rivista di statistica ufficiale

n. 1
2011

Temi trattati

The dynamics of inflation components and their comparability among countries: the case of the HICP
Carlo De Gregorio

Methodological aspects and empirical evidence in the use of administrative sources to estimate price dynamics in external trade
Paola Anitori, Carlo De Gregorio

Un indice sintetico non compensativo per la misura della dotazione infrastrutturale: un'applicazione in ambito sanitario
Matteo Mazziotta, Adriano Pareto

Istat
Istituto Nazionale
di Statistica

rivista di statistica ufficiale

n. 1
2011

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. 1/2011

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 Dicembre 2011
presso il Centro stampa dell'Istat
Via Tuscolana 1788 - Roma
Copie 300

The dynamics of inflation components and their comparability among countries: the case of the HICP[1]

Carlo De Gregorio[2]

Abstract

The paper presents a method to classify the dynamic patterns of the whole set of sub-indices that make up the harmonised index of consumer prices (HICP) and to assess the comparability of the national series referred to a same elementary aggregate of the target consumption expenditure. We consider more than 1.200 series in the period 2004-2008, referred to nearly 100 elementary aggregates and 14 EU countries. Based on the values assumed by a set of variability indicators compiled for each series, we explore the outcomes of a classification of sub.-indices obtained through the sequential use of principal components and hierarchical clustering. From a measurement of the degree of heterogeneity of the series within each elementary aggregate, we derive a list of the most problematic aggregates which possibly deserve further methodological harmonisation in order to provide comparable estimates. Three distinct case studies are more thoroughly analysed and discussed: air transport, package holidays and electricity.

Keywords: Consumer price index, European Union, Harmonisation, Industrial policies, Cluster analysis, Principal components, Variability

1. Introduction

Notwithstanding its aggregative structure, the consumer price index (CPI) has been traditionally seen in the literature as a typical macroeconomic indicator to be used mainly as a tool to target monetary policies, while much less importance has been given to its potential uses to support industrial, competition or consumer protection policy purposes. These uses seem nevertheless to have been gaining ground in the last years at least in the European Union (EC 2009; EUROSTAT 2009), also due to the presence of a harmonised CPI (the HICP) produced by 30 countries according to a common legal basis. They seem to pave the way for further extensions of the scope of CPIs well beyond the needs of monetary policy.

---

[1] This paper is dedicated to the memory of Carmina Munzi, colleague and friend at Price statistics in ISTAT, and to the rest of the colleagues in charge from 2004 to 2007 of centralised price collection, with whom I shared fundamental experiences and training on this field (Patrizia Caredda, Stefania Fatello, Rosanna Lo Conte, Maurizio Massaroni, Stefano Mosca, Francesca Rossetti, Paola Zavagnini). The empirical core of the paper has been prepared during a period of secondment at the Price statistics unit of EUROSTAT, where it has been used to address the activity of a Task Force on HICP Sampling: a preliminary version has been discussed at the Workshop on Exploratory Data analysis and Visualisation (Vienna, 27-28 May 2010) organised by EUROSTAT and Statistik Austria. I also wish to thank Alexandre Makaronidis, Keith Hayes, Emilio Di Meglio, the colleagues of the Methodologist Network of EUROSTAT and the anonymous referee for their comments and suggestions. The author is anyway the sole responsible for the views expressed in the paper which are not necessarily those of EUROSTAT nor of ISTAT.

[2] Senior Researcher (Istat), e-mail: cadegreg@istat.it.

As any other CPI, the HICP is obtained in each country from the weighted average of several hundreds of sub-indices, each referred to an elementary aggregate derived from an exhaustive partition of the target consumer expenditure. Two main points deriving from this structure will be stressed in this paper. The first derives from the acknowledgement that CPIs derive from the aggregation of possibly largely diversified price dynamics: flatly or smoothly time-linear, erratic or seasonal fluctuations, sharp rises or declines, and so forth. Is it possible to measure this heterogeneity and to derive a classification of the behaviours of the whole set of sub-indices? Secondly, it may be reasonable to expect that the sub-indices referred to a same elementary aggregate behave according to similar rules across EU or at least that they can be compared in order to give an economic explanation to different behaviours of a same market in different countries. In other words, are the national series referred for instance to cars really comparable? Aren't such comparisons affected by heterogeneities in the methods applied to estimate them? Do the sub-indices of the HICP represent a useful tool for comparative market analysis and policy purposes?

The paragraphs that follow propose a method to classify the behaviour of sub-indices and to evaluate their comparability across countries, with some suggestions concerning the possible improvements to make it viable the use of sub-indices as a policy tool. For a given elementary aggregate, the possible lack of methodological harmony can have manifold causes which cannot be properly detected without a detailed look into the microdata and into the sampling designs adopted by the National Statistical Institutes (NSIs): but microdata are rarely available, and also the detailed and systematic descriptions of survey designs are quite rare.[3] Given this premise, the approach which is proposed here is based on the examination of the actual behaviour of more than 1.200 HICP series in the period 2004-2008, broken down by country and by about one hundred exhaustive consumption purposes.[4] The series are classified on the basis of the values assumed by a set of indicators which are analysed by means of the sequential use of principal components and hierarchical clustering.[5]

In the second paragraph some background of this discussion is provided. Paragraph 3 provides a description of the HICP database adopted in the analysis and of the list of the indicators on the behaviour of the series. Paragraph 4 describes the results of principal components analysis on this database and a classification of behavioural patterns. Paragraph 5 presents a ranking of the elementary aggregates according to their cross country heterogeneity. Three case studies concerning the markets for air transports, package holidays and electricity are finally discussed, with the further purpose of evaluating the information content of national HICPs and their adequacy for policy support.

---

3 See as an example De Gregorio et al. (2008) on a subset of the CPI surveys currently run by ISTAT.

4 They correspond to the classes (four digits) of the COICOP classification (EUROSTAT 2001, 281-318), that is to the highest level of detail with which data are publicly available for download from the website of EUROSTAT: these data are thus available for analysis and comparisons to any user. This is also the same detail with which Member states deliver their monthly data to EUROSTAT, according to an agreed calendar.

5 The approach used in this paper has been adopted in Anitori et al. (2004) and it has been applied to price statistics in Brunetti et al. (2004) and in ISTAT (2005, p.42-44).

2. Comparability and variability issues concerning CPI sub-indices

With the exception of a pioneering work by Dalèn (1998), at the dawn of the HICP, systematic and detailed longitudinal comparative analyses of CPI sub-indices are not yet available in the literature, neither with reference to the HICP nor to any other group of national CPIs. With different objectives, the cross section variability of CPI components has been recently examined with at least two distinct purposes. On one side the literature on the "relative price variability" – see for recent overviews Fielding et al. (2001), Nautz et al. (2006), Caraballo et al. (2006) – has investigated the nature of the relationships between the changes occurring in relative prices and the overall inflationary processes, and their implications on the economy depending on the type of inflationary regime. On a different perspective – mainly originated from the central banks' need to target inflation, see for example Wynne (1999) and Roger (1998) – the literature on "core inflation" dealt with the task of finding out a consistent theoretical definition and a viable measurement of the inflationary process in the medium and longer term: practical solutions have focussed mainly on the classification of sub-indices according to their variability in order to freeze the effect of the most volatile ones.

Both these approaches focussed on a single country and on typically macroeconomic purposes, mainly tied to monetary policy. The Harmonised Index of Consumer Prices (HICP) – compiled in 30 European countries[6] within a common legal basis (EUROSTAT 2001) – offers instead the opportunity to move the focus towards the analysis of the behaviour of national sub-indices. It is well known that in general national CPIs are not perfectly comparable. Basic principles might differ, for instance, in the target consumption expenditure, in the classification, in weighting, in the time reference. Such differences make the aggregate indices quite peculiar to each country; the sub-indices on the contrary are far less concerned with general principles and more with national habits having to do with the organisation and the methodology of the numerous distinct surveys which feed CPI estimates. With the binding legal framework of the HICP - whose monthly series started in 1996- many methodological aspects have been harmonised and many heterogeneities have formally disappeared, starting from more general aspects: a common classification; the coverage in terms of consumption expenditure, population, territory; the annual chaining; the weighting. Also some more specific issues have been harmonised, concerning the treatment of tariffs, some coverage aspects in selected sectors (such as insurance, health, education, social protection), the price collection period. Given this huge harmonisation effort, the HICP has no doubt become the price index with the largest background literature, legal basis, documentation and debate.

Not differently from national CPIs, the HICP is primarily adopted as a macroeconomic indicator, for setting and evaluating monetary policies within the euro area: nevertheless, its use for a larger set of policies was already envisaged in its founding documents.[7] The availability of indices with a relatively high degree of harmonisation across the EU, regularly disseminated on the web, broken down by country and by about 100 exhaustive sub-indices, has induced their use to provide evidence for consumer market monitoring, and

[6] That is in the 27 countries participating to the EU plus Norway, Iceland and Switzerland.

[7] See the Council Regulation (EC) No 2494/95 of 23 October 1995 concerning harmonized indices of consumer prices. Official Journal, L 257, October 10, p. 1–4.

namely to support competition and consumer protection policies.[8] Such developments appear quite natural, given the growing importance of these policies within the EU and the availability of a huge amount of comparable information on the dynamics of consumer prices in each member state.

Nevertheless, methodological harmonisation of the HICP is still a work in progress, and several sources of methodological heterogeneity among countries may still persist. Consumer markets are in fact numerous and very much differentiated: they require specific methodological approaches and solutions to the provision of estimates. Lack of harmonisation might concern more general aspects (such as the definition of the population parameter) or the approach to sampling, the treatment of seasonal pricing, the alternative use of list or bargain prices, the management of replacements, quality adjustments, the aggregation formula and may reflect differences in the interpretation of the functioning of consumer markets. As a matter of fact, compared with other branch of statistics, price statistics has received by far less attention on more specific but crucial methodological grounds, such as sample design and variance of the estimates. These issues are only recently acquiring a growing and systematic importance,[9] since scholars on price indices have largely explored other mainstream subjects: from the properties of estimates and indices (Allen 1975; Balk 2008) to quality adjustment and hedonic modelling, the latter having been generally treated as an issue formally distinct from sampling. One of the most promising developments of the HICP has to do with the fact that in the latest years the focus of harmonisation has been moved towards a more methodological ground, running straight into the issue of the estimation of the HICP.[10] Within the process of methodological harmonisation, the whole set of aspects involved in the production of CPI estimates has been treated within a unitary approach, based on the principle that sample design, sample selection and the other typical technicalities of price statistics (elementary aggregation, replacements, quality adjustment etc.) need a common foundation based primarily on proper and coherent definitions of the target population and of the parameter to be estimated.[11]

3. The data base and the indicators

3.1 The construction of the HICP

The HICP is a fixed base chained Laspeyres index, where the fixed base is given by the average of a reference year, namely 2005. In the month *m* of year *t* the aggregate HICP is given by the chaining of an annual link and a fixed base index:

[8] Recently the European Commission has been pushing for a deeper use of CPI microdata to feed analyses both for competition policy and consumer protection. This has brought recently to an in-depth analysis of the food chain (EC 2009) to the use of the microdata of price surveys to estimate average price levels and to compare the efficiency of retailing in a larger set of consumer markets. A convergence in the purpose and scope of CPI and Purchasing Power Parities (PPP) estimates is increasingly considered as an area of further developments in price statistics.

[9] One of the main concerns for the production of price index estimates has to do with the definition of the target population and of the parameter to be estimated: see Ribe (2000) on this subject. As infrequent examples of statistical theory applied to price statistics see also Baskin (1996) and Norberg (2004).

[10] See the 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, L 296, November 15, p. 22-26.

[11] See Ribe (2000) and Dalen (2001).

$$X_{m,t} = X_{12,t-1} H_{m,t} \tag{1}$$

where X is the fixed base HICP and H is the annual link based in the month of December of the previous year (EUROSTAT 2001, p.175-97). The latter is derived as the weighted mean of an exhaustive set of sub-indices:

$$H_{m,t} = \sum_j h_{j;m,t} w_{j,t} \tag{2}$$

where with h we label the links corresponding to the elementary aggregate j. The normalised expenditure weights (w) change every year[12] and the fixed base sub-indices x are given by:

$$x_{m,t} = x_{12,t-1} \, {}^{12,t-1}h_{m,t} \qquad (3)^{13}$$

3.2 The data base

The series used for this analysis are publicly and freely available from EUROSTAT web site. They correspond to the monthly indices based on year 2005 and referred to the 94 classes of the COICOP-HICP classification (the highest available breakdown): 14 countries were considered, 11 from the euro area[14] plus Denmark, Sweden and the UK. The download has regarded the 60 months from January 2004 to December 2008. As a consequence, the data base includes as a whole 1.257 series and about 75.000 monthly indices.

All the selected countries show in the period a relatively homogeneous overall price dynamics.[15] Table 1 resumes the main HICP figures in the target period: the range of the inflation rates across countries has kept constantly below three percentage points, and their coefficient of variation even decreased in 2008 notwithstanding the higher inflation.

From here on, the original series of the fixed base indices are labelled with x_i, where $i \in \mathrm{I}_x = \{1,\ldots,n_x\}$ counts the months from January 2004 to December 2008 (hence $n_x = 60$). The series of the yearly rates have been derived from the original fixed base series as follows:

$$y_i = {x_i}/{x_{i-12}} \tag{4}$$

---

12 In a Laspeyres context, these weights are referred to the consumption expenditure of year t-1.

13 As a consequence – and since the weighting structure changes every year - the fixed base HICP cannot be derived as a weighted average of the fixed base sub-indices but throughout the chain-linking of the weighted averages of the annual links of the sub-indices.

14 Excluding Luxembourg, Cyprus, Malta, Slovenia and Slovakia.

15 This happens in period 2004-2007, characterised by very low inflation rates, but still remains valid with the further inclusion of year 2008, which has been characterised by a sudden and short lived surge in inflation. Separate analyses have been conducted on 2004-2007 data and to the whole period extended to 2008: the results obtained in the classification of the indices and in the measurement of heterogeneity almost identical. Here we propose the entire five years from 2004 to 2008.

where $i \in \mathrm{I}_y = \{13,...,n_x\}$ and $n_y = n_x -12=48$. Accordingly, the series of the monthly rates are derived as:

$$z_i = {x_i}/{x_{i-1}} \qquad (5)$$

where $i \in \mathrm{I}_z = \{2,...,n_x\}$ and $n_z = n_x -1=59$.[16]

Table 1 - HICP annual averages, by year and country (years 2004-2008)

COUNTRY	Average index (base: 2005=100)				Yearly rate of change			
	2004	2006	2007	2008	2005	2006	2007	2008
Austria	97.9	101.7	103.9	107.3	2.1	1.7	2.2	3.2
Belgium	97.5	102.3	104.2	108.9	2.5	2.3	1.8	4.5
Germany	98.1	101.8	104.1	107.0	1.9	1.8	2.3	2.8
Denmark	98.3	101.8	103.5	107.3	1.7	1.8	1.7	3.6
Spain	96.7	103.6	106.5	110.9	3.4	3.6	2.8	4.1
Finland	99.2	101.3	102.9	106.9	0.8	1.3	1.6	3.9
France	98.1	101.9	103.6	106.8	1.9	1.9	1.6	3.2
Greece	96.6	103.3	106.4	110.9	3.5	3.3	3.0	4.2
Ireland	97.9	102.7	105.6	109.0	2.2	2.7	2.9	3.1
Italy	97.8	102.2	104.3	108.0	2.2	2.2	2.0	3.5
The Nederland's	98.5	101.7	103.3	105.5	1.5	1.7	1.6	2.2
Portugal	97.9	103.0	105.5	108.3	2.1	3.0	2.4	2.7
Sweden	99.2	101.5	103.2	106.7	0.8	1.5	1.7	3.3
UK	98.0	102.3	104.7	108.4	2.1	2.3	2.3	3.5
Range	2.61	2.29	3.63	5.36	2.72	2.29	1.41	2.28
Standard deviation	0.71	0.68	1.14	1.50	0.74	0.68	0.49	0.62
Coefficient of variation (%)	0.7	0.7	1.1	1.4	36.2	30.4	22.9	18.0

Source: Elaborations on Eurostat HICP data base

3.3 Some indicators

Relative variability

Three classes of indicators have been calculated for original, monthly and yearly rates series: relative variability indicators, time-linear dynamics indicators and other miscellaneous dynamics indicators. In the formulas that follow, the symbol k (with $k \in \{x, y, z\}$) will be used to represent the three series above.

