle quattro classi latenti individuate. La Tavola 5 riporta le stime così ottenute per la sola regione Umbria. In particolare, i *Non autosufficienti* sono stimati al 1.1.2000 in 13.755 persone e costituiscono l'1,7% della popolazione di riferimento. Vale la pena osservare che mentre la percentuale di popolazione anziana su quella di almeno 6 anni è del 23,0%, tale percentuale scende al 17,7% nel gruppo degli *Autosufficienti* ma sale al 73,4% tra i *Disabili lievi*; al 83,2% tra i *Non autosufficienti iniziali* e al 80,1% tra i *Non autosufficienti*. È evidente, perciò, come la non autosufficienza sia un fenomeno caratteristico della condizione anziana.

2.3. Validazione delle classi latenti

Al fine di analizzare la validità della classificazione effettuata sono state calcolate le frequenze percentuali di particolari attributi e le medie aritmetiche di alcune variabili quantitative nelle sottopopolazioni di unità assegnate alle diverse classi latenti. I valori ottenuti si riferiscono all'intero campione analizzato costituito dalle unità campionate in Umbria e nelle altre regioni della ripartizione Centro.

Un primo insieme di risultati è riportato nella Tavola 6. Le prime cinque righe (variabili) sono relative alle quattro forme di disabilità individuate dall'ISTAT (Confinamento, Difficoltà nel movimento, Difficoltà nelle funzioni, Difficoltà sensoriali) e alla Disabilità di qualsiasi tipo. Ad esempio, con riferimento a quest'ultima, la percentuale delle persone disabili è del 4,9% nella popolazione totale con 6 anni e più, ma scende allo 0,4% nella classe degli *Autosufficienti* per salire al 100% nella classe dei *Non autosufficienti*. In quest'ultima classe poi, quasi tutti gli individui sono disabili nelle funzioni, mentre nel 77,5% dei casi sono disabili per confinamento.

Questi dati sono confermati anche dalla frequenza delle modalità di risposta "Molto male" e "Male" alla domanda "Come va in generale la sua salute?". "Molto male" è massimamente presente nella classe dei *Non autosufficienti* (36,7%), mentre la risposta "Male" è più presente nella classe dei *Non autosufficienti iniziali*. La percentuale degli individui che hanno risposto "Male" o "Molto male" è in totale del 3,9% nella classe degli *Autosufficienti*, del 30,2% nella classe dei *Disabili lievi*, del 56,1% nella classe dei *Non autosufficienti iniziali* e sale al 79,9% nella classe dei *Non autosufficienti*.

Tavola 4 - Probabilità delle diverse modalità di risposta per ciascun indicatore condizionate alle classi latenti indicate

Classe Autosufficienti

	Modalità di risposta			
Quesito	***0***	***1***	***2***	***3***
Costretto in casa, sulla sedia, a letto	0,999	0,001	0,000	0,000
Distanza più lunga	0,996	0,004	0,000	
Scendere o salire rampa di scale	0,993	0,007	0,000	0,000
Chinarsi a terra	0,994	0,006	0,000	0,000
Mettersi e alzarsi dal letto	1,000	0,000	0,000	
Sedersi e alzarsi dalla sedia	1,000	0,000	0,000	
Vestirsi e spogliarsi	0,999	0,001	0,000	
Fare il bagno o la doccia	0,997	0,002	0,000	
Lavarsi le mani e il viso	1,000	0,000	0,000	
Mangiare e tagliando il cibo	0,999	0,001	0,000	
Masticare	0,994	0,005	0,001	0,000
Parlare	0,997	0,003	0,000	0,000
Sentire	0,982	0,017	0,001	
Vedere	0,995	0,004	0,001	

Classe Disabili lievi

	Modalità di risposta			
Quesito	***0***	***1***	***2***	***3***
Costretto in casa, sulla sedia, a letto	0,986	0,014	0,000	0,000
Distanza più lunga	0,682	0,304	0,014	
Scendere o salire rampa di scale	0,337	0,582	0,076	0,005
Chinarsi a terra	0,405	0,527	0,051	0,018
Mettersi e alzarsi dal letto	0,910	0,089	0,001	
Sedersi e alzarsi dalla sedia	0,967	0,033	0,000	
Vestirsi e spogliarsi	0,908	0,090	0,002	
Fare il bagno o la doccia	0,744	0,218	0,039	
Lavarsi le mani e il viso	0,988	0,012	0,001	
Mangiare e tagliando il cibo	0,941	0,058	0,001	
Masticare	0,805	0,171	0,022	0,002
Parlare	0,953	0,034	0,008	0,006
Sentire	0,821	0,150	0,028	
Vedere	0,926	0,066	0,008	

Classe *Non Autosufficienti iniziali*

	Modalità di risposta			
Quesito	***0***	***1***	***2***	***3***
Costretto in casa, sulla sedia, a letto	0,822	0,165	0,013	0,000
Distanza più lunga	0,269	0,501	0,229	
Scendere o salire rampa di scale	0,065	0,347	0,416	0,172
Chinarsi a terra	0,063	0,359	0,436	0,143
Mettersi e alzarsi dal letto	0,235	0,718	0,047	
Sedersi e alzarsi dalla sedia	0,436	0,550	0,014	
Vestirsi e spogliarsi	0,372	0,536	0,092	
Fare il bagno o la doccia	0,217	0,437	0,346	
Lavarsi le mani e il viso	0,829	0,146	0,025	
Mangiare e tagliando il cibo	0,771	0,189	0,039	
Masticare	0,588	0,312	0,091	0,009
Parlare	0,852	0,098	0,036	0,014
Sentire	0,719	0,223	0,057	
Vedere	0,800	0,174	0,026	

Tavola 4 - segue

Classe *Non Autosufficienti*				
	Modalità di risposta			
Quesito	***0***	***1***	***2***	***3***
Costretto in casa, sulla sedia, a letto	0,215	0,246	0,270	0,269
Distanza più lunga	0,006	0,081	0,914	
Scendere o salire rampa di scale	0,000	0,005	0,086	0,909
Chinarsi a terra	0,000	0,017	0,071	0,912
Mettersi e alzarsi dal letto	0,005	0,109	0,885	
Sedersi e alzarsi dalla sedia	0,037	0,274	0,689	
Vestirsi e spogliarsi	0,018	0,100	0,882	
Fare il bagno o la doccia	0,018	0,012	0,969	
Lavarsi le mani e il viso	0,150	0,367	0,483	
Mangiare e tagliando il cibo	0,223	0,325	0,451	
Masticare	0,328	0,307	0,298	0,067
Parlare	0,461	0,262	0,192	0,084
Sentire	0,615	0,218	0,167	
Vedere	0,590	0,300	0,110	

Tavola 5 - Consistenza numerica per la regione Umbria degli individui con 6 anni e più appartenenti alle quattro classi latenti individuate, alla data della rilevazione Multiscopo (1.1.2000)

Livello di autosufficienza	Numero di unità	% sul totale
Autosufficienti	706.598	89,58
Disabili lievi	44.140	5,60
Non autosufficienti iniziali	24.333	3,08
Non autosufficienti	13.755	1,74
Totale (individui con 6 anni e più)	788.826	100,00

Fonte: nostra elaborazione su dati ISTAT – Multiscopo 1999-2000

Tavola 6 - Frequenza percentuale di alcune forme di disabilità e di alcuni giudizi sul proprio stato di salute e valore medio di alcuni indici di salute percepita nella quattro classi latenti individuate

Variabili	Livelli di autosufficienza				Sul totale popolaz. con 6 anni e più
	Autosuff.	Disabili lievi	Non autosuff. iniziali	Non autosuff.	
Confinamento (%)	0,1	1,3	17,9	77,5	1,9
Difficoltà nel movimento (%)	0,0	4,7	37,8	44,4	2,2
Difficoltà nelle funzioni (%)	0,0	4,3	39,2	90,0	2,9
Difficoltà sensoriali (%)	0,2	4,5	8,7	25,5	1,2
Disabilità di qualsiasi tipo (%)	0,4	14,0	65,3	100,0	4,9
Come va in generale la sua salute?					
male (%)	3,4	26,2	45,5	43,2	6,8
molto male (%)	0,5	4,2	10,6	36,7	1,6
Indice di Benessere fisico (PCS12) (val. medio)	52,2	38,8	30,9	26,4	50,2
Indice di Benessere psichico (MCS12) (val. medio)	50,8	45,0	40,2	35,9	49,8

Fonte: nostra elaborazione su dati ISTAT – Multiscopo 1999-2000

Le successive 2 righe (variabili) della Tavola 6 contengono i valori medi dei due indici PCS12 *(Psysical Component Summary)* ed MCS12 (*Mental Component Summary*) calcolati nella Multiscopo. Si tratta di due indici ricavati da un questionario messo a punto dall'OMS per valutare la percezione da parte di una persona delle sue condizioni psico-fisiche. La versione utilizzata è chiamata SF12 (*Short Form Health Survey*), perché basata su 12 quesiti riguardanti la sfera fisica e psichica degli individui. Tramite un opportuno algoritmo le risposte ai quesiti vengono tradotte nei due indici PCS12 ed MCS12 (ISTAT 2002; p. 136). Maggiore è il valore dell'indice, migliore è la condizione fisica o psichica percepita. Valori intorno a 50 corrispondono a buone condizioni di salute, mentre valori intorno a 20 denotano pessime condizioni di salute. Nella Tavola 6 il valore medio del PCS12 è pari a 52,2 nella classe degli *Autosufficienti* e scende sino a 26,2 in quella dei *Non autosufficienti*. I corrispondenti valori dell'MCS12 sono 50,8 e 35,9.