[16] Yearly and monthly rates of change are conventionally used to summarise price dynamics: they are the most visible and commented outcomes of CPI estimates.

As concerns relative variability,[17] a first set of two indicators (the coefficient of variation and the Gini mean difference) use all the information available for each series:

$$CV_k = \frac{\sigma_k}{\mu_k} = \frac{\sqrt{\frac{\sum_{i \in I_k} (k_i - \mu_k)^2}{n_k}}}{\frac{\sum_{i \in I_k} k_i}{n_k}} \qquad (6)$$

$$RGini_k = \frac{\frac{1}{\binom{n_k}{2}} \sum_{\substack{i<j \\ i,j \in I_k}} \left| k_i - k_j \right|}{\mu_k} \qquad (7)$$

where μ and σ are respectively the mean and the standard deviation of the series, while n is the total number of monthly observations.

The remaining indicators are built by trimming away part of the data of each series. The range considers only the extremes:

$$RRange_k = \frac{\max(k_i) - \min(k_i)}{\mu_k} \qquad (8)$$

the MAD (median absolute difference from the median) takes the median of all the distances from the median:

$$RMAD_k = \frac{median(|k_i - median(k_i)|)}{median(k_i)} \qquad (9)$$

the interquartile range uses only the extreme quartiles (Q_1 and Q_3) of each distribution:

$$RQrange_k = \frac{Q_3(k_i) - Q_1(k_i)}{median(k_i)} \qquad (10)$$

Time-linear dynamics

A second set of indicators has been adopted in order to account for some aspects of the linear relationships with time attributable to each series. In particular, the coefficient of determination R^2 evaluates the goodness of fit provided by the model:

$$x_i = \alpha + \beta i + \varepsilon_i \qquad (11)$$

[17] The absolute value of the underlying variability indicator has been divided by a level indicator (the mean or the median). It is important to remark that expressions from (6) to (10) measure variability independently of the order of the observations and simply on the basis of the distribution of their values.

where i labels the months. The autocorrelation coefficients, with lags equal to 1 and 12 months, are also estimated using the following model:[18]

$$x_i = -A_1 x_{i-1} - A_2 x_{i-2} \dots - A_{12} x_{i-12} + \varepsilon_i \tag{12}$$

Other dynamics indicators

Two further indicators summarise the difference, within each country, between the yearly rates of each sub-index and those of the corresponding aggregate HICP. For the i-th month we have:

$$d_i = y_i - y_i^{CP00} \tag{13}$$

where y_i^{CP00} is the yearly rate of growth of the aggregate HICP in month i. The Euclidean distance between the two series is derived by averaging the squares of these residuals:

$$D2_y = \sqrt{\frac{\sum_i d_i^2}{n_y}} \tag{14}$$

while the mean of the residuals delivers an evaluation of the systematic discrepancies between them:

$$D1_y = \frac{\sum_i d_i}{n_y} \tag{15}$$

A last indicator has the objective to evaluate the propensity of each index to move monthly and it is given by the share of months in which each series has shown a change with respect to the preceding month. In particular:

$$M_z = \frac{\sum_i M_{iz}}{n_z} \tag{16}$$

where

$$M_{iz} = \begin{cases} 1 \; if \; z_i \neq 1 \\ 0 \; if \; z_i = 1 \end{cases} \tag{17}$$

[18] The coefficient A_{12} is used in order to give a rough indication on the yearly seasonal nature of the series, while A_1 is used in order to evaluate short term autocorrelation. Given this formula, a negative value for the coefficients of this linear relation implies a positive autocorrelation, and vice versa. This aspect should be reminded when commenting tables and principal factors in par.4.

Overall features of indices' variability

The final matrix that is used in the analysis is thus composed by 21 indicators referred to 1.257 series. Relative variability is in general higher for fixed base indices (Table 2).[19] The average values assumed by these indicators are all higher than their median, given a positive asymmetry generated by a limited number of series with high or very high levels of relative variability and a large number of series whose behaviour is instead very smooth. The asymmetry is more evident for yearly and monthly rates. Most fixed base series show an approximately time-linear longitudinal pattern: R^2 is very near to unity for 25% of the series, its mean is well below the median and the first quartile is still high. An appreciable lag 12 autocorrelation appears in less then 10% of the series.[20]

The distribution of the distances from aggregate series is also positively skewed. Series diverge in fact on average by four percentage points (pp) from the corresponding aggregate HICP; but, while for a half of them this average is lower than 2.5 pp, for 10% of the series *D2* is over eight pp. As a consequence, the median is lower than the arithmetic mean. The distribution of *D1* appears instead nearly symmetric. Nevertheless, both 5%-tails reveal the existence of series which diverge systematically from the aggregate HICP: in particular, the number of sub-indices running below the aggregated index appears slightly higher. Finally, according to the variable *M*, on average series have a 77% propensity to monthly movement: the distribution is in this case negatively skewed.[21]

Two objectives are pursued in the following paragraphs 4 and 5: to obtain a classification of HICP sub-indices according to the characteristics of their behaviour and to measure the degree of heterogeneity across countries of the series belonging to a same elementary aggregate. The classification is obtained by applying a clustering algorithm to the scores obtained by the prior application of principal components (PC) to the matrix of indicators.[22] This paragraph describes the main results of PC and some partitions of the original data derived from cluster analysis. In paragraph 5 the coordinates corresponding to the main principal components are used to provide a measure of the entropy within each elementary aggregate and to highlight the aggregates that might need more urgently further methodological harmonisation.

---

[19] This is obvious since yearly and monthly rates refer to more constrained time spans.

[20] Given expression (12), negative coefficients imply positive autocorrelation.

[21] There is a tail of a 10% of the series which tend to show some movement only in one month out of four. For the estimates of some sub-indices, many NSIs in fact collect prices only on a quarterly basis instead of monthly. Furthermore, imputation is rarely adopted to replace missing monthly observations and last available estimates are often simply carried over. EUROSTAT has ruled out such practices. See the Council regulation (EC) No 701/2006 of 25 April 2006 laying down detailed rules for the implementation of Regulation (EC) No 2494/95 as regards the temporal coverage of price collection in the harmonised index of consumer prices. Official Journal, L 122, November 15, p. 3-4.

[22] For the use of this, so called, "tandem analysis" to business statistics see Anitori et al. (2004). As regards more specific applications to price statistics, see Brunetti et al. (2004) and ISTAT (2005, p.42-44).

Table 2 - Averages and percentiles of the distribution of variability indicators

INDICATOR	Mean	Percentiles									Median/ Mean
		1st	5th	10th	Q1	Median	Q3	90th	95th	99th	
						INDEX LEVELS					
CV	5,4	0,7	1,2	1,6	2,7	4,4	6,5	9,9	14,1	24,9	0,80
RQrange	8,1	0,5	1,3	1,9	3,6	6,4	10,0	15,2	21,9	38,7	0,79
RGini	6,1	0,7	1,3	1,8	3,0	4,9	7,5	11,1	15,7	27,5	0,80
RMad	3,8	0,2	0,5	0,8	1,6	2,9	4,8	7,1	10,0	18,6	0,77
RRange	19,2	2,3	4,4	5,9	9,3	14,6	22,6	36,9	50,5	94,5	0,76
R^2	0,73	0,00	0,07	0,19	0,60	0,88	0,96	0,98	0,99	1,00	1,19
A_1	-0,8	-1,3	-1,1	-1,1	-1,0	-0,8	-0,7	-0,4	-0,3	0,0	1,05
A_{12}	0,0	-0,5	-0,3	-0,2	-0,1	0,0	0,1	0,1	0,2	0,2	-0,71
						YEARLY CHANGES					
CV	2,5	0,3	0,4	0,6	0,9	1,6	3,0	5,4	7,6	15,3	0,64
RQrange	3,6	0,3	0,5	0,7	1,2	2,2	4,2	7,5	11,0	21,7	0,63
RGini	2,8	0,3	0,5	0,7	1,0	1,8	3,3	5,9	8,5	16,9	0,64
RMad	1,5	0,1	0,2	0,3	0,6	0,9	1,7	3,1	4,5	9,1	0,64
RRange	9,6	0,8	1,7	2,2	3,5	6,1	11,6	20,5	29,7	57,3	0,64
D2	4,0	0,7	1,0	1,2	1,6	2,5	4,5	8,2	12,0	22,0	0,63
D1	-0,4	-16,4	-6,4	-3,7	-1,6	-0,2	1,3	3,0	4,8	11,2	0,48
						MONTHLY CHANGES					
CV	1,4	0,1	0,2	0,3	0,4	0,7	1,4	3,4	5,6	9,3	0,52
RQrange	1,1	0,0	0,0	0,0	0,2	0,4	1,0	2,9	4,8	9,6	0,39
RGini	1,3	0,1	0,2	0,3	0,4	0,6	1,2	3,4	5,7	9,6	0,46
RMad	0,5	0,0	0,0	0,0	0,1	0,2	0,5	1,3	2,3	4,5	0,36
RRange	7,7	0,6	1,1	1,3	2,3	4,2	8,4	18,7	26,5	53,6	0,55
M	0,77	0,09	0,15	0,28	0,65	0,90	0,97	1,00	1,00	1,00	1,17

Source: Elaborations on Eurostat HICP data base

4. Main results from multivariate data analysis

4.1 Principal components

No weights have been applied to this analysis.[23] The first two principal components account for 68.1% of total inertia, the first five overcome 90% (Table 3).[24] The first factor might be interpreted as expressing general variability: higher scores identify series with a high overall dynamics, especially in the level of the fixed base index and in its yearly growth rates. Furthermore this factor has a positive correlation with the average quadratic distance from the overall HICP, while no significant correlation emerges with the variables concerned with time-linear dynamics.

23 Some experiments have been made by weighting series, but no relevant changes in the overall results occurred. The structure of the data base is quite robust, and remains substantially unvaried if expenditure and/or country weights are introduced.

24 The sum of the eigenvalues of the first ten principal components nearly reaches 99% of total inertia.

Table 3 - Correlation between the first five principal components and the original variables

VARIABLE	Principal components				
	Prin1	Prin2	Prin3	Prin4	Prin5
Eigenvalue (%)	50.0	18.1	11.3	7.0	4.2
		INDEX LEVELS			
CV	0.84	0.46	0.27	0.02	0.01
Qrange	0.71	0.52	0.38	-0.02	-0.01
Gini	0.82	0.47	0.30	0.01	0.02
MAD	0.66	0.52	0.46	-0.06	0.05
Range	0.88	0.33	0.24	0.00	0.04
R^2	-0.21	0.66	0.28	0.09	-0.08
A_1	0.08	-0.51	0.53	0.10	-0.03
A_{12}	-0.01	0.37	-0.53	0.07	-0.05
		YEARLY CHANGES			
CV	0.88	0.05	-0.41	0.00	0.07
Qrange	0.84	0.12	-0.44	0.05	-0.05
Gini	0.89	0.05	-0.41	0.10	-0.06
MAD	0.83	0.06	-0.35	-0.18	0.23
Range	0.91	-0.06	-0.33	-0.01	0.04
D2	0.87	0.34	-0.01	-0.19	0.23
D1	0.14	-0.04	-0.50	0.16	-0.16
		MONTHLY CHANGES			
CV	0.74	-0.57	0.15	0.29	0.41
Qrange	0.67	-0.59	0.17	0.34	-0.14
Gini	0.72	-0.62	0.18	-0.47	-0.11
MAD	0.66	-0.59	0.15	-0.08	-0.22
Range	0.75	-0.50	0.07	0.60	0.52
M	0.21	-0.17	0.03	-0.72	0.44

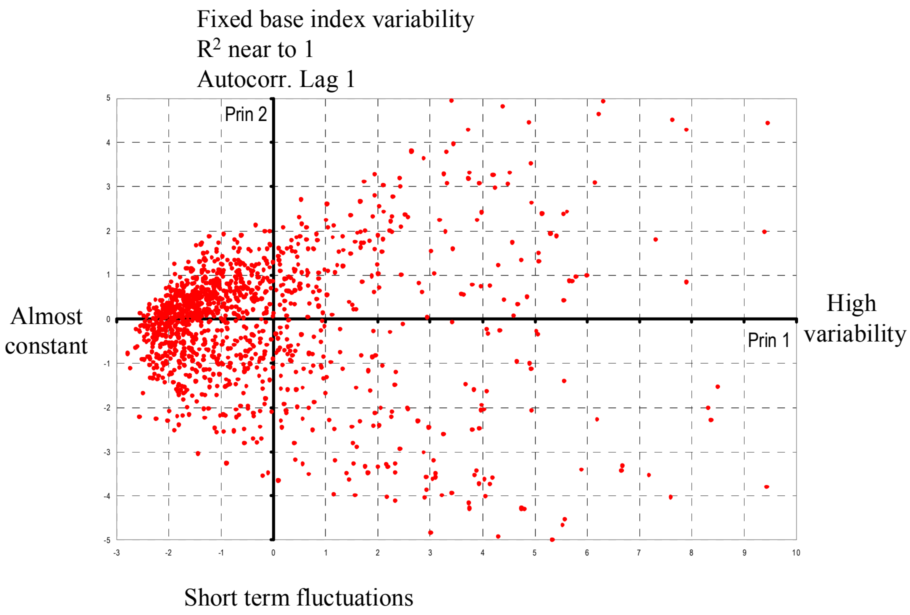
Source: Elaborations on Eurostat HICP data base

The second factor accounts for the relationship between overall variability and time-linear dynamics: in particular, the scores are directly related with the variability of the fixed base index, and negatively related with the variability of its monthly changes. This component shows also a positive relation with R^2 and with lag 1 autocorrelation. Higher scores identify series with a regular trend - positively or negatively sloped - and consequently a higher variability in the level of the fixed base index as expressed in particular by *MAD* and *QRange*. Negative scores at the opposite identify fluctuations and, as a consequence, index levels have a smaller range.

In a similar way, the third principal component shows a negative relation with the variability of yearly rates and a positive one with that of fixed base indices. The higher the score on this axis, the more the yearly rates tend to be constant, the lag 12 autocorrelation high and the lag one low. Positive values of this factor characterise series whose behaviour has some yearly seasonality and whose yearly changes tend to be systematically lower than those of the corresponding aggregate HICP. On the negative side the opposite applies: it is possible to find there series with quite irregular yearly rates of change, which tend to be higher than the aggregate HICP.[25]

Most sub-indices show negative scores on the first component (relatively low variability profiles), while their distribution on the second appears more even (Figure 1).[26] The more we move to positive values of the first, the more the coordinates on the second component tend to increase their dispersion: the position of each series depends respectively on whether it assumes a monotone (north-east) or an erratic (south-east) short term dynamics.[27]

Figure 1 - The scatter of HICP series in the space given by the first two components



Source: Elaborations on Eurostat HICP data base

25 The fourth factor captures the inverse association between seasonality and faster yearly growth on one side and the frequency of monthly changes on the other, while the fifth shows the direct association between these last two variables.