Nella Tavola 7 è riportata la frequenza di modalità di risposta relative ad aspetti delle condizioni di salute degli individui nelle quattro classi considerate. Le malattie neurologiche sono praticamente assenti nelle prime due classi ma sono presenti in più di un quarto dei *Non autosufficienti*. A conferma di questo dato, la presenza di forme di invalidità da insufficienza mentale è al 23,2% tra i *Non autosufficienti*, mentre è molto sporadica nelle altre classi. Un trend analogo lo si osserva per l'invalidità per cecità e sordità. Da osservare poi l'alta percentuale dell'invalidità motoria nella classe dei *Non Autosufficienti*, al 65%.

Tavola 7 - Frequenza percentuale di alcuni attributi e valore medio del numero di cronicità nelle quattro classi latenti individuate

Variabili	Livelli di autosufficienza				Sul totale popolaz. Con 6 anni e più
	Autosuff.	Disabili Lievi	Non autosuff. Iniziali	Non autosuff.	
Presenza di Riduzione autonomia (%)	2,7	28,3	58,7	87,7	7,3
Presenza di Invalidità da cecità (%)	0,1	1,8	5,6	14,6	0,6
Presenza di Invalidità da sordità (%)	0,9	7,8	9,6	15,5	1,9
Presenza di Invalidità da insufficienza mentale (%)	0,2	1,9	7,1	23,2	0,9
Presenza di Invalidità motoria (%)	0,6	8,0	23,0	65,0	2,8
Presenza di Parkinson, Alzheimer, epilessia, perdita della memoria, ecc. (%)	0,6	2,8	8,4	28,7	1,3
Presenza di almeno una cronicità (%)	51,9	86,9	91,6	94,7	56,1
Numero delle cronicità (media)	1,3	3,8	4,7	4,9	1,6

Fonte: nostra elaborazione su dati ISTAT – Multiscopo 1999-2000

La Tavola 8 riporta, infine, alcuni indici relativi a variabili anagrafiche e sociali. L'età media nelle diverse classi sale dal valore di 41,3 nella prima classe a quello di 75,8 anni dell'ultima classe. Evidenti sono pure le differenze di genere. Mentre i maschi sono il 48,2% della popolazione con 6 anni e più, tale percentuale scende drasticamente nelle classi dei *Non autosufficienti iniziali* e dei *Non autosufficienti*, a significare che la non autosufficienza è con maggiore prevalenza donna. Si osservi poi l'alta percentuale nell'ultima classe rispetto alle altre della presenza nella famiglia di collaboratori domestici e di persone addette all'assistenza di anziani. Gli individui, invece, che hanno dichiarato di essere usciti da casa per frequentare luoghi di intrattenimento o di culto negli ultimi tre mesi sono il 67,0% tra gli *Autosufficienti*, ma tale percentuale scende drasticamente al 7,7% nella classe dei *Non autosufficienti*. Per quanto riguarda la percezione dell'adeguatezza delle proprie condizioni economiche, se si eccettua la classe degli *Autosufficienti*, nelle altre tre classi la percentuale di coloro che danno un giudizio di inadeguatezza è circa la stessa e sensibilmente superiore. Un dato molto interessante è la frequenza percentuale degli individui che hanno fruito di servizi di assistenza domiciliare sanitaria negli ultimi tre mesi. Come si vede il servizio è principalmente fruito da individui appartenenti alla classe dei *Non autosufficienti* (15,4%).

Tavola 8 - Età media e frequenza percentuale di alcuni attributi nelle quattro classi latenti individuate

Variabili	Livelli di autosufficienza				Sul totale popolaz. con 6 anni e più
	Autosuff.	Disabili lievi	Non autosuff. iniziali	Non autosuff.	
Età (media)	41,3	65,4	72,4	75,8	42,2
Maschio (%)	49,5	43,8	29,6	29,3	48,2
Collaboratore domestico (%)	8,1	10,8	16,6	18,2	8,7
Assistenza anziano (%)	0,5	1,2	8,2	27,5	1,3
Almeno una uscita per tempo libero (%)	67,0	27,6	17,2	7,7	62,0
Risorse economiche scarse o assolutamente scarse (%)	25,1	41,1	46,0	41,7	27,0
Assistenza domiciliare sanitaria (%)	0,1	1,1	5,6	15,4	0,6

Fonte: nostra elaborazione su dati ISTAT – Multiscopo 1999-2000

3. Stime per piccole aree

In questo paragrafo si propone un metodo di stima del numero delle persone assegnate alle quattro classi latenti nelle Aziende Sanitarie Locali (ASL) dell'Umbria al 1.1.2005. I dati campionari umbri non sono però sufficienti per una stima diretta della consistenza delle diverse classi a livello subregionale in quanto gli errori relativi di campionamento sarebbero particolarmente grandi, specialmente quando interessa anche il dettaglio per classe d'età. Diventa perciò necessario ricorrere a metodi di stima per piccole aree utilizzando informazioni extra campionarie disponibili per esse. In Montanari (2005), per ciascuna ASL è stata utilizzata una stima sintetica post-stratificata per sesso e classe d'età basata sui tassi di non autosufficienza specifici stimati a livello di intera regione.