26 Only 31% of the series show a positive coordinate on the first factor while 58% have a positive one on the second.

27 Something similar could be said with reference to the form of the scatter in the space given by the first and third factor. The variability on the third factor is quite reduced, also in comparison with the second one. In any case, as overall variability increases along the first axes, the third one discriminates between more regular seasonal behaviours (positive values) and occasional sharp level shifts (negative).

4.2 Clustering series

We explored the possibility to classify series according to their behaviour without any a priori hypothesis on the structure of the classification itself. For this purpose we opted for the use of hierarchical aggregative clustering algorithms in order to investigate alternative segmentations of the series. The results that we present here are referred to the application of the Ward (1963) algorithm, although there is a high similarity with those obtained with average linkage clustering. The Ward algorithm has been applied to the main principal components: alternative solution have been explored including from a minimum of five up to twelve components (99.6% of total variability), and the results appeared quite robust with only marginal changes. We will discuss then the twelve components case and a hierarchy of partitions ranging from two to nine clusters (see Table 4).[28]

The partition in two main clusters, A and B, explains 27.4% of total variance. It separates the series in two exhaustive sets: the series with a very low variability (cluster A) and those with an appreciable variability. Cluster A includes the large majority of the series, more than 900 corresponding nearly to the 75% of the total (see Table 5). At a deeper level of detail, each main cluster can be split in three distinct clusters, and in the case of cluster B the breakdown has been brought further to six clusters. The partition in four clusters, with the original cluster B split in three subsets (BB, BC and BD), brings a considerable gain in the share of the variance between (near to 50% of total variance: see Table 4). The further partition in nine clusters brings this share over 67%.[29]

Table 4 - Cluster aggregation

SPLIT	IN	Variance between	Clusters	
			No.	Name
Population	A and B	27.4%	2	A, B
B	(BC+BD) and BB	38.0%	3	A, BB, (BC+BD)
(BC+BD)	BC and BD	46.7%	4	A, BB, BC, BD
BD	BD1 and BD2	54.4%	5	A, BB, BC, BD1, BD2
A	(A1+A2) and A3	58.3%	6	(A1+A2), A3, BB, BC, BD1, BD2
BB	BB1 And BB2	61.8%	7	(A1+A2), A3, BB1, BB2, BC, BD1, BD2
BC	BC1 and BC2	64.9%	8	(A1+A2), A3, BB1, BB2, BC1, BC2, BD1, BD2
(A1+A2)	A1 and A2	67.2%	9	A1, A2, A3, BB1, BB2, BC1, BC2, BD1, BD3

Source: Elaborations on Eurostat HICP data base

[28] Further partitions have also been examined, although the gains are rather smooth (the marginal contribution of further partitions was less then 2% of total variability): it is possible to explain 75% of total variance with 16 clusters, which mainly derive from splitting A1, A2, BB1, BC1 and BD1.

[29] Further partitions have also been examined, although the gains are rather smooth: it is possible to explain 75% of total variance with 16 clusters, who mainly derive from splitting clusters A1, A2, BB1, BC1 and BD1.

Table 5 - Cluster size and centre

CLUSTER	DESCRIPTION	Series		Principal components				
		No.	%	Prin1	Prin2	Prin3	Prin4	Prin5
Total		1257	100					
A	Low variability	921	73.3	-1.5	0.6	0.2	-0.3	0.6
B	Higher variability	336	26.7	3.9	0.0	0.0	0.0	-0.1
A1	Low variability discrete	256	20.4	-1.6	0.1	0.6	1.3	-0.2
A2	Low variability monotone	393	31.3	-1.5	0.6	0.2	-0.3	0.6
A3	Low variability irregular	272	21.6	-1.2	-0.9	-0.8	-0.8	-0.5
BB	Medium variability non time linear	108	8.6	4.2	-4.0	1.3	-0.2	0.1
BC	Medium variability time-linear	180	14.3	2.7	1.1	-1.4	0.4	0.1
BD	High variability	48	3.8	8.2	4.4	2.3	-1.3	-0.9
BB1	Regular seasonal	92	7.3	3.1	-3.7	1.3	-0.2	0.1
BB2	Irregular seasonality and fluctuations	16	1.3	10.7	-5.9	1.0	-0.4	0.5
BC1	Marked irregular time-linear trend	61	4.9	5.5	1.2	1.1	-1.0	0.3
BC2	Irregular time-linear trend	119	9.5	1.2	1.1	-1.0	0.3	0.0
BD1	Strongly downward time-linear trend	36	2.9	5.6	5.0	4.4	-2.5	-1.6
BD2	Strongly irregular yearly rates	12	1.0	16.2	2.7	-4.1	2.4	1.1

Source: Elaborations on Eurostat HICP data base

Lower variability (A)

The sub-indices in the lower variability domain on average show a relative variability between two to five times lower than the average of cluster B, with a low propensity to monthly change: they in fact have mainly negative scores on the first principal component. This huge area of low variability can be further split in three distinct clusters (A1, A2 and A3), on the basis of a slight differentiation of the scores on the second and third component.

Cluster A1 identifies sub-indices with small and discrete movements: it contains 20.4% of the series, which have on average the lowest coordinates on the first component (Table 5). They are characterised by a very low propensity to move (Table 6). Yearly changes are particularly weak and monthly changes as well, especially if extremes are trimmed off.

Cluster A2 concentrates 31.3% of the series, which can be labelled as "Monotone" since they show very small and regular monthly changes. As a consequence, the indicators relating to index levels - and based on the range - are a slightly higher than the rest of cluster A. The yearly rates lay on average shortly below the aggregate HICP. In this cluster and in A1, series have on average a R^2 close to unity.

The behaviour of the 21% of the sub-indices which are included in cluster A3 shows some irregularity. Yearly and monthly changes are more volatile and the series fluctuate moderately: the time-linearity of the series, as measured by the R^2, appears very limited. The range of index levels is consequently more constrained: their variability is on average almost one half with respect to the rest of cluster A. The inflationary profile of these series lays well below the aggregate HICP.

Medium variability and non linearity (BB)

In considering the remaining 336 series classified in the main cluster B it is possible to find a first domain composed by series (8.6% of the total) showing a moderate variability - the double or more with respect to cluster A - associated with remarkable fluctuations. Clusters BB1 and BB2 are in fact characterised by lower scores on the second component. R^2 is relatively low while the variability of the monthly rates is very high: both clusters contain the series with the highest short term variability.

In particular, cluster BB2 (1.3% of the series) is characterised by fluctuations and quite irregular seasonal patterns. The average variability of yearly changes in BB2 triples that registered in cluster BB1 (7.3% of the total), which is characterised by a regular seasonality. The series included in BB2 are in fact more variable, not only with regard to the levels of the indices but also to their yearly and monthly changes. Both BB clusters show on average a weaker inflationary profile with respect to the aggregate HICP (especially BB2).

Medium variability and time-linear dynamics (BC)

A marked linear evolution with time (R^2 is on average relatively high, near to 0.7) is shown in clusters BC1 (4.9% of the series) and BC2 (9.5%). Their variability is similar or slightly higher with respect to cluster BB1. The sub-indices in BC1 are more volatile, with a slightly lower R^2. Lag 1 autocorrelation is in both clusters very high. The yearly rates are more irregular in BC1 while the variability of monthly changes is significantly lower as compared to BB. These series, and in particular those in BC1, have an upwards dynamics significantly stronger as compared to the aggregate HICP.

High variability (BD)

The clusters BD1 and BD2 include as a whole 3.8% of the series all characterised by very large movements: *CV* is on average over 20% and the relative range is over 70%. Series in BD2 are characterised by large fluctuations in the yearly rates. On the contrary those in BD1 denote a remarkable downward trend: these sub-indices have an R^2 very close to one and a negative time-linear trend, whose regularity is witnessed by the comparatively smaller variability of monthly and yearly changes. On the other side, the yearly growth rates in BD2 are much higher *vis-à-vis* the aggregate HICP, and they come along with large fluctuations both on monthly and yearly rates.

Table 6 - Average indicators by cluster

INDICATOR	Clusters								
	A1	A2	A3	BB1	BB2	BC1	BC2	BD1	BD2
	FIXED BASE INDEX (BASE: AVERAGE 2005=100)								
CV	4.0	4.4	2.4	6.6	11.8	11.8	7.5	20.9	24.4
RQrange	6.2	7.2	3.1	9.0	16.4	16.2	11.1	34.4	31.3
RGini	4.6	5.0	2.6	7.3	13.1	13.0	8.1	24.1	26.9
RMad	3.0	3.4	1.4	4.1	7.7	7.5	4.6	17.1	12.3
RRange	13.4	14.9	9.1	27.3	53.8	42.4	23.8	72.4	95.7
R^2	0.91	0.93	0.41	0.34	0.22	0.66	0.76	0.98	0.74
A_1	-0.67	-0.86	-0.83	-0.47	-0.49	-0.94	-0.95	-0.82	-0.88
A_{12}	-0.12	0.02	0.02	-0.15	-0.20	0.08	0.04	0.02	0.09
	YEARLY CHANGE MOVING BASE INDEX								
CV	1.2	1.2	2.1	3.4	9.4	8.4	4.5	3.8	17.5
RQrange	1.7	1.7	2.9	4.0	11.9	12.5	6.5	5.4	28.6
RGini	1.3	1.3	2.3	3.6	10.5	9.1	4.8	4.2	19.4
RMad	0.7	0.7	1.2	1.9	5.6	4.7	2.3	2.5	10.8
RRange	4.4	4.4	8.0	15.7	41.6	30.0	14.9	14.4	64.3
D2	2.0	2.3	3.0	4.2	10.0	10.7	5.5	16.3	26.3
D1	0.1	-0.3	-1.8	-0.8	-1.5	5.1	1.9	-15.9	16.6
	MONTHLY CHANGE MOVING BASE INDEX								
CV	0.8	0.5	0.9	5.5	9.6	2.8	1.3	1.4	6.1
RQrange	0.3	0.5	0.8	4.8	10.9	2.0	0.6	1.4	4.6
RGini	0.6	0.5	0.9	5.8	10.5	2.5	1.0	1.5	5.8
RMad	0.1	0.2	0.4	2.2	5.3	1.0	0.2	0.6	2.2
RRange	4.3	3.0	5.5	26.5	47.9	16.4	8.3	7.4	35.1
M	0.45	0.86	0.86	0.94	0.98	0.80	0.71	0.98	0.88

Source: Elaborations on Eurostat HICP data base

5. The analysis of elementary aggregates

5.1. Heterogeneity of sub-indices

The country series referred to a same elementary aggregate are often positioned in separate clusters, and this is one of the aspects which appear as more interesting in order to target further improvements in harmonisation. If we consider the less detailed partition in two clusters (A and B), it is possible to identify only 25 elementary aggregates (slightly more than one out of four) whose national series are all concentrated in a same cluster (almost exclusively in cluster A) while the remaining 69 are split in the two clusters (see Table 7) and are thus characterised by the existence of some heterogeneous patterns in national series.

If we consider the further partition in four clusters, 24 elementary aggregates still belong to a same cluster while, at the opposite, 13 aggregates are separated in more than two distinct clusters. The most detailed partition in nine clusters shrinks to two the number of elementary aggregates whose national series all belong to a same cluster, while 44 aggregates are split in more than three clusters.

Table 7 - Splitting of elementary aggregates into clusters

COUNTRY SERIES	Partitions		
	2 clusters	4 clusters	9 clusters
All in the same cluster	25	24	2
In two clusters	69	57	15
In three clusters		13	33
In four clusters			34
In five clusters			8
In six clusters			2
Total aggregates	**94**	**94**	**94**

In order to derive a measure of the heterogeneity within each elementary aggregate j, it is proposed here an indicator based on the Euclidean distance (ED) of each series from the centre of the corresponding elementary aggregate on the space determined by the first Q principal components. Country distances have been then averaged in order to obtain the mean quadratic distance within each aggregate. In particular:

$$ED_j^q = \sum_g \sqrt{\sum_{q=1}^{Q} \left(c_{qjg} - \bar{c}_{qj}\right)^2} \tag{18}$$

where c_{qjg} are the coordinates on the principal factor q of the series for COICOP j in country g, while:

$$\bar{c}_{qj} = \frac{\sum_g c_{qjg}}{n_g} \tag{19}$$

are the coordinates of the centre of COICOP class j on factor q, derived as an unweighted average of country coordinates.[30] Furthermore, n_g is the number of countries taken into consideration ($n_g = 14$).

Table 8 ranks the most heterogeneous aggregates. In particular, "passenger transportation by air" and "package holidays" seem to represent by far the sectors in which the behaviour of the national series is the most heterogeneous. Transports by sea and inland waters also show high heterogeneity; the same applies to hi-tech and books, to clothing and to some goods and services related to housing. Fruit and vegetables are also in this list of heterogeneous cases, like accommodation, hospital services and many public utilities.

[30] Other distance measures have been tried, all bringing roughly the same results.

Table 8 - The 25 elementary aggregates with the highest heterogeneity

DESCRIPTION	ED	DESCRIPTION	ED
Passenger transport by air	6.04	Garments	3.01
Package holidays	5.01	Vegetables	3.00
Hospital services	4.06	Electricity	3.00
Clothing materials	4.04	Equip. for the reception, recording and reproduction	3.00
Heat energy	4.04	Shoes and other footwear	2.09
Passenger transport by sea and inland waterway	4.02	Books	2.08
Photographic and cinematographic equip. and optical instr.	4.01	Education	2.07
Gas	3.09	Refuse collection	2.07
Liquid fuels	3.06	Other articles of clothing and clothing accessories	2.06
Accommodation services	3.04	Water supply	2.06
Information processing equipment	3.03	Jewellery, clocks and watches	2.05
Gardens, plants and flowers	3.01	Fruit	2.04
Sewerage collection	3.01		

Source: Elaborations on Eurostat HICP data base

How can we find an explanation for heterogeneity? Does it derive from differences in the actual behaviour of national markets? Or is it generated by some methodological reason revealing a lack of harmonisation, for example in sampling design, stratification, elementary aggregation, replacements, weighting etc?

This analysis can help to focus on strong disharmony due to sampling issues, where it can hardly be explained by differences in the functioning of national markets. Nevertheless, it cannot be straightforwardly assumed that high homogeneity is synonymous of sound methodological harmonisation. For instance, restaurants make up the most homogeneous aggregate (Table 9), although the task of estimating the corresponding sub-index is rather difficult. High homogeneity also concerns some foodstuff (i.e. meat, wine, bread, etc.), fuels (which mainly obey to the international markets and whose prices can be quite easily observed with a zero influence of replacements). Cars are also in this group of homogeneous aggregates, due the smooth behaviour of the estimates based on list prices.[31]

[31] NSIs may in fact be adopting similar but not methodologically correct designs, as it may happen for example with cars, where list prices are widely used instead of actual bargain prices. The use of the latter would probably induce heterogeneity in the behaviour of the indices (De Gregorio 2010).

Table 9 - The 25 elementary aggregates with the lowest heterogeneity

DESCRIPTION	ED	DESCRIPTION	ED
Restaurants, cafés and the like	0.07	Materials for the maintenance and repair of the dwelling	1.03
Pets and related products	0.09	Fuels and lubricants for personal transport equipment	1.03
Spare parts for personal transport equipment	1.00	Food products not elsewhere classified	1.03
Cultural services	1.00	Cleaning, repair and hire of clothing	1.03
Meat	1.00	Motor cars	1.03
Wine	1.01	Services for the maintenance and repair of the dwelling	1.03
Maintenance and repair of personal transport equipment	1.01	Non-durable household goods	1.03
Bread and cereals	1.01	Mineral waters, soft drinks, fruit and vegetable juices	1.04
Appliances for personal care	1.01	Tools, equipment and miscellaneous accessories	1.04
Motor cycles, bicycles and animal drawn vehicles	1.02	Recreational and sporting services	1.04
Newspapers and periodicals	1.02	Miscellaneous printed matter and stationery materials	1.04
Sugar, jam, honey, chocolate and confectionery	1.02	Milk, cheese and eggs	1.04
Canteens	1.02		

Source: Elaborations on Eurostat HICP data base

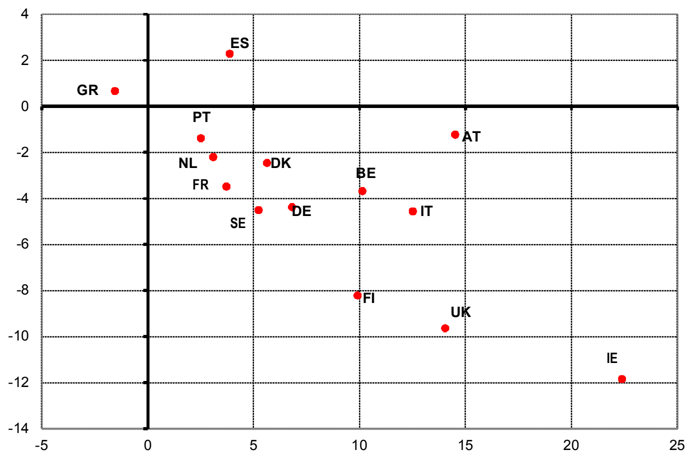
5.2 Some case studies

The possible reasons that lay behind the lack of homogeneity may concern the behaviour of national markets or the approach to sampling, or both. Three different cases are here presented: air transport, where heterogeneity seems to derive from specific methodological issues; package holidays, where a different sort of methodological differences appears; electricity, where the peculiarity of national markets seems to prevail over methodological explanations.

Passenger transportation by air

National indices provide very different pictures of air transports across the EU. This can be visualised in the two dimensions space given by the first two components (Fig. 2). Countries appear all distributed along the diagonal running from north-west to south-east. With the only exception of Greece and Spain, all countries have positive coordinates on the first principal component and negative on the second. The series with a higher variability in the level of the index appear influenced by short term fluctuations and in some cases by seasonal behaviours. Greece is the only country with a negative coordinate on the first factor: its sub-index moves discretely and remains constant for several months (nearly one year). Spain, whose sub-index has a relatively low variability, is the only case of positive coordinate on the second component: in fact it has a time-linear pattern, with an upward trend. The remaining countries all show very sharp monthly changes and fluctuations: the differences among them mainly concern the intensity of movements.

Figure 2 - The coordinates of HICP series for air transports on the space given by the first two principal components, by country



Source: Elaborations on Eurostat HICP data base

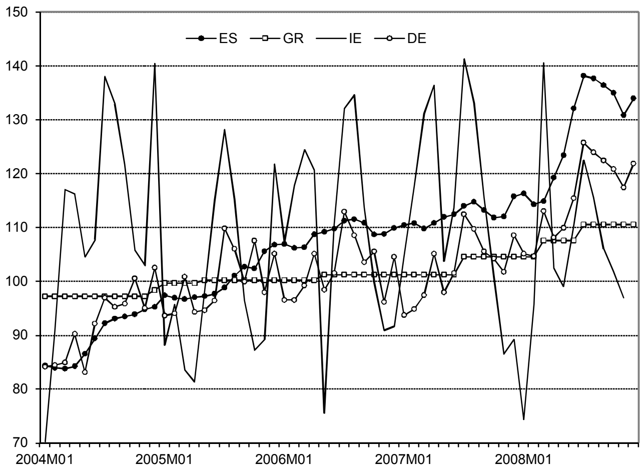
Nevertheless, such heterogeneous behaviours look quite implausible and this impairs the explanatory power of any euro area or EU average for the price index of this sector. Several elements might help to explain them. The characteristics of national markets may play some role: the market power of incumbent national carriers may vary across countries, as well as the barriers to entry and the degree of penetration of low cost carriers. Some differences might occur if the size of the country, that is the existence of a national market, is taken into consideration. Nevertheless competition and the appearance of new operators has taken place almost everywhere in Europe, and especially in the lower segments, like those accounted for by low cost operators. This has brought everywhere competitive pressures and shorter term pricing policies. For this reason, it is hard to maintain that in the EU coexist markets where price dynamics are contemporarily those shown in figure 3.

It seems likely that the lack of homogeneity within this elementary aggregate cannot be satisfyingly explained on the basis of structural differences in national markets: this is quite evident if we compare Greece and Ireland although it might be also interesting to explain the differences among less heterogeneous series. Methodological explanations are very probably playing a major role. Although the literature on CPI methodology for this sector is quite small[32] and cannot help much, it is easy to guess that these explanations may involve the sampling design: namely, the size and selection of the sample, the use of sampling stages and stratified designs, the price collection technique, the methods for replacements and elementary aggregation, the implicit hypotheses concerning the elasticity of consumer behaviours. Some answers to easy questions may clarify things. For example: how many airlines are included in the sample? And on how many routes? Which prices are followed?

[32] See for example Lent at al. 2005, Good et al. 2008.

The official prices, or the actual prices, or those derived from purchase simulations on the internet? How does it take place the selection of a route? Which types of destinations were considered in the sample? How is it defined the stratification? How is it considered the presence of a plurality of offers on the same route? Are the prices from different carriers directly compared?

Figure 3 - HICP series for air transports for Spain, Greece, Ireland and Germany (Jan 2004-December 2008) (base: average year 2005=100)



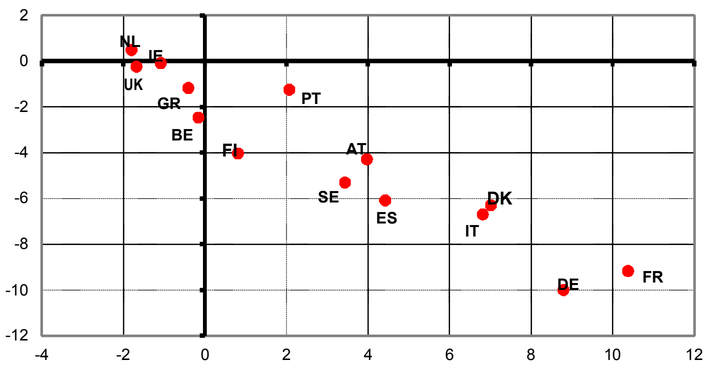
Source: Elaborations on Eurostat HICP data base

It is evident that quite different approaches coexist, and that there is a wide scope for harmonising them within the existing HICP legal basis. Moreover, this market has been given a great deal of attention from the EU Commission, for instance for what concerns transport policy, competition and passenger rights: one might ask whether the HICP is adequate to describe this market and to drive policies. The answer seems negative, but at the same time many elements are at hand in order to bring on the necessary improvements and a common view on the criteria to be used to provide these estimates.

Package holidays

The case of package holidays looks at least in part similar to air transport. Countries tend to be positioned on the diagonal north-west/south-east, although the scale of figure 4 is significantly reduced. Five countries have negative scores on the first principal component, and they all show similar patterns: very smooth index as in the case of Ireland, UK and The Nederland's, with some fluctuations as in the case of Belgium and Greece which have a negative coordinate on the second component. On the contrary, the rest of the countries show fluctuations and seasonality which become particularly strong in the case of Germany and France.

Figure 4 - The coordinates of HICP series for package holidays on the space given by the first two principal components, by country



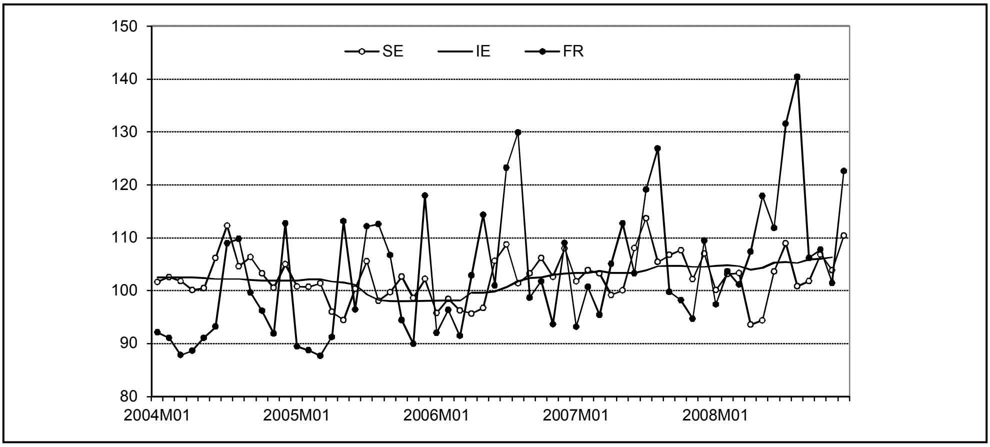
Source: Elaborations on Eurostat HICP data base

If we look at the behaviour of the series (Fig. 5), it is possible to appreciate the large differences in the estimated price dynamics. It looks very unlikely to attribute them to the working of national markets: foreign destinations are in fact very similar across European tour operators. Methodological aspects are certainly playing an all but secondary role: in particular, the way in which the sub-indices account for the seasonal nature of pricing in this sector appears to be a relevant issue. The use seasonal baskets[33] may explain the absence of fluctuations in the behaviour of some indices. Other sources of heterogeneity may derive from the treatment of missing observations and from the possibly different imputation techniques adopted to treat seasonal products. A lack of harmonisation may also concern the type of stratification and sample size. Differences may arise on the management of replacements, which appear to be quite frequent in this market: the adoption of overlap or bridged overlap on one side, and direct comparisons or other imputation techniques on the other, may bring sharply different results in terms of overall variability. A common approach, and in particular a proper common definition of consumption segments, might help considerably to harmonise estimates.[34]

[33] On the seasonal basket approach followed in Ireland in the case of package holidays, see CSO (2003). On the treatment of package holidays in the Italian HICP see De Gregorio et al. (2008), p.28-29.

[34] Having a good price index for this aggregate has also some importance for policy. Package holidays have in fact been the object of the focus of EC policies, especially concerning consumer protection. Common rules to be followed by tour operators in order to protect consumers were defined since 1990 within the normative body on package travel, package holidays and package tours. See EC (1999, 2007b).

Figure 5 - HICP series for package holidays for Sweden, Ireland and France (Jan 2004-December 2008) (base: average year 2005=100)

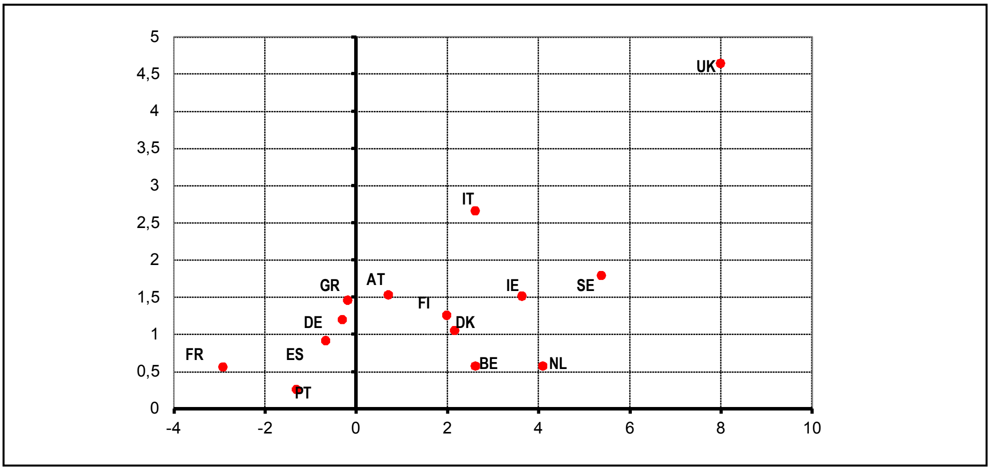


Source: Elaborations on Eurostat HICP data base

Electricity

In the case of electricity, countries all show positive scores on the second component, with indices that generally change smoothly and with time-linear patterns (Fig. 6). Five countries have negative values on the first component: in particular the French index moves discretely and remains constant for very long periods; something similar happens to the series of Portugal, Spain, Greece and Germany. Although discrete shapes seem to prevail, other countries show more variability, mainly due to a higher frequency of their movements and a more regular (upward) trend. The highest overall variability is shown by the UK series, where the sub-index shows a regular upward trend with a major fluctuation towards the end of 2008 (Fig. 7).

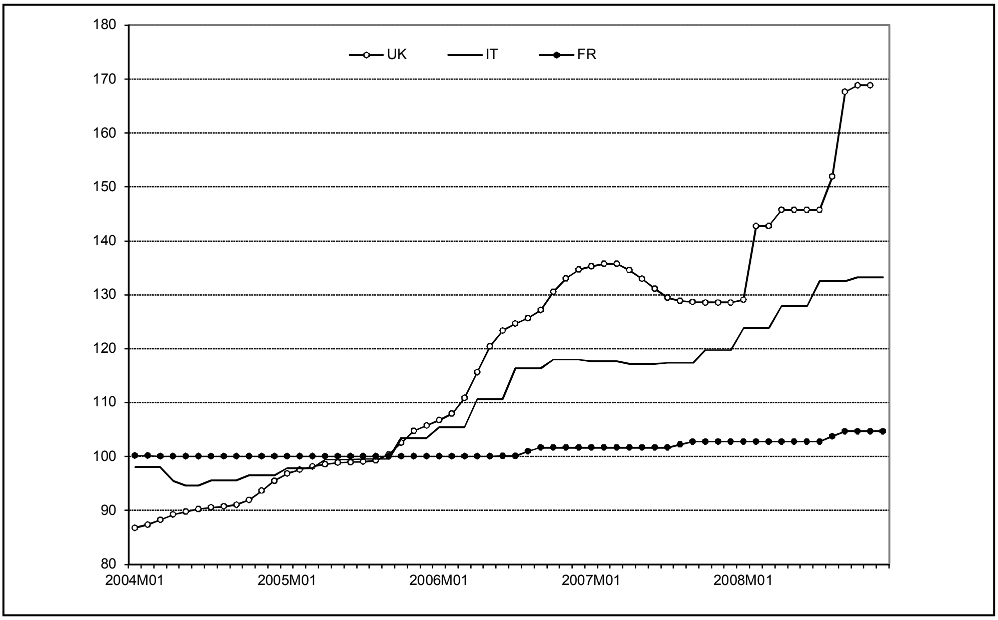
Figure 6 - The coordinates of HICP series for electricity on the space given by the first two principal components, by country



Source: Elaborations on Eurostat HICP data base

An accurate description of national features in electricity markets may be very helpful in explaining the differences in the behaviour of these series. From 1996 the EC has agreed for liberalisation, and since then a constant monitoring of the market has been provided. In 2007 the Commission has proposed further liberalisation. In the meantime several cases have been analysed by the European antitrust authority.[35] As a consequence, in some countries pricing in the market for electricity has, at least partly, a legal basis. Often consumers can choose between the regulated market and the liberalised one, where structurally different tariff structures usually are applied: differences in the behaviour of sub-indices might thus be determined by differences in the degree of regulation or by other structural differences that keep on existing on national markets. It follows that electricity is likely to be affected to a lesser extent by methodological heterogeneities, at least as compared to air transport or package holidays. Nevertheless, the issue of harmonisation is anyway crucial since the lack of homogeneity in the series may also be partly attributed to differences in sampling design and in particular to a different degree of coverage of the different segments of the whole market. The use of consumer profiles appears a promising perspective, even more if we consider the increasing use of non linear pricing (for instance, differentiated hourly tariffs).