Nella letteratura statistica corrente sono stati proposti modelli più sofisticati in grado di tenere conto della natura categorica e ordinata della variabile risposta qual è quella della variabile indicatrice della classe latente di appartenenza (si veda ad esempio Rao, 2003). Questa può essere modellata in relazione a variabili di cui sia nota la distribuzione sul territorio e a tale scopo sono particolarmente adatti i modelli di tipo logistico multinomiale (McCullagh e Nelder, 1990) e ordinale (*proportional-odds*, Agresti, 2002). Il primo permette di modellare simultaneamente le probabilità di appartenere alle quattro classi latenti sfruttando la relazione che esiste fra le frequenze relative dei diversi livelli della variabile risposta e le variabili indipendenti. Il secondo è un modello più parsimonioso che è stato proposto in letteratura per gestire variabili risposta categoriche le cui modalità presentino un ordinamento naturale (nel nostro caso crescente non autosufficienza). Le variabili che possono essere inserite nel modello sono quelle per cui si dispone della distribuzione sul territorio. In base ai dati disponibili è stato possibile modellare le probabilità di appartenere alle quattro classi latenti in funzione delle sole variabili età e sesso, ma nulla impedisce di utilizzare ulteriori covariate utili allo scopo se disponibili.

Le variabili età e sesso sono state dapprima inserite in un modello di tipo ordinale. Il modello è stato interpolato sul campione allargato di 23.352 individui con 6 anni d'età e più. Nel nostro caso, questo tipo di modello ipotizza che l'effetto delle covariate sia il medesimo per ciascuna classe. Tale ipotesi, piuttosto forte, è stata testata ed è risultata non confermata dai dati impiegati. Si è quindi deciso di adottare un modello logistico multinomiale. In particolare la variabile sesso entra nel modello come variabile qualitativa categorica con 2 modalità, mentre la variabile età è stata inserita fra le covariate come variabile quantitativa ipotizzando per essa una forma polinomiale di terzo grado. In aggiunta, per tener conto della variabilità esistente fra le ASL non catturata dalle variabili esplicative, è stata inserita una variabile categorica indicatrice della ASL di appartenenza. Tale variabile presenta 5 livelli, uno per ciascuna della 4 ASL umbre, più uno in corrispondenza dell'insieme delle altre regioni confinanti con l'Umbria. Nella letteratura sulla stima per piccole aree questo effetto viene solitamente modellato inserendo un effetto casuale di area e considerando, di conseguenza, tali parametri del modello come variabili casuali (si veda ad esempio, Rao, 2003). In presenza di un numero così esiguo di aree, tuttavia, tale approccio non è adeguato. In definitiva, il modello impiegato può essere scritto come segue:

$$\log\frac{p_j}{p_1} = \eta_j = a_{0j} + a_{1j}sesso_F + a_{2j}asl_1 + a_{3j}asl_2 + a_{4j}asl_3 + a_{5j}asl_4 + a_{6j}eta + a_{7j}eta^2 + a_{8j}eta^3$$

dove p_j é la probabilità di appartenere alla classe j-esima, per j = 2, 3, 4, e p_1 è quella di appartenere alla prima classe degli *Autosufficienti* (classe di riferimento); la variabile esplicativa $sesso_F$ è la variabile indicatrice di sesso femminile mentre asl_l, l = 1, …, 4, è la variabile indicatrice dell'appartenenza all'ASL l-esima (categoria di riferimento sono le regioni confinanti con l'Umbria); infine, la variabile eta^r è l'età della persona elevata a potenza r con r =1, 2, 3. Stante la parametrizzazione adottata, il termine a_{0j} è il termine noto del polinomio di terzo grado nella variabile età che modella il logaritmo del rapporto tra p_j e p_1 (*odds*) per un individuo maschio residente in un comune delle regioni confinanti. Si osservi che il modello stima coefficienti a_i, per i = 0, …, 8, diversi per ciascuna classe latente.

La significatività delle variabili impiegate viene verificata mediante il test del rapporto di verosimiglianza, la cui distribuzione è di tipo chi-quadrato con un numero di gradi di libertà pari al numero dei parametri stimati per ciascuna variabile. Tutte le variabili impiegate sono risultate significative. La Tavola 9 riporta i valori stimati dei parametri a_{ij} per $i = 0, \ldots, 8$ e $j = 2, 3, 4$ delle variabili impiegate nel modello. L'effetto sulla scala logistica di appartenere alle categorie riportate nella tavola è in aumento o diminuzione delle probabilità di appartenere alla classe latente corrispondente rispetto a quella degli autosufficienti, a seconda del segno del parametro stimato. Si ricordi che i valori dei parametri come da Tavola 9 sono nell'ordine del predittore lineare η_j e, quindi, rappresentano il logaritmo degli *odds*. Questo significa che, in relazione ad una classe latente, se un parametro assume un valore positivo, la presenza di quella categoria rende più probabile appartenere a quella classe rispetto alla classe degli autosufficienti con riferimento all'individuo base. Modelli più complessi con interazioni fra le variabili non sono risultati significativi.

Tavola 9 - Parametri stimati dal modello logistico multinomiale per piccole aree

Variabile	Categoria (dove applicabile)	Classe latente	Parametro
Intercetta*		Disabili lievi	-0,114
		Non autosufficienti iniziali	-3,938
		Non autosufficienti	-6,249
Sesso	Femmina	Disabili lievi	0,173
		Non autosufficienti iniziali	0,596
		Non autosufficienti	0,456
ASL	ASL1	Disabili lievi	-0,256
		Non autosufficienti iniziali	0,227
		Non autosufficienti	-0,554
	ASL2	Disabili lievi	-0,654
		Non autosufficienti iniziali	-0,361
		Non autosufficienti	-0,419
	ASL3	Disabili lievi	0,068
		Non autosufficienti iniziali	0,043
		Non autosufficienti	-0,324
	ASL4	Disabili lievi	-0,261
		Non autosufficienti iniziali	-0,267
		Non autosufficienti	0,004
Età	-	Disabili lievi	-0,419
		Non autosufficienti iniziali	-0,194
		Non autosufficienti	-0,030
Età ^2	-	Disabili lievi	0,009614
		Non autosufficienti iniziali	0,004350
		Non autosufficienti	0,000259
Età ^3	-	Disabili lievi	-0,000055
		Non autosufficienti iniziali	-0,000019
		Non autosufficienti	0,000010

*L'individuo base è maschio e residente in un comune delle regioni Lazio, Marche, Toscana. La classe latente di riferimento è quella degli *Autosufficienti*

Sulla base del modello interpolato è possibile costruire la Figura 2 che mostra l'andamento delle probabilità stimate di appartenere ad ognuna delle quattro classi in funzione dell'età per un individuo maschio e residente in un comune delle regioni confinanti con l'Umbria. Disponendo delle probabilità di appartenere a ciascuna classe latente a seconda dell'età, del sesso e dell'ASL di residenza è ora possibile stimare il numero atteso delle persone appartenenti alle diverse classi latenti in una data sottopopolazione. A tale scopo, attraverso i parametri del modello viene determinata dapprima la probabilità di appartenere a ciascuna delle quattro classi latenti per un individuo appartenente a ciascuna sottoclasse individuata dall'incrocio delle variabili ASL, età e sesso. Moltiplicando poi tali probabilità per l'ammontare di popolazione di ciascuna sottoclasse nella sottopopolazione considerata e sommando sulle sottoclassi si ottiene il numero atteso totale di persone da classificarsi nelle quattro classi latenti nel territorio o sottopopolazione di interesse. In questo modo sono state ottenute le stime a livello di ASL al 1° gennaio 2005 riportate nella Tavola 10. Tali stime sono ovviamente calcolabili per date diverse purché si disponga della distribuzione per sesso ed età della popolazione di ciascuna ASL.

Figura 2. - Probabilità (stimata) di appartenere a ciascuna classe latente rispetto all'età per l'individuo base (maschio, residente in un comune delle regioni confinanti).

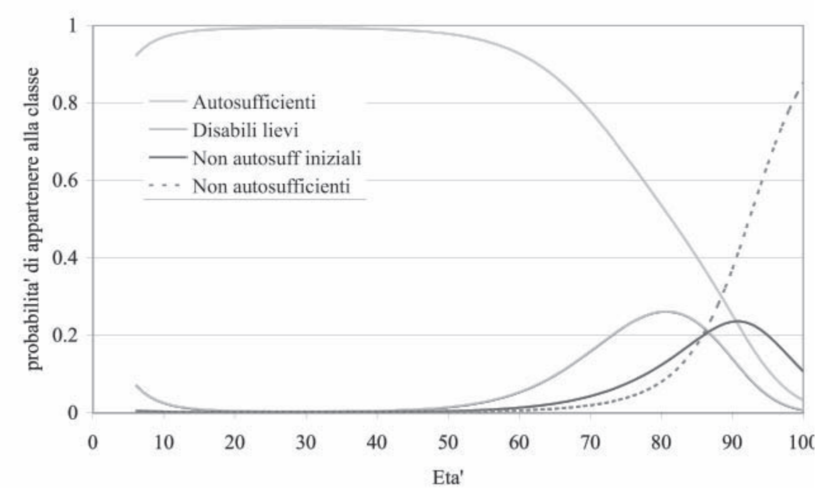


Tavola 10 - Frequenza assoluta e percentuale della popolazione umbra con 6 anni e più, Autosufficiente, Disabile lieve, Non autosufficiente iniziale, Non autosufficiente per ASL al 1.1.2005 (stime).