Figure 7 - HICP series for electricity for United kingdom, Ireland and France (Jan 2004-December 2008) (base: average year 2005=100)



Source: Elaborations on Eurostat HICP data base

35 For an overview of EC electricity policy see http://ec.europa.eu/competition/sectors/energy/electricity/electricity_en.html. See also the sector inquiry, started in 2005 and presented in 2007 (EC, 2007a).

Conclusions

Monetary policy is progressively losing its exclusivity in the use of consumer price statistics, since multiple uses of CPI data concerning industrial policies and consumer markets monitoring are rapidly gaining ground. This brings important challenges to the role of price indices as a valid support to such policies, since the main focus shifts from the aggregate price index to its components, defined with a high level of detail and regularly disseminated by NSI's and EUROSTAT. This implies the need to guarantee the methodological soundness of such indices and a reasonable ground for comparability: in the case of the EU, where the HICP project has been developed since 1996, the need of comparability is an increasingly crucial issue.

Given this premise, two quite interlocked subjects have been focussed in this paper: first of all, the classification of the dynamic behaviour of the almost one hundred sub-indices from which the HICP monthly estimates are compiled; secondly, the evaluation of whether it makes sense to compare across countries the series referred to a same elementary aggregate and to average them by groups of countries (e.g., to obtain the index for air transports in the euro zone or the EU).

By considering the monthly HICP series referred to 14 countries and nearly one hundred elementary aggregates in the period 2004-2008, it results that almost three quarter of the series are very homogeneous and show monthly profiles which are very near to the corresponding national HICP, with very flat and nearly time-linear patterns. The remaining sub-indices demonstrate more vivacity. Less than 10% of the series shows some seasonality and nearly 20% some irregular time-linear evolution. Strong and very strong variability is concentrated in less than 4% of the sub-indices. This picture should reflect the differences in the functioning of the numerous consumer markets covered by the HICP, and in many cases it really does: housing or clothing show in general a much higher variability than, for instance, furniture or tobacco.

Nevertheless, it is quite frequent to meet heterogeneous paths in the national series referred to a same aggregate: it can be said that about one half of the elementary aggregates present some heterogeneity across countries. Should this heterogeneity be generated by different approaches to the production of the estimates and not by underlying differences in the respective markets, it has to be recognised that the use of sub-indices may give in some cases small support or even wrong messages if used to build or evaluate industrial and market policies. The homogeneity and heterogeneity of national series are not in itself a signal of something that is going wrong in the process of HICP harmonisation. Nevertheless, the areas where it can be found a larger lack of homogeneity deserve a deeper investigation.

The cases of air transports and package holidays are paradigmatic of a lack of methodological harmony. For both markets it may be fruitful to work in order to induce common approaches to the estimates by NSIs: it appears clear enough that not all the differences among countries can be explained on the basis of local, market specific, features. Electricity, although it shows a strong differentiation in the behaviour of country series, is on the contrary a market where national regulatory frameworks are probably very important in determining price dynamics. It can be expected that some of these structural difference might fade away with time if European policies will spur the convergence of national markets.

At the opposite side, it must be recognised that a high homogeneity of the series referred to a same elementary aggregate does not necessarily mean that the sub-indices are perfectly harmonised and methodologically coherent. For example, a complex market like the one for new cars, where smooth series are common to all the countries, is probably treated by most member states in the same way, following list prices and thus avoiding all the perturbations in bargain prices which are induced by dealers' policies. Moreover, the frequent turnover of models, which can be seen as a potential source of variability, is often neutralised by the frequent use of overlapping technique for replacements.[36]

The main question at this point is: does the aggregation of sub-indices across countries make sense (for instance at EU or euro zone level)? This question is likely to have, at least in some cases, a negative answer, as it happens for air transport and package holidays. In order to avoid this, it seems reasonable to provide a common framework to manage sampling issues, especially because all these data are provided freely to any user and this implies that they are deemed to be representative and statistically founded. Furthermore, a great deal of consumer markets is the object of explicit EC policies, concerning competition, consumer protection and other sector specific issues.

Policy issues are, on the other side, a very important tool in the definition of the approaches adopted to design the proper methodological framework for estimating price changes. It appears in fact fruitful to base them on a sound analysis of market mechanisms, on supply and demand side and on the expected impact of policies.

As concerns the methods adopted for this analysis, they can be enriched with other sets of indicators which may be deemed to be useful to better discriminate the behaviour of the indices. This enrichment may be obtained by means of an increase in the number of indicators, exploiting more thoroughly the tools offered, for instance, by time series analysis.

[36] See De Gregorio (2010).

References

Allen R.G.D. (1975), *Index Numbers in Theory and Practice*. MacMillan, London.

Anitori P., De Gregorio C. (2004), "A proposal of classification of wholesale trade enterprises on the base of structural and performance indicators". In Bock, H.H., Chiodi M., Mineo A. (eds.) *Advances in Multivariate Data Analysis*, Springer, Berlin, 169-180.

Balk B.M. (2008), *Price and quantity index numbers. Models for measuring aggregate change and difference*, Cambridge University Press, Cambridge.

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

Brunetti A., De Gregorio C., Monducci R., Nuccitelli A., Polidoro F., Ricci R. (2004), "Official CPI and users: recent developments in ISTAT strategy", paper presented at the *OECD Seminar: "Inflation Measures: Too High - Too Low - Internationally Comparable?"*, June 21-22, Paris.

Caraballo M.A., Dabùs C., Usabiaga C. (2006), "Relative prices and inflation: new evidence from different inflationary contexts", *Applied Economics*, Vol. 38, issue 16, 1931-44

CSO. *See* Central Statistical Office of Ireland.

CSO (2003), "Consumer Price Index. Package Holidays. The seasonal baskets approach". *Information notice*, May.

Dalén J. (1998), "Studies in the comparability of consumer price indices", *International Statistical Review*, 66, 1, 83-113.

Dalén J. (2001), "Statistical targets for price indexes in dynamic universes", paper presented at the *Sixth Meeting of the International Working Group on Price Indices*, April 2-6, Canberra, Australia.

De Gregorio C. (2010), "List price and local price collection: are they really distinct alternatives?", *Rivista di Statistica Ufficiale*, n.1, 35-52.

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, p. 34.

EC. *See* Commission of the European Communities.

EC (1999), *Report on the Implementation of Directive 90/314/EEC on Package Travel and Holiday Tours in the Domestic Legislation of EC Member States.* SEC(1999) 1800. http://ec.europa.eu/consumers/cons_int/safe_shop/pack_trav/pack_trav02_en.pdf.

EC (2006a), *Developments in car retailing and after-sales markets under Regulation N° 1400/2002*. http://ec.europa.eu/competition/sectors/motor_vehicles/documents/retailing.html.

EC (2007a), *DG Competition report on Energy sector inquiry*. http://ec.europa.eu/competition/sectors/energy/inquiry/full_report_part1.pdf.

EC (2007b), *Working document on the Council Directive 90/314/EEC of 13 June 1990 on package travel, package holidays and package tours*. http://ec.europa.eu/consumers/rights/commission_working_document_final26-07-2007.pdf.

EC (2008a), *Analyses of the European air transport market. Annual Report 2007* http://ec.europa.eu/competition/sectors/motor_vehicles/documents/evaluation_report_en.pdf.

EC (2008b), *Commission evaluation report on the operation of regulation (ec) n° 1400/2002 concerning motor vehicle distribution and servicing*.

EC (2009), A *better functioning food supply chain in Europe*. Communication from the Commission to the European Parliament, the Council, the European economic and social committee and the Committee of the regions, COM(2009) 591, provisional version, http://ec.europa.eu/economy_finance/publications/publication16061_en.pdf.

EUROSTAT (2001), *Compendium of HICP reference documents*. European Communities, Luxembourg, 354.

EUROSTAT (2009), *Consumer prices research. An experimental analysis into the measurement of indicative price levels for consumer products*. Directorate D: Economic and Regional Statistics Unit D-4: Price statistics, http://epp.eurostat.ec.europa.eu/portal/page/portal/hicp/documents/Tab/Tab/04_METH_CPR_-_FEB_2009_WEB_0.pdf.

Fielding D., Mizen P. (2001), "The relationship between price dispersion and inflation. A reassessment", *European University Institute Working Paper ECO*, No. 10.

Good D.H., Sickles R.C., Weiher J.C. (2008), "A Hedonic Price Index for Airline Travel", *Review of Income and Wealth*, 54:3, 438-465

ISTAT (2005), *Rapporto annuale. La situazione del Paese nel 2004,* Roma.

Lent J., Dorfman H. (2005), "A transaction price index for air travel", *Monthly Labor Review*, June, 16-31

Nautz D., Scharff J. (2006), "Inflation and relative price variability in the euro area: evidence from a panel threshold model", *Deutsche Bundesbank Discussion paper*, series 1: Economic studies, No. 14, 36.

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

Ribe M. (2000), "Price index, statistical inference and harmonisation", Paper presented at the *Fourth International Conference on Methodological Issues in Official Statistics*, October 12-13, Stockholm.

Roger S. (1998), "Core Inflation: Concepts, Uses and Measurement", *Reserve Bank of New Zealand Discussion Paper,* No. G98/9

Ward J.H. (1963), "Hierarchical grouping to optimize an objective function", *Journal of the American Statistical Association*, No.58, 236-44

Wynne M.A. (1999), "Core inflation: a review of some conceptual issues". *European Central Bank Working paper*, No.5, May, 27.

Methodological aspects and empirical evidence in the use of administrative sources to estimate price dynamics in external trade [1]

Paola Anitori,[2] Carlo De Gregorio[3]

Abstract

This paper explores the possibility to widen the current offer of unit value and price indicators in external trade throughout a more intensive and efficient use of the available administrative sources. In particular, it firstly overviews the opportunities offered by a fuller exploitation of custom records linkage to other sources, such as business registers. Secondly, it presents a suitable conceptual and methodological framework for the estimates of price indicators based on custom records regarding, on the one hand, the definition of the reference population and the target parameter and, on the other hand, a statistically founded method to supersede the lack of details on the characteristics of the products. A generalised method of univariate density estimation and clustering is proposed to estimate homogeneous market segments on the basis of the elementary unit value distribution. A basic version of such method is applied to a subset of Italian custom records referred to years 2005-2007 and it is used to simulate estimates of export and import price dynamics.

Keywords: Unit values, Import and export price indices, Non parametric density estimates, Univariate clustering, Market segments, Administrative sources.

1. Introduction

In the last decade the main international organizations have supported a more intensive use of administrative records for statistical purposes in order to reduce the burden on enterprises and the costs of sample surveys.[4] Administrative data, in fact, are usually very detailed and often easily accessed: if properly treated, they allow the national statistical institutes to meet the increasing demand for additional information. Care must be taken in reconciling them to the statistical needs, but they can be a valid tool to supplement more traditional statistics.

This work discusses the possibility of producing alternative indicators on the behaviour of the prices of goods traded on international markets by means of a more intensive use of

[1] The authors wish to thank Roberto Monducci for his useful comments to an earlier draft and the anonymous referees for their suggestions. The authors remain the sole responsible of the view expressed in this work, which do not necessarily reflect the point of view of ISTAT.

[2] Senior Researcher (ISTAT), anitori@istat.it.

[3] Senior Researcher (ISTAT), cadegreg@istat.it.

[4] See for instance UN (2003), Eurostat (2005), OECD (2008) and EC (2009): the latter in particular confirms and supports the strategic role of administrative sources in driving innovations and quality improvements in official statistics. This approach is clearly fostered also by the MEETS program (see the Decision n. 1297/2008/ec of the European Parliament and of the Council of 16 December 2008 on a Programme for the Modernisation of European Enterprise and Trade Statistics).

the available administrative data. Interesting opportunities are in fact at hand for more detailed analysis of unit value dynamics throughout the linkage of custom records to other statistical sources, such as business registers and other structural business statistics, in order to improve the offer of suitable indicators for competition analyses. More thoroughly, the paper focuses on the possibility of using custom data to estimate price indices for imported and exported goods, and on the appropriate conceptual and methodological framework which is needed for this task: on the one hand a proper statistical definition of both the reference population and the target parameter and, on the other hand, a statistically founded method to supersede the lack of details on the characteristics of the products.

Two main methodological challenges are tackled here. The first one regards the specification of a suitable conceptual framework for price indicators based on unit values: the issue is dealt with in par.2 where a solution has been tailored according to the most recent methodological achievements of the Harmonised Index of Consumer Prices (hereafter HICP) in terms of target population, target parameter and statistical design. The second challenge is about the tools to be used for a full exploitation of the information conveyed by custom data in order to supersede some of the main shortcomings of unit values: in par.3 we propose a method based on the estimation of the univariate density function of unit values and on the use of a clustering algorithm to partition the population in market segments, exploiting the qualitative information available in custom records. Par.4 discusses the empirical results of this approach applied to a large subset of Italian import and export data referred to three important groups of products – namely, textiles, footwear and machinery for general purposes – in the period 2005-2007. An evaluation of this approach and some suggestions for further research are finally drawn in the concluding remarks.

2. UVIs and XMPIs: standard theory and new perspectives

2.1 Preliminary remarks

In the economic field, foreign trade statistics are probably one of the best examples of a cost-effective way of producing information through administrative records. Custom data are quite exhaustive and updated and, once they are reconciled to the statistical needs, they provide most of the official information on the international position of a country. Nevertheless, with respect to certain areas of research, custom records have not been used to their full extent yet: new perspectives, for instance, are opened by linking business and trade statistics; others can be imagined in the field of competitiveness analysis.

With respect to the latter, one of the oldest debate in literature is about the use of custom data to produce measures of international price changes. Traditionally the problem of price competitiveness has been framed - and solutions have been proposed - within the macro-economic perspective of national accounts.[5] Nevertheless, the discussion has recently taken

[5] See for example EUROSTAT (2001b) and IMF (2009). EUROSTAT has undertaken for many years a series of actions to enable Member States to produce reliable and comparable price and volume measures of National Accounts aggregates. Commission Decision 98/715 clarified the principles to be applied in measuring prices and volumes for most products of the ESA95 and set up a research programme for the remaining issues. EUROSTAT (2001b) provided a Manual on price and volume measures in National Accounts to give more detailed guidelines in this field. The

a different turn due to the existence of a strong demand for more detailed information able, for example, to explain the latest economic turmoil: whenever drastic economic events occur, the attention is often focused primarily on the behaviour of price relatives and on their influence on the inflation mechanism across countries.[6]

By the same token, in some countries the debate is fixed on the opportunity to asses the "quality upgrading" of exported products in order to understand to what extent pure price competition is still the main strategy pursued by firms on foreign markets (Giovannetti, 2007). Beyond single experiences, though, what is emerging is that detailed demands imply detailed answers that do not have to be necessarily satisfied through the usual approach based on macro-founded indicators (such as traditional price and unit value indexes). More articulated contexts need to be built following, for instance, micro-economic approaches where custom data can be used to their maximum extent (firm-level single transactions) in order to reduce costs and take advantage of both their high level of detail and large availability.

Sample surveys, in fact, are usually quite expensive and consequently sample size is often relatively small: results, then, are expected to be reliable and sound only for the parameters for which the sample scheme has been designed. In the domain of price statistics, for instance, a limited sample size may assign too relevant a role to the management of non responses, replacements and quality adjustments that, on their part, may generate a bias in the estimates if not adequately controlled for within a well structured and coherent statistical design. On the contrary, the use of very large data bases (almost of a census nature such as those from custom records) to produce some measurement of "price" dynamics in a micro-economic frame must be compelled to a definition of the statistical domains consistent with the peculiarity of the data source (therefore, not necessarily identical to what is commonly adopted for a sample survey).