ASL	Totale	Autosuff.	Disabili lievi	Non autosuff. iniziali	Non autosuff.	Autosuff. %	Disabili Lievi %	Non autosuff. iniziali %	Non autosuff. %	Totale
1	124.274	108.984	7.255	6.369	1.666	87,70	5,84	5,13	1,34	100
2	329.001	298.121	14.335	10.722	5.823	90,61	4,36	3,26	1,77	100
3	148.701	127.639	12.012	6.484	2.567	85,84	8,08	4,36	1,73	100
4	213.884	187.318	13.738	7.454	5.375	87,58	6,42	3,48	2,51	100
Umbria	815.860	722.061	47.340	31.028	15.431	88,50	5,80	3,80	1,89	100

Fonte: nostra elaborazione su dati ISTAT – Popolazione residente al 1° gennaio 2005

In modo analogo è stata ottenuta la Tavola 11 dove vengono riportate le distribuzioni di frequenza assoluta e percentuale secondo l'età all'interno delle varie classi latenti. Come si può osservare, le persone di 65 e più anni sono il 24,5% della popolazione con almeno 6 anni di età ma tale percentuale sale al 77,1% tra i disabili lievi; al 86,4% tra i non autosufficienti iniziali e al 90,2% tra i non autosufficienti. Si osservi che passando dai dati di popolazione utilizzati da ISTAT per il calcolo dei pesi di riporto all'universo a quelli della Popolazione residente al 1° gennaio 2005, il numero stimato dei non autosufficienti sale da 13.755 a 15.431 e ciò è anche da imputare al fatto che nel frattempo è stato svolto il Censimento della Popolazione del 2001 con una evidente discontinuità nella serie storica della popolazione presente a causa delle modifiche apportate alla struttura della popolazione per sesso ed età. Ovviamente, la metodologia può essere applicata a domini di studio più piccoli come ad esempio quelli che si ottengono incrociando l'ASL con la classe d'età, oppure i distretti sanitari o i comuni.

Tavola 11 - Totale e distribuzione percentuale della popolazione umbra con 6 anni e più, Autosufficiente, Disabile lieve, Non autosufficiente iniziale, Non autosufficiente per età al 01.01.2005 (stime).

Età	Popolazione 6+		Autosufficienti		Disabili lievi		Non autosuf. in.		Non autosuffic.	
6 – 13	55.949	*6,9*	53.660	*7,4*	1.911	*4,0*	302	*1,0*	76	*0,5*
14 – 24	88.830	*10,9*	87.965	*12,2*	530	*1,1*	226	*0,7*	108	*0,7*
25 – 44	250.944	*30,8*	248.863	*34,5*	1.058	*2,2*	667	*2,1*	356	*2,3*
45 – 64	220.357	*27,0*	209.025	*28,9*	7.335	*15,5*	3.034	*9,8*	964	*6,2*
65 – 74	100.623	*12,3*	77.117	*10,7*	15.071	*31,8*	6.552	*21,1*	1.883	*12,2*
75 e oltre	99.157	*12,2*	45.431	*6,3*	21.435	*45,3*	20.248	*65,3*	12.043	*78,0*
Totale	815.860	*100,0*	722.061	*100,0*	47.340	*100,0*	31.029	*100,0*	15.431	*100,0*

Fonte: nostra elaborazione su dati ISTAT – Popolazione residente al 1° gennaio 2005

4. Conclusioni

Questo studio propone una metodologia di analisi del fenomeno della non autosufficienza e una applicazione al caso dell'Umbria sulla base dei dati ricavati dall'indagine Multiscopo dell'ISTAT 1999-2000 sulle condizioni di salute e il ricorso ai servizi sanitari, che viene eseguita con cadenza quinquennale. Attraverso la metodologia delle classi latenti sono state identificate quattro categorie di persone relativamente alla dimensione dell'autosufficienza: *Autosufficienti*, *Disabili lievi*, *Non autosufficienti iniziali* e *Non autosufficienti*. Queste categorie sono state ulteriormente analizzate per definirne caratteristiche e numerosità.

Allo scopo di stimare il numero delle persone assegnate alle diverse classi di autosufficienza in una sottopopolazione di interesse, come può essere quella delle ASL della regione dell'Umbria, è stato messo a punto, sulla base della situazione al momento della indagine multiscopo 1999-2000, un modello logistico multinomiale che ha consentito di calcolare la probabilità di ricadere nelle diverse classi sulla base dell'età in anni compiuti e di altre caratteristiche della persona e del contesto in cui essa vive. L'applicazione di tale modello alla popolazione residente in ciascuna ASL al 1.1.2005 ha consentito una stima delle persone non autosufficienti presenti a questa data utile ai fini della programmazione socio-sanitaria. La procedura descritta si configura come una metodologia di stima per piccole aree in quanto applicabile a domini territoriali più piccoli come i distretti sanitari o i comuni.

La classificazione delle unità è stata condotta senza tenere conto in modo esplicito del piano di campionamento. È intenzione degli autori proseguire la ricerca studiando in dettaglio il ruolo dei pesi campionari nell'analisi in classi latenti impiegando, ad esempio, le metodologie illustrate in Vermunt (2002). Inoltre, il modello per piccole aree è stato interpolato impiegando la variabile di classificazione nelle quattro classi latenti come se fosse una variabile osservata. È allo studio al riguardo l'impiego di modelli a classi latenti che includano le covariate e gli effetti casuali utili alla stima per le piccole aree già al momento della definizione della classificazione. Inoltre, l'impiego di opportuni modelli di Rasch (1960, 1977) potrebbe permettere, una volta valutata l'autosufficienza su una scala ordinata, di analizzare i fattori socio-sanitari e di contesto maggiormente associati con la stessa.