Recent innovations in price collection techniques - and the ensuing access to a larger amount of price quotes – are inducing a more extensive use of unit values (hereafter UV) for the estimation of elementary price indices[7] stimulating the need of providing further methodological approaches.[8] The international literature on the use of UVs as price measures in micro-economic contexts, though, is comparatively meagre especially from the external

---

Commission Decision 98/715 clarified the principles for the measurement of prices and volumes based on the results of the research programme. The Commission has thus identified the most appropriate estimation methods to be applied (A methods), the alternatives which may be used if the most appropriate methods cannot be implemented (B methods) and the methods which shall no longer be used by 2006 (C methods).In a National Accounts context the preferred measure of price changes is a proper price index estimated from a sample survey. Unit value indices are considered as a secondary choice (B) if price indices cannot be made available. In this respect Member States are also expected to produce innovations and methodological advancements on B methods in order to improve reliability of the estimates (EFC 2010). See also Von der Lippe (2007).

6 In this frame, for instance, the European Commission has recently promoted the monitoring of import and consumption price developments in the EU in order to measure the extent to which consumers eventually benefited from import price decreases (as in textiles and clothing), to quantify linkages between tariffs and consumer price inflation, to analyse possible explanations for the magnitude of these linkages, or to identify factors linking consumer prices to future changes in EU import protection policy. This strategy was outlined in EC (2006) where it is argued that trade liberalisation and more open trade regimes would benefit European citizens in several ways, including through lower prices. See on this issue also Francois et al. (2007) and Institute for International and Development Economics (2007).

7 We refer to improvements in consumer price collection through scanner data (De Haan et al. 1997; Fenwick 2001; Jain et al. 2001; Silver et al. 2002; Triplett 2001; Koskimäki et al. 2003) and other sources like e-commerce and administrative or private databases.

8 Francoise et al. (2007).

trade perspective. The reason is twofold. On the one hand, the use of foreign trade data at firm level is quite a recent achievement boosted by a wider availability of statistical business registers and the improvement of record linkage techniques between business and trade statistics. On the other hand, the use of such measures limited to macro-economic contexts has prevailed for a long time, leaving a more structured micro-economic debate to lag behind.

In Italy, the idea of a more efficient use of custom data stems from the debate about the quality of measurements in official statistics with regard to the economic performance of the country.[9] Economists and users have remarked their dissatisfaction for the lack of import prices indicators (MPIs) and criticized the fact that production price indexes on non-domestic markets (XPIs) are not taken into account in compiling national accounts estimates for deflation purposes, the current practice relating to foreign trade components being, in fact, centred on the use of unit value indices (hereafter UVIs).[10] The recurring objection, by far the dominant point of view, is that UVIs are not proper price indexes and will never be, but unfortunately no alternative solution has been proposed yet to compensate for the lack of MPIs.[11] The debate, clearly shaped within the traditional accounting context, is characterized by a direct opposition of pro and cons of the two categories of indicators.

It is not the purpose of this paper to go into the methodological and theoretical underpinnings of the discussion on deflation practices, for which international manuals already give their valid contribution, nor to insist on methodological elements characterising UVIs as opposed to export/import price indexes (XMPIs). On the contrary, what we intend to explore is the possibility to propose measurements of external trade price dynamics in a micro-oriented context based on a more intensive use of custom data at their finest level of detail (firm level), in order to build up a system of conceptually sound and coherent indicators to be made available for analytical purposes.

2.2 UVI components and segmentation

Given the data set of the micro data corresponding to the transactions taking place in year *y*, with regard to a specific flow (import or export), each generic element can be represented through the following vector:[12]

$$t = [V_t \ , Q_t \ ; c_t \ , CN8_t \ , f_t \ , m_t \ ] \qquad (1)$$

where *V* is the value and *Q* the corresponding traded quantities, *c* is the foreign country, CN8 is the eight-digit code of the official commodity classification[13], *f* is the firm identifier,

9 The methodology to derive UVIs has been recently reviewed by ISTAT on the basis of sounder techniques for outlier detection and correction (Anitori et al. 2008, 2010).

10 Banca d'Italia (2008), ISAE (2009). In Italy XPI were released in 2008 (base year 2005=100) but the series are too short to allow a reliable use for NA purposes. Moreover the lack of MPI would imply an asymmetry in the deflation of the import component of GDP when compared to export. This is the reason why NA decided to use UVIs for deflating the corresponding aggregates raising a never ending discussion on the real amount of the industrial value added.

11 For an earlier analysis of this issue see Siegel (1991).

12 Vector *t* reproduces the typical structure of a custom record.

13 The EU official commodity classification adopted in foreign trade statistics is the Combined Nomenclature, directly derived from the UN Harmonised System, where goods are classified up to 8-digit level of detail (around 9.500 codes).

and *m* is the month in which the transaction takes place. From each vector *t* the unit value of a single transaction is derived as follows:[14]

$$P_t = \frac{V_t}{Q_t} \tag{2}$$

Omitting for simplicity the suffixes concerning the country and the firm's code, for the reference month *m* the unit value UV_m^{CN8} of the generic product is equal to the total value of all the transactions of the domain in the observed month divided by the corresponding total quantities[15]:

$$UV_m^{CN8} = \frac{\sum_{\forall t|m_t=m} V_t^{CN8}}{\sum_{\forall t|m_t=m} Q_t^{CN8}} = \frac{\sum_{\forall t|m_t=m} \left(\frac{V_t^{CN8}}{Q_t^i} \right) Q_t^{CN8}}{\sum_{\forall t|m_t=m} Q_t^{CN8}} = \sum_{\forall t|m_t=m} P_t^{CN8} q_t^{CN8} \tag{3}$$

where $q_t^{CN8} = \frac{Q_t^{CN8}}{\sum_{\forall t|m_t=m} Q_t^{CN8}}$ are normalised quantity weights. The unit value index in the reference period *0* is equal to the ratio of the respective UVs derived from (3):

$$UVI_m^{CN8} = \frac{UV_m^{CN8}}{UV_0^{CN8}} = \frac{\sum_{\forall t|m_t=m} P_t^{CN8} q_t^{CN8}}{\sum_{\forall t|m_t=0} P_t^{CN8} q_t^{CN8}} \tag{4}$$

Following the Laspeyres formula, further aggregations of the CN8 headings are derived as value-weighted averages:

$$UVI_m = \frac{\sum_{CN8} UVI_m^{CN8} V_0^{CN8}}{\sum_{CN8} V_0^{CN8}} = \sum_{CN8} UVI_m^{CN8} v_0^{CN8} \tag{5}$$

where $v_0^{CN8} = \frac{V_0^{CN8}}{\sum_{CN8} V_0^{CN8}}$ are the normalised value weights.

Due to its custom-made nature, though, no further distinction of the items of a product is possible within each 8-digit heading. Hereafter we shall refer to the classification with CN followed by the number of digits.

[14] With reference to the single transaction we adopted the symbol *P*, which usually suggests a reference to a price, to denote the elementary unit value. Actually, the unit value defined in (2) can be interpreted as a price only if it is assumed that every transaction is referred to homogeneous items. In principle homogeneity cannot be ensured nor somehow assumed to be the rule. Hereafter we shall conventionally adopt the symbol *P* to represent the unit values computed within a presumably homogeneous context.

[15] Expression (3) holds if $Q_t^i > 0$, which can be easily assumed by default.

Assume now that, on the basis of some specific criterion, the set of the transactions referred to each single domain can be partitioned in K^{CN8} distinct and exhaustive segments defined according to some specified criteria.[16] Then expression (3) can be rewritten as follows:

$$UV_m^{CN8} = \sum_{k=1}^{K^{CN8}} \sum_{\forall t \in k | m_t = m} P_t^{CN8} q_t^{CN8} = \sum_k \overline{P}_m^{CN8,k} q_m^{CN8,k} \tag{6}$$

where $\overline{P}_m^{CN8,k}$ and $q_m^{CN8,k}$ are respectively the average UV and the normalised volume-weight of the *k*-th segment (*k=1,2,……K*) in month *m*. By adapting this result to expression (4) we obtain:

$$UVI_m^{CN8} = \frac{\sum_k \overline{P}_m^{CN8,k} q_m^{CN8,k}}{\sum_k \overline{P}_0^{CN8,k} q_0^{CN8,k}} = \frac{\sum_k \left(\frac{\overline{P}_m^{CN8,k}}{\overline{P}_0^{CN8,k}} \right) \overline{P}_0^{CN8,k} q_m^{CN8,k}}{\sum_k \overline{P}_0^{CN8,k} q_0^{CN8,k}} =$$

$$= \sum_k I_m^{CN8,k} v_0^{CN8,k} + \sum_k I_m^{CN8,k} v_0^{CN8,k} \left(\frac{q_m^{CN8,k} - q_0^{CN8,k}}{q_0^{CN8,k}} \right) = I_m^{CN8} + R_m^{CN8} \tag{7}$$

where $I_m^{CN8,k}$ and $v_0^{CN8,k}$ are respectively the UVI and the normalised value weight of the *k*-th subset in the base year.[17] The first term on the right side of (7) is the Laspeyres weighted average of the within-segment UVIs. The second term measures the weighted co-variation between the UVIs and the change in the market shares (in terms of volumes) of their respective segment: this component is by definition null if market shares do not change. It also represents that part of unit values' dynamics in a specified domain which depends on the composition effects among the segments.[18]

Further aggregations of the elementary domains might easily be obtained from (7):

$$UVI_m = \sum_{CN8} I_m^{CN8} v^{CN8} + \sum_{CN8} R_m^{CN8} v^{CN8} = I_m + R_m \tag{8}$$

The criteria actually adopted for partitioning are very important for the meaning of the results obtained from (7). If the segmentation criteria are held constant across all domains it is possible to compare aggregated UVIs dynamics by segment and to measure the

[16] Such criteria may have to do with any information contained in vector *t* or with any auxiliary variable obtained by linking custom data with other sources such as for instance a business register.

[17] Expression (7) only applies to the segments existing in the base period. Entirely new segments which eventually appear in the current month must be either mapped into one of the already existing segments or imputed.

[18] In the literature on price statistics this last term is known as "unit value bias" and it is referred to any comparison between a UVI and the embedded price index (Parniczky 1980; Balk 1998 2008; Jain at al. 2001; Silver et al. 2002; Bradley 2005; Silver 2007). We prefer to treat this term simply as a residual, in order to stress the fact that it marks a drawback in the use of unit values as price indices but not an intrinsic negative property of unit values *per sé*.

aggregated composition effect among segments; in other words, $I_m^{CN8,k}$ can be aggregated also by segment:

$$I_m^k = \sum_{CN8} I_m^{CN8,k} v^{CN8,k} \tag{9}$$

and similarly for R_m^k. This happens for instance when segments are selected according to some auxiliary variables contained in the available data set (i.e. size classes of transaction values) or defined by the linkage with other statistical sources, such as firm level data from business registers or business statistics (i.e.: number of employees, class of turnover, region, legal type, main economic activity etc.). In such cases, in fact, neither conceptual nor methodological issues emerge in computing (9) - and, consequently, R_m^k - and these indicators can be seen as an enrichment of the information regularly disseminated.

On the contrary, what is more demanding from a methodological point of view is to build indicators of import and export price dynamics. For this purpose, the partition of the transactions into segments must follow the characteristics of homogeneity of the products: in this case, the within component (*I*) could be envisaged as very close to a price index. If such partition in segments is also economically and conceptually founded, expression (7) can be used to derive both the price component in the UVIs dynamics and an the estimate of the residual (*R*), the result depending on the theoretical and conceptual approach adopted to define the index and on the accuracy with which the segments are defined. In the remainder of the paper we develop some steps in this direction providing both a generalised methodological framework and a relatively basic method of partitioning transactions into market segments, testing the results of an empirical application. In particular, the provision of a suitable statistical frame for price indices estimation from custom data will be discussed in par. 2.4.

2.3 UVIs and XMPIs in the standard theory

It has been clearly stated both in literature and in international manuals that UVIs cannot be considered as proper estimates of foreign trade price indexes due to conceptual and methodological differences that prevent the definition of a common statistical ground for the estimation of the population parameter.[19] The recognised "superiority" of a proper XMPIs relies essentially on the evidence that most of the information needed to build up

[19] See for example EUROSTAT (2001a), IMF(2009), UN(1981), and the contributions of Parniczky (1980), Bradley (2005), Silver (2007), Balk (2008). International manuals admit the adoption of UVI only if it is proved that the products which they refer to are sufficiently homogeneous and stable over time; in general, though, the practice is rather discouraged. "*UVIs (...) do not generally control for changes in the product mix within one item, leading to quality changes mistakenly included in the price component. Their coverage of products is generally complete, but even at the most detailed level of trade classification can often include a range of different products. Where the products within an item of the trade classification may appear to be homogeneous this may in reality not be the case as products of similar description may be of very different quality. It may be possible to construct more homogeneous UVIs if the country of origin (or destination) is also taken into account. UVIs are clearly unsuitable for products that are unique or change quickly in specification. (....) The use of UVIs for some products is clearly inappropriate. (....)For product groups that are sufficiently homogeneous over time, UVIs can (...) be considered B methods. The volatility of the UVIs should be examined as a test for suitability (...).*" (EUROSTAT 2001a, p. 51-2).

the estimates comes from a direct and purposive survey where some sources of potential bias can be controlled for *ex ante*. A UVI instead, being some sort of a by-product, does not guarantee the same quality standard despite some well known and appreciated advantages. Nevertheless, it has been repeatedly suggested the idea that a more sophisticated treatment of unit values - by means of a complete exploitation of the content of custom records - might offer a good alternative for the estimation of price indices.[20] The study of unit values behaviour within a highly detailed group of products is, in particular, focussed on as an important element in determining the feasibility of their use as price indicators. On the one hand, most of the attraction in the use of custom records relies on the potential exhaustiveness of the source. Custom data provide an almost full coverage of the transactions on which the target population of a foreign trade index should be based on, whilst survey data collections provide estimates usually based on relatively small samples whose size is conditioned by budget and burden constraints. Custom data also immediately account for the entry of new products (and the exit of old ones), making it possible a continuous and updated monitoring of the markets; direct surveys, instead, rely on intense quality adjustment and replacement techniques to deal with entries and exits. Furthermore, custom records provide a full coverage of the products traded (direct surveys generally regard only manufactured goods) and, finally, they include all the firms dealing with external markets, whilst XMPI surveys tend to exclude trading companies (especially at export).

On the other hand, the main drawbacks in the use of custom records derive from the fact that a direct enquiry to firms for some important details is precluded: the impossibility of checking the specific conditions related to the single transactions, known in literature as price determining characteristics (i.e: the terms of sales, the time of contract, the transaction quantity etc…), the evidence of misclassifications and compilation errors can undermine the accuracy of the estimates. In particular, the lack of details about the product for which a "price" must be obtained is the most serious shortcoming. Although commodity classifications are very detailed, the impossibility of identifying the characteristics of the items classified within a same elementary code does not allow to check if its internal structure remains stable over time. This implies that an increase/decrease in the UV may be due to unidentifiable non-price effects that impair the measurement of pure price changes (Balk 1998; Von der Lippe 2001; IMF 2009): in other words, the UV of a specific product code is potentially biased with respect to the correspondent price if no homogeneity of the items transacted within that code is assured - that is, if composition effects are at work.[21]

---

[20] IMF (2009, p. 85-7) suggests to put more efforts on stratification (geographical or by firm) and data validation (treatment of outliers). EUROSTAT (2001a, p. 51-2) traces a solution based on a combined use of UVIs and price indices (estimated through a survey) for goods with stable UVIs and on the estimate of adjustment factors to be applied to the rest of the products. EFC (2010, pag.38-39) suggests stronger efforts on the improvement of B-methods: "*In order to accelerate the availability of more reliable and comparable data, it was proposed to consider the B output methods as reference methods, and to include the results of A methods in satellite accounts, while pursuing the development of more robust and harmonised explicit quality adjustment methods. All efforts should now be focused on further specifying and implementing B output methods that are solid theoretically and that can be applied consistently by all Member States, in particular in the context of the revision of the European System of Accounts*.".

[21] More precisely, one or more of the following conditions must be respected in order to avoid any bias in the use of a UVI to estimate a price index: quantities must increase proportionally between the price reference and the price reporting periods; the prices in the reference period must be all equal; absence of correlation between quantity relatives and price relatives (Balk 2008, p. 74).

This is the central issue in the conceptual frame supporting the statistical design of a price survey: under specific and sometimes elaborated assumptions, all is based on the strict control of items' characteristics so that any change is somehow managed and eventually reflected on the estimates (matched-model approach).[22] The driving practical principle of the matched-model approach is that, with regard to the products traded by a single firm, no quality change must intervene to modify the item's characteristics in order for a price difference between two points in time to be a "pure" price change. This assumption guarantees a coherence of the measurement not only at firm level but also for aggregations across firms selling or buying the items belonging to the same product heading: even if the items are not identical across firms, the aggregation of the elementary price indices within each heading is not affected by non-price effects. However, since in practice the frequency of the shifts in the commodity characteristics is quite high - especially at import - several methods for quality adjustment are recommended by international manuals in order to maintain measurement coherence over time.[23].

Despite the fact that at some stage of the survey strategy, firms and broad product categories are selected with probabilistic criteria, the final selection of the items for which price quotes are collected is usually purposive (i.e., it is made together with some representative of the firms and chosen among the most traded over a long period): in this way the usual problems related to the classifications adopted in foreign trade statistics are overcome and the aim of a pure price index better defined. The fact that the final sample of items is not randomly chosen can generate biased estimates, unless it is assumed that firms' pricing policies regarding the selected items do not differ systematically from the rest of the items traded and that the firm does not change its available offer.

Through the matched-model approach the universe from which the survey is designed has its own structure defined, depending on the particular set of commodities, price determining characteristics, industries etc. the statistician is interested in. Several possible universes can be identified according to the aim of the index, but once the population is chosen the structure is clear. Changes in this structure may occur (i.e: new products coming in/old products going out) but adjustments are in principle always possible to preserve the original assumptions.

In this context, the very nature of custom data makes them not compatible with the matched model approach and, more generally, with any approach based on a strict control over items' characteristics: if no items' specification is possible, composition effects are almost impossible to be controlled for. Custom data do not allow like with like comparisons.

---

22 No changes in the data collection, in items specifications, in the reporting firm, in the country of origin/destination of the goods etc. are allowed. In other words, all the *price determining characteristics* of the good must be kept constant over time (see IMF 2009). "*Pricing matched commodities constrains the selection of commodities to a static Universe of commodities given by the intersection of the two sets of commodities existing in the two periods compared. This static universe by definition excludes both new commodities and disappearing commodities, and in both cases their price behavior is likely to diverge from that of the matched commodities. Price indices have to try to take account of the price behaviour of new and disappearing commodities so far as possible*." IMF (2009, p. 27).

23 Inappropriate or over adjustments in XMPIs can be a problem often resulting in smoother time series; unfortunately, the impact of quality adjustments on price estimates is very seldom assessed even if this information could be very important to evaluate the accuracy of the estimation.

2.4 Conceptual issues in a HICP perspective

It follows that in order to give a sound statistical meaning to tentative price measures based on custom records, some alternative statistical structure of the universe is needed. To define such a structure it might be fruitful to refer to the most recent developments in the methodological and legal framework of the European Harmonised Index of Consumer Prices (HICP)[24] where design based approaches have been largely discussed and formalised. The aim is pursued without renouncing to the exploitation of the richness of administrative sources to its maximum extent.

The central issue in the HICP methodological and legal framework is the identification of a structure of the target population based on the concept of consumption segment. Namely a "consumption segment by purpose" or "consumption segment" is *"(...) a set of transactions relating to product-offers which, on the grounds of common properties, are deemed to serve a common purpose, in the sense that they are marketed for predominant use in similar situations, can largely be described by a common specification and may be considered by consumers as equivalent.*".[25] Every universe can be partitioned into exhaustive consumption segments where each segment is characterised by clusters of transactions which are deemed to be homogeneous with regard to marketing targets, consumption purposes and product characteristics.

Consumption segments are entities relatively stable over time and represent the fixed objects to be followed in a Laspeyres perspective. The target parameter - a monthly Laspeyres index - is obtained as the ratio of two simulated consumption expenditure (Ribe 2000) consistently defined by mapping the transactions within each consumption segment in the weight reference-period onto the product-offers available in the same segment, respectively, in the price reference and price reporting periods. From an economic perspective, consumption segments define a structure in the target population which depends solely on the functioning of the market and on the purposes of consumers.

Although the HICP Regulation recognises that some ambiguities still concern the definition of consumption segments,[26] and although the definition of the mapping function remains an open issue, the above theoretical structure establishes a logical correspondence between pairs of objects in two periods making it possible to compare like with like; as a consequence, sampling, grossing up, percentage changes etc. from this frame make full sense from a statistical point of view, once the underlying logic is agreed upon. Another key point of this structure is the fact that the adoption of consumption segments fosters the strategic role of a case-by-case analysis of consumer markets (as it happens, for example, with the structure of supply and demand, the marketing approaches and the segmentations adopted by producers and dealers etc.).

[24] See EUROSTAT (2001b), EC Regulation n.1334/2007.

[25] EC Regulation 1334/2007. A product-offer identifies the object of each transaction between two parties, namely *"(...) a specified good or service that is offered for purchase at a stated price, in a specific outlet or by a specific provider, under specific terms of supply, and thus defines a unique entity at any one time"* (EC Regulation 1334/2007, art.1.1): the set of all the transactions, and consequently the product-offers they refer to, contribute to characterise the statistical target population.

[26] EC Regulation 1334/2007 explicitly recognises that *"(...) an ambiguity in this concept concerns the level of aggregation at which it is defined and applied"*. This structured framework for the definition of HICP statistical universe remains in fact a theoretical tool open to a range of different solutions, while methodological and empirical research is still needed in order to provide complete statistical designs. The consumption segments represent also the population used for replacements and substitutions as each of them can be seen as a self-standing sub-universe.

On these premises, it is interesting to explore the possibility to structure a reference universe for custom data in a similar fashion, defining economic objects as clusters of transactions, homogeneous according to some rules, and labelling them as market segments.[27] The partition in subgroups will obviously need to be exhaustive within the domains they refer to. According to a HICP-type of approach these market segments represent the fixed entities whose price level has to be followed over time. Once the structure is defined in a weight reference year it will be replicated on to the set of available transactions in any current period in order to give sense to the target index number. This will help to overcome the main uncertainty pending on any "price" measure obtained from custom records (mostly related to the fact that any reference to some clear target population never seem to be straightforward) giving a new relevance to the measurement of the target parameter.

Assuming the viewpoint of a firm dealing with external markets, a criterion is needed to split the sets of transactions relating to each single reference domain into market segments. For each separate flow we decided to identify as a single domain the set of all transactions relating to a specific CN8 heading traded with a certain geographical area (hereafter GEO)[28] in a specific time period. This level of detail was judged, in fact, most appropriate to give evidence of the existence of different markets (and implicitly to deal with different items in the same CN8 heading), the only caution being to ensure the presence of a sufficient number of observations within each domain.[29]

3. The partition of UVIs target population into market segments: a generalisation of the clustering approach

Consider a generic domain d of transactions, defined by a given pair of CN8 and GEO in a given period. Consider also a transformation of the associated unit values (2):[30]

$$x_t = G(P_t) \tag{10}$$

Assume now that the domain is partitioned in K market segments[31] and that the values of x_t^k ($k=1,2,...,K$) within each segment are the realisations of a random variable with

[27] Reference to market segments as opposed to the consumption segments defined for the HICP is due to the fact that in external trade frame the point of view is traditionally that of firms and not of consumers.

[28] We chose to adopt the same geographical breakdown used for ISTAT UVIs series (Istat, 2003) in order to enable comparisons with the official figures. The breakdown includes both single countries and *ad hoc* areas. The general rule allows to single out all the 27 countries of the EU area and for each continent (or subcontinent) countries for which foreign trade with Italy is particularly important or shows a growing trend. We acknowledge that in some cases it would have been more appropriate to tailor this breakdown on our specific investigation needs in order to improve the coverage in specific areas, this aspect being of paramount importance and a relevant issue for further investigations.

[29] It is possible to identify finer strata (i.e: firm level domains by CN8 and GEO) if the number of transactions included in each of them allowed it. Here information at firm level is maintained but it will be used once markets segments are identified. Note that the possibility to link custom data to business registers enables to consider groups of firms on the basis of characteristics like activity sector, size, region, etc.. which could be stratification variables themselves. The only constraint remains, obviously, the number of observations within each strata.

[30] G is usually a log-transformation given that the distributions of unit values are frequently asymmetric and positively skewed.

[31] A single domain could be represented, for instance, by exports of cotton T-shirts (CN8) to Germany in each month of 2007. Distinct market segments could be represented by the (unknown) quality score (very low, low, medium, high ecc.) of T-shirts presumably reflected in the level of the UVs.

unknown density function $f^k(x)$. Assume also that, due to the transformation (10), the functions f^k are all well-behaved, i.e. nearly symmetric around their mode $\bar{x}^k$ (with $\bar{x}^k \geq \bar{x}^{k-1}$) and with a relatively small standard deviation σ^k. The overall density function computed for any value *x* in the whole domain is derived from the weighted average of the *K* densities:

$$f(x) = \sum_k f^k(x) q^k \tag{11}$$

where q^k are the normalised volume weights of each market segment in the *k-th* domain. The objective is to provide an estimate $\hat{f}(x)$ of $f(x)$ and to derive from it a possible partition in market segments. Since both the true partition of the domain and the functions $f^k(x)$ are all unknown, $\hat{f}(x)$ can be estimated on the basis of the empirical distribution of the data collected in the target period: such an estimate can be obtained by means of non parametric methods - for instance, some type of univariate kernel method (fixed or variable) or by means of a nearest neighbour estimate or with other dedicated approaches.[32]

While the density functions $f^k(x)$ can be assumed to be roughly symmetric around their mode, $f(x)$ and its estimate $\hat{f}(x)$ are generally asymmetric with a plurality of local maxima.[33] The problem is to check for the local maxima of $\hat{f}(x)$, to examine the saddles around them and to assess the possibility to treat it as the mode of a function $\hat{f}^{\hat{k}}(x)$ referred to the estimated market segment $\hat{k}$ defined on the basis of the some measures of homogeneity and concentration of the unit values in the domain. The set $\{\bar{x}^{\hat{k}}\}$ of the $\hat{K}$ estimated local maxima selected from the analysis of $\hat{f}(x)$ is used as a set of market segment identifiers:[34] let $\pi^{\hat{k}}(x)$ be the probability that each observed *x* in the domain belongs to the generic segment $\hat{k}$.[35] The estimate of the average unit value for $\hat{k}$ in a given year is then obtained as:

---

32 Bean et al. (1980), Silvermann (1981, 1986), Izenman et al. (1991), Terrell et al. (1992), Jones et al. (1996).

33 It is well known that for density estimation a global window width kernel estimator may not perform well when the underlying empirical density features require different amounts of smoothing at different locations points (Wand et al. 1991). This is certainly likely to happen with UV or price distributions, usually characterised by log-normal densities and, a strongly positive skewness. One of the anonymous referees has rightly pointed out that "*the number of modes (and minimums) on a kernel-estimated function does depend on the scale (original or logarithmic) of the data*" with possible alterations of the original multi-modal structure of the data. We certainly agree on the need for further research in this direction and especially in the automated detection of false maxima, eventually determined by the overlapping tails of two contiguous $\hat{f}(x)$ and more likely to happen with log-transformed data.

34 The choice of the notation relating to *K* wants to stress the fact that the estimation of the market segments is the core methodological issue.

35 The selection of these segments might for example be obtained with a clustering algorithm (Pisani, 1993) and/or by a discriminant analysis. These approaches will be tested as further developments of the results exposed in this paper.

$$\overline{P}^{\hat{k}} = \frac{\sum_{t\in\hat{k}} x_t Q_t \pi_t^{\hat{k}}}{\sum_{t\in\hat{k}} Q_t \pi_t^{\hat{k}}} = \sum_{t\in\hat{k}} x_t q_t^{\hat{k}} \tag{12}$$

The monthly series of the average prices $\overline{P}_m^{\hat{k}}$ is accordingly derived and, as a consequence, the monthly price indices by segment can be compiled as follows:

$$I_m^{\hat{k}} = \frac{\overline{P}_m^{\hat{k}}}{\overline{P}^{\hat{k}}} \tag{13}$$

and the index for the whole domain is given by:

$$I_m = \sum_{\hat{k}} I_m^{\hat{k}} w^{\hat{k}} \tag{14}$$

where $w^{\hat{k}}$ are the normalised value weights of the $\hat{k}$-th segments. It is important to notice that the segmentation is domain-specific and that there is no possibility to compare segments of different domains.[36]

As an example, we report the outcome of a case study related to three distinct CN8 headings, each belonging to one of the groups of products we decided to focus on, namely Other textiles, Footwear and Machinery for general purposes. These groups represent part of the *core* of Italian trade although they have quite different characteristics: in particular, the first two typically address consumer markets whilst the latter concerns mainly investment goods tailored on industry needs.[37]

For each heading, two domains have been selected among those with a larger number of records according to the country of destination. Figure 1 plots the estimated density distributions $\hat{f}(x)$ of the log-transformed unit values (10) for each domain. Such estimates have been obtained by means of a uniform kernel of fixed radius.[38] It is possible to notice that within each CN8 the shape of $\hat{f}(x)$ changes as we consider different destinations. In the case of Other textiles, for instance, this is particularly evident: export UVs relating to country A have in fact a bi-modal distribution - corresponding to quite different UV levels (respectively 13.5 and 30 euro) - well separated by low density saddles; in this specific case, the delimitation of at least two segments appears feasible.

UVs referred to country B show, on the contrary, an unimodal density distribution (the mode being almost coincident with the lowest of country A) with some noise in the right tail:

---

36 This implies that expression (9) cannot be applied and that it is not possible to derive aggregate indices by segment.