Una ulteriore direzione di sviluppo del presente lavoro che si intende esplorare è quella della ricerca e dell'utilizzo di dati amministrativi e/o rilevati ad hoc, da una parte, per meglio caratterizzare le persone assegnate alle varie classi e i loro bisogni di cura, anche al fine di poter consentire una programmazione sanitaria più aderente alle reali necessità delle persone e suddividere le risorse attribuite ai vari territori su basi più quantitative, aggirando così le eterogeneità che si riscontrano nella gestione del fenomeno della non autosufficienza a livello di Aziende sanitarie, di province e di regioni. Dall'altra parte per perfezionare la metodologia di stima per piccole aree. Ad esempio si può pensare di utilizzare gli archivi delle indennità di accompagnamento e delle invalidità civili in genere, gli archivi delle prestazioni sociali erogate dai comuni, oppure ancora quelli relativi ai dati di stock e di flusso delle persone istituzionalizzate. Maggiore è il patrimonio informativo di cui si potrà disporre, maggiore saranno le possibilità di analisi e stima dei fenomeni che qui interessano.

Riferimenti bibliografici

Agresti A. (2002), *Categorical data analysis*. John Wiley & Sons, New York.

Branch L. G. e Meyers A. R. (1987), "Assessing physical function in the elderly". *Clinics in Geriatrics Medicine* 3 (1): 29-51.

Deville, J. C. e Sarndal, C. E. (1992) "Calibration estimators in survey sampling", *Journal of the American Statistical Association*, 85, 376–382.

Goodman L. A. (1974), "Exploratory latent structure analysis using both identifiable and unidentifiable models". *Biometrika*, 61, 215-231.

Heinen T. (1996), *Latent class and discrete latent trait models: Similarities and differences*. Thousand Oaks, California, Sage.

ISTAT (2002), Le condizioni di salute della popolazione – Indagine Multiscopo sulle famiglie "Condizioni di salute e ricorso ai servizi sanitari". Anni 1999 – 2000, *Informazioni n. 12*, Roma.

ISTAT (2004), *Studio sulla tematica della "Non autosufficienza",* a cura del Sistema Informativo sulla Disabilità, disponibile all'indirizzo:
http://www.disabilitaincifre.it/descrizioni/altriapprofondimenti.asp

ISTAT (2005a), "Le previsioni regionali della popolazione residente 1.1.2001-1.1.2050". A cura di M. Marsili e M. P. Sorvillo, disponibile all'indirizzo:
http://demo.istat.it/prev/index.html

ISTAT (2005b), "Sistema sanitario e salute della popolazione (4a edizione)". Pagina Web *http://www.istat.it/sanita/sociosan/*

Katz S., Ford A. B., Moskowitz R.W., Jackson B.A. e Jaffe M.W. (1963), "Studies of illness in the aged. The Index of ADL: A standardized measure of biological and psychosocial function". *JAMA* 185(12):914-919.

Lazarsfeld P. F. e Henry N.W. (1968), *Latent structure analysis*. Boston, Houghton Mifflin.

Masters G. N. (1982), "A Rasch model for partial credit scoring". *Psychometrika, 47,* 149-174.

McCullagh P. e Nelder J. A. (1990), *Generalized Linear Models*. Chapman and Hall, London.

McCutcheon A. C. (1987), *Latent class analysis*. Beverly Hills, Sage.

Montanari G. E. (2005), "La stima della non autosufficienza in Umbria sulle base dell'indagine Multiscopo dell'ISTAT", Rapporto di Ricerca.

Montanari G. E., Ranalli M.G., Eusebi P., Cicioni C. e Catanelli M. (2006), Stima della non autosufficienza negli anziani in Umbria, in "Umbria: demografia e stima della non autosufficienza negli anziani". DVSS – Monografia nr. 1, Regione dell'Umbria.

Rao J. N .K. (2003), *Small area estimation*. Wiley.

Rasch G. (1960), *Probabilistic Models for Some Intelligence and Attainment Tests*. The Danish Institute of Educational Research, Copenhagen.

Rasch G. (1977), "On specific objectivity. An attempt at formalizing the request for generality and validity of scientific statements". In Blegvad, M. (ed.). The Danish Yearbook of Philosophy, 58-94. Munksgaard, Copenhagen.

Uebersax J. S. (1993), "Statistical modeling of expert ratings on medical treatment appropriateness". *Journal of the American Statistical Association*, 88, 421-427.

Vermunt J. K. (2002), "Latent class analysis of complex sample survey data: application to dietary data. Comment to the Authors". *Journal of the American Statistical Association*, 97, 736-737.

Vermunt J. K. e Magidson J. (2002), "Latent class cluster analysis". Chapter B1 in Hagenaars and McCutcheon, eds., *Advances in latent class models*. Cambridge, Cambridge University Press.

Von Davier M. (2001), WINMIRA 2001, Pagina Web *http://winmira.von-davier.de*

Wiener J. M., Hanley R.J., Clark R. e van Nostrand J. (1990), "Measuring the activity of daily living: comparison across national surveys". *Journal of Gerontology: Social Sciences* 45 (6): S229-S237

World Health Organisation (WHO) (1980), *International Classification of Impairments, Disabilities and Handicaps*. World Health Organisation: Geneva.

Norme redazionali

La Rivista di Statistica Ufficiale pubblica contributi originali nella sezione “Temi trattati” ed eventuali discussioni a largo spettro nella sezione “Interventi”. Possono essere pubblicati articoli oggetto di comunicazioni a convegni, riportandone il riferimento specifico. Gli articoli devono essere fatti pervenire al Comitato di redazione delle pubblicazioni scientifiche Istat corredati, a parte, da una nota informativa dell’Autore contenente: appartenenza ad istituzioni, attività prevalente, qualifica, indirizzo, casella di posta elettronica, recapito telefonico e l’autorizzazione alla pubblicazione firmata dagli Autori. Ogni articolo prima della pubblicazione dovrà ricevere il parere favorevole di un referente scelto tra gli esperti dei diversi temi affrontati. Gli originali, anche se non pubblicati, non si restituiscono.

Per l’impaginazione dei lavori gli autori sono tenuti a conformarsi rigorosamente agli standard editoriali fissati dal Comitato di redazione e contenuti nel file Template.doc disponibile on line o su richiesta. In base a tali standard la lunghezza dei contributi originali per entrambe le sezioni dovrà essere limitata entro le 30-35 pagine.

Tutti i lavori devono essere corredati di un sommario nella lingua in cui sono redatti (non più di 12 righe); quelli in italiano dovranno prevedere anche un *Abstract* in inglese. La bibliografia, in ordine alfabetico per autore, deve essere riportata in elenco a parte alla fine dell’articolo. Quando nel testo si fa riferimento ad una pubblicazione citata nell’elenco, si metta in parentesi tonda il nome dell’autore, l’anno di pubblicazione ed eventualmente la pagina citata. Ad esempio (Bianchi, 1987, Rossi, 1988, p. 55). Quando l’autore compare più volte nello stesso anno l’ordine verrà dato dall’aggiunta di una lettera minuscola accanto all’anno di pubblicazione. Ad esempio (Bianchi, 1987a, 1987b).

Nella bibliografia le citazioni di libri e articoli vanno indicate nel seguente modo. Per i libri: cognome dell’autore seguito dall’iniziale in maiuscolo del nome, il titolo in corsivo dell’opera, l’editore, il luogo di edizione e l’anno di pubblicazione. Per gli articoli: dopo l’indicazione dell’autore si riporta il titolo tra virgolette, il titolo completo in corsivo della rivista, il numero del fascicolo e l’anno di pubblicazione. Nei riferimenti bibliografici non si devono usare abbreviazioni.

Nel testo dovrà essere di norma utilizzato il corsivo per le parole in lingua straniera e il corsivo o grassetto per quei termini o locuzioni che si vogliano porre in particolare evidenza (non vanno adoperati, per tali scopi, il maiuscolo, la sottolineatura o altro).

Gli articoli pubblicati impegnano esclusivamente gli Autori, le opinioni espresse non implicano alcuna responsabilità da parte dell’Istat.

Per contattare il Comitato di redazione delle pubblicazioni scientifiche Istat e per inviare lavori: rivista@istat.it. Oppure scrivere a:

Segreteria del Comitato di redazione della Rivista di Statistica Ufficiale
All’attenzione di Gilda Sonetti
Via Cesare Balbo, 16
00184 Roma

La Rivista di Statistica Ufficiale accoglie lavori che hanno come oggetto la misurazione e la comprensione dei fenomeni sociali, demografici, economici ed ambientali, la costruzione di sistemi informativi e di indicatori come supporto per le decisioni pubbliche e private, nonché le questioni di natura metodologica, tecnologica e istituzionale connesse ai processi di produzione delle informazioni statistiche e rilevanti ai fini del perseguimento dei fini della statistica ufficiale.

La Rivista di Statistica Ufficiale si propone di promuovere la collaborazione tra il mondo della ricerca scientifica, gli utilizzatori dell'informazione statistica e la statistica ufficiale, al fine di migliorare la qualità e l'analisi dei dati.

La pubblicazione nasce nel 1992 come collana di monografie "Quaderni di Ricerca ISTAT". Nel 1999 la collana viene affidata ad un editore esterno e diviene quadrimestrale con la denominazione "Quaderni di Ricerca - Rivista di Statistica Ufficiale". L'attuale denominazione, "Rivista di Statistica Ufficiale", viene assunta a partire dal n. 1/2006 e l'Istat torna ad essere editore in proprio della pubblicazione.

1B012011001000000

ISSN 1828-1982

€ 10,00