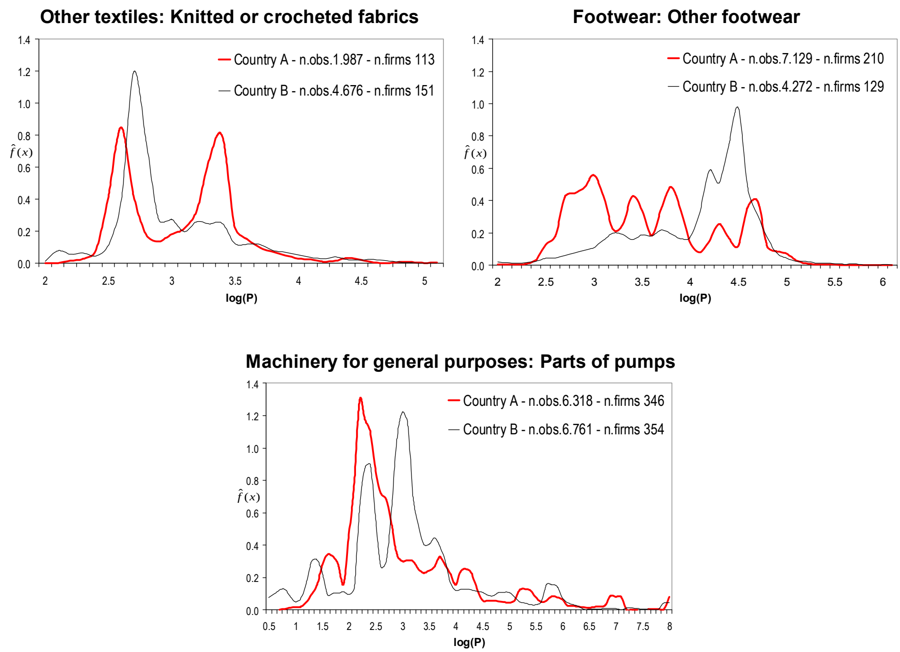
37 In 2005, at export, Other textile represented about one third of total Textiles (10% at import) , Footwear more than half of Leather goods (53% at import) and Machinery for general purposes is more than 30% of Total machinery and apparel (29% at import). The CN8 headings considered in the case study are: “Knitted or crocheted fabrics” (code 60041000) for Other textiles, “Other footwear” (code 64039996) for Footwear and “Parts of pumps” (code 84139100) for Machinery for general purposes. The groups correspond to the 3-digits of the EU Classification of Products by Activity (CPA) which is linked to the CN through a corresponding table. See, http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC.

38 See Silverman (1986). The observations have been weighted by the corresponding volumes and standardised, while the radius has been set to 1% of the standard deviation.

the isolation of at least one single segment appears here very reasonable. For Footwear, the example reports quite different shapes of $\hat{f}(x)$: in particular, the density referred to country A has at least four well separated local maxima, ranging from 20 to 110 euro. The one relating to country B is by far more concentrated on a high-priced segment (distribution around the mode reveals a lower kurtosis with a high-density saddle point on the left): in this specific case, a single segment would imply too large a range in UV levels (from 60 to 120 euro) and further efforts are needed to improve the discriminatory power of the method. As for Machinery, $\hat{f}(x)$ appears to be unimodal with respect to country A with a large positive tail, whilst two local maxima appear for the density function relating to country B (corresponding to a range from 11 to 20 euro), sharply separated by low density saddles. In both cases the segments appear well defined, although slightly asymmetric for country A.

The example suggests that further investigation are needed whenever the partition algorithm is not sufficiently discriminating; in such cases a more intensive use of firm level information and further investigations on the distribution of the variable may be necessary in order to improve the classification method and to avoid excessively time-consuming elaborations usually associated to the application of density estimation and cluster analysis to such large and complex data sets.

Figure 1 - Density function $\hat{f}(x)$ of the log-transformed export unit values by CN8 heading and country of destination. Year 2007



4. An application to a subset of custom data

4.1 Main features of the data base

Based on the general approach outlined in par. 3, in this paragraph we test a straightforward, time saving and effective method of partitioning each domain into market segments in order to compile an estimate of both the price dynamics and the residual composition effect seen in expression (7) above. For this purpose, we use custom data referred to all import and export transactions observed between 2005 and 2007 relating to the CN8 headings belonging to the groups of products singled out in the example above (see Figure 1).

Table 1 outlines the main structural characteristics of the available information in 2005 (base year of the forthcoming index).[39] Several hundred thousands of records are available in the CPA groups under scrutiny and, as expected, the number of transactions at export nearly doubles the corresponding elementary data at import. The amount of domains - defined by crossing the CN8 headings and more than fifty GEO modalities - ranges from 3,000 to more than 9,000 at export and from 2,000 to more than 5,000 at import, showing in both cases larger absolute values in textiles and machinery. In footwear, though, the amount of transactions is the most relevant despite the lower number of products and firms. Within each CPA group, trade is highly concentrated on few CN8 and firms, giving evidence of a certain number of low-value transactions, occasional operators and low volumes.[40]

Table 1 - Main characteristics of custom data set, by type of flow and CPA group. Year 2005

MAIN GROUPS OF PRODUCTS (CPA)	No of products traded (CN8)	No of trans-actions	Firms		Domains			
					CN8-GEO		Firm-CN8-GEO	
			N.	Trans-actions *(avg)*	N.	Trans-actions *(avg)*	N.	Trans-actions *(avg)*
			EXPORT					
Other textiles	301	481.874	10.800	*44,6*	9.583	*50,3*	75.636	*6,4*
Footwear	73	892.016	6.869	*129,9*	3.245	*274,9*	100.607	*8,9*
Machinery for g.p.	224	823.958	13.353	*61,7*	7.223	*114,1*	118.868	*6,9*
			IMPORT					
Other textiles	302	247.560	10.099	*24,5*	5.381	*46,0*	56.997	*4,3*
Footwear	73	227.988	6.149	*37,1*	1.913	*119,2*	25.666	*8,9*
Machinery for g.p.	223	413.892	12.637	*32,8*	4.186	*98,9*	5.183	*7,8*

Source: elaboration based on custom data

Although the variability of UVs within a single CN8 is generally extremely high,[41] the auxiliary information included in custom data may help to reduce it drastically. The results

[39] These characteristics remain stable also in the following two years.

[40] As example we refer to the case of Other textiles at export where about one third of the products sold with more than 1.000 outward transactions account for 87,8% of total transactions and 85,9% of values. Figures by GEO show that one fourth of the areas accounts for 95% of the total value of transactions. Finally, only 10 to 15% of firms have more than 100 transactions: those firms account, though, for 70-90% of the total value traded depending on the CPA group.

[41] In three headings out of four the UV have a coefficient of variation higher than 100%; in a quarter of the headings it is higher than 500%.

of the analysis of variance reported below in table 2 show that, within each CN8 most of the UV variability is accounted for by firm and GEO (the F-test is highly significant for most products in both cases). At a 5% significance level, firm effect is considerable for more than 90% of the headings both at import and at export, whilst GEO effects appear in general more significant at export or, specifically, for imported footwear. Interaction effects on the contrary are significant in less then 20% of the headings (below 10% for textiles).

Due to the high coverage and to the explanatory power of the firm level information, these results seem to encourage the search for solutions based on selective designs where original domains characterised by a low number of observations may be aggregated, for instance, by properly defined geographical areas, without biasing the estimates.

Table 2 - Analysis of variance on the distribution of unit values, by type of flow and CPA group. Years 2005-2007 *(number of CN8; % share of CN8 headings with 5% significant F-test)*

YEARS	Other textiles				Footwear				Machinery for g.p.			
	No of CN8	Firm effect	GEO effect	Inter-action	No of CN8	Firm effect	GEO effect	Inter-action	No of CN8	Firm effect	GEO effect	Inter-action
					EXPORT							
2005	106	96,2	65,1	36,8	67	100,0	67,2	49,3	106	92,5	66,0	41,5
2006	112	98,2	60,7	32,1	69	100,0	68,1	60,8	106	94,3	67,9	45,3
2007	108	94,4	69,4	33,3	68	98,5	70,6	58,8	108	87,0	63,9	41,6
					IMPORT							
2005	69	100,0	33,3	2,9	36	100,0	63,8	19,4	80	91,3	41,3	18,8
2006	72	95,8	27,7	6,9	40	100,0	67,5	20,0	86	93,0	44,2	17,4
2007	75	98,7	42,7	9,3	41	100,0	68,3	9,8	89	92,1	44,9	13,5

Note: The test has been limited to the CN8 headings with at least 1.000 observations and to the domains (firm-CN8-GEO) with at least 20 observations

4.2 A simplified approach for identifying market segments

The general method outlined in par. 3 provides a partition of each domain in several market segments on the basis of the underlying distribution of unit values (eventually transformed according to expression (10). It relies, in particular, on the estimation of the domain's density function. The use of unit value densities implies the assumption that consistent differences among unit values within a single domain are an indirect evidence of the existence of several market segments. Although this is not necessarily true for all domains, we deemed it conceivable in general, due to the fact that pricing policies on foreign markets are usually constrained by international competition; in particular, with regard to a specific product sold in (or bought from) a specific country by several firms, we reckon not unlikely to assume that differences in price levels may reflect differences in the product offers, the latter being supposed on their part to intercept different market targets. The more homogeneous the product heading, the truer; more generally, we believe that if the amount of observations at firm level in each domain is very large this assumption can be still reasonable and hold.

Nevertheless, from a practical point of view, care must be taken to ensure an accurate screening of the characteristics of the domains' density distributions and, consequently, of the validation of the statistical procedures adopted to classify observations into segments.

The method of partitioning proposed here is based on a deterministic linear algorithm applied on annual data with the aim to obtain estimates, at CPA 3-digit level, of the UVI components *I* and *R* as defined in (7), that is the price index referred to the specific partition in market segments (i.e. the average within-segments component) and the additive residual measuring the composition effects between market segments in a same domain. For each domain the method singles out one subset of homogeneous transactions as follows.

Consider the distribution of the transformed UVs (10) associated to the transactions of each domain *d* as defined in par. 3.[42] Call x^q the *q*-th percentile of their distribution.[43] Given an integer $j \geq 1$ relatively small, we consider the estimate of the density distribution of transformed UVs as a function of the *j* inter-percentile relative range (suffix relating to the domain *d* is here suppressed):

$$\hat{f}\left(x^q\right) = \lambda - \frac{x^q - x^{q-j}}{(x^q + x^{q-j})/2} \tag{15}$$

where *q=j,…,100* and λ is a positive scale parameter.[44] Select the percentile $x^{q'}$ that maximises the function (4.1) over a predefined threshold $(\lambda - \mu)$:[45]

$$x^{q'} \mid \hat{f}\left(x^{q'}\right) = \max_{x^q}\left[\hat{f}\left(x^q\right)\right] \quad and \quad \hat{f}\left(x^{q'}\right) > \lambda - \mu \tag{16}$$

When $x^{q'}$ exists, the centre of the interval ($x^{q'-j}$, $x^{q'}$) is consequently considered as a major accumulation point of the domain:[46]

$$x' = \frac{\left(x^{q'} - x^{q'-j}\right)}{2} \tag{17}$$

We use (4.3) as the centre of a price interval $\left(x'_L , x'_U\right)$ which marks the borders of the main market segment in *d*, labelled with $\hat{k}^d$.[47] Several alternatives have been tested concerning the criteria to derive the above interval. Two classes of solutions have been considered in our empirical work: symmetric intervals and asymmetric intervals with constrained UV's variability.[48]

---

42 Only the domains with at least 200 transactions have been considered in the analysis. Nevertheless, the weights of the excluded domains are used for upper level aggregations of the indexes.

43 In this case we considered bounded quantity weighted percentiles. In any case, $x^0 = \min(x)$ and $x^{100} = \max(x)$.

44 The *j* inter-percentile relative range is the ratio between the difference between the *q*-th and the *(q-j)*-th percentiles and their average: it varies between 0 and 1. In our applications we set *j=4* on the basis of empirical evaluations. The smaller *j* the higher the probability of identifying false accumulation points; a higher *j* increases the smoothness of the density function. As a matter of fact *j* plays the role of the bandwidth in kernel density estimation (Silverman 1986).

45 In the elaborations, μ was set to 0.1.

46 In principle the occurrence of more minimums is possible, especially for $\hat{f}(x) = 0$: in this cases the main accumulation point is chosen taking into account the values of function $\hat{f}(x)$ in an extended interval.

47 The notation stresses the fact that the main segment *k* is domain-specific.

48 The bounds of symmetric intervals are expressed in function of x' : $x'_L = x'(1-\vartheta)$ and $x'_H = x'(1+\vartheta)$, with $1 > \vartheta > 0$. The bound of asymmetric intervals are determined adding to segment *S* all the transactions with a lower distance from x' until a variability threshold is reached.

Given this interval, all transactions belonging to the segment $\hat{k}^d$ are determined as follows:

$$if \quad x_t \Big| \; x'_L \leq x_t \leq x'_U \quad \Rightarrow \quad t \in \hat{k}^d$$

4.3 Price index estimates and chaining

The price index of domain *d* is estimated using the set of all the transactions belonging to the market segment $\hat{k}^d$ as determined in par. 4.2. With reference to a base year *y* and to a reporting month *m*, the average price in segment $\hat{k}^d$ for firm *f* in month *m* is estimated as follows:[49]

$$\hat{\bar{P}}_{m;y}^{d,f} = E_{\forall t \in \hat{k}^d | m_t = m}\left(P_t\right) \tag{18}$$

The corresponding elementary price index at time *m* is the ratio of the average prices compiled in (18) in the current period *m* and in the base year *y*:

$$\hat{I}_{m;y}^{d,f} = \frac{\hat{\bar{P}}_{m;y}^{d,f}}{\hat{\bar{P}}_{y;y}^{d,f}} \tag{19}$$

Further aggregations of these indices are obtained as weighted averages of (19), where weights are proportional to the corresponding values in the base year; in particular, weights include also the transactions not directly used for the computation of the indices – i.e. those not belonging to segment $\hat{k}^d$. For instance, indexes by CN8 are computed as follows:

$$\hat{I}_{m;y}^{CN8} = \sum_{GEO} v_y^{CN8,GEO} \sum_f \hat{I}_{m;y}^{d,f} v_y^{d,f} \,, \tag{20}$$

and similar formulas are applied to calculate indices at CPA level and/or larger geographical areas.

By chaining, the series referred to month *m* in year *y+n* on a common reference base *y* is obtained as:[50]

$$\hat{I}_{y+n,m;y} = \hat{I}_{y+n,m;y} \prod_{q=1}^{n} \hat{\hat{I}}_{y+q;y+q-1} \tag{21}$$

[49] The mean used to aggregate price quotes may assume various possible forms, for instance a geometric or arithmetic mean, weighted or unweighted. We opted for a weighted arithmetic mean.

[50] For simplicity we have put the bases of the indices equal to 1 instead of the usual 100. We opted here for a chain link expressed by the index of year *y+q* based on the previous year: for the estimation of the link the transactions in year *y+q* are mapped into the segments estimated for year *y+q-1*.

4.4 Coverage and empirical outcomes

In 2005[51] more than 600 thousands observations and 2.300 domains for export and 260 thousands observations and 1.000 domains for import were used for the estimation of the indices with the “percentile method” discussed above (Table 3). Coverage is higher for EU flows, where about 30% of total custom records were used due to the fact that intra-community trade is traditionally more intense (in fact, the number of transactions per domain is higher). A more appropriate geographical breakdown referred to Extra-EU trade would probably improve coverage in the relating flows also in terms of number of domains.[52] It can also be observed that for EU coverage is lower in terms of values: this is probably due to the fact that the method tends to select “low UV” segments. Import and export indexes have been estimated by domain and then aggregated by CPA group and main area (EU27, EMU and Rest of the World) according to the formulas described in par.4.3 in order to compare them with official UVIs. In particular, indices produced by means of the “percentile method” (labelled as “NEW” in the tables and figures below) were compared both with official UVIs and XPIs at export, but only with official UVIs at import due to the lack of the corresponding official MPIs.

Table 3 - Domains and observations used for index estimation by flow, product and area. Year 2005 *(no. of domains and transactions; % share on total transactions and values)*

CPA GROUPS	EU27				Rest of the World			
	Domains	Observations			Domains	Observations		
		No.	% on total trans.	% on total value		No.	% on total trans.	% on total value
			EXPORT					
Other textiles	769	123.854	33,1	26,9	222	17.506	16,3	20,0
Footwear	626	231.953	35,4	21,1	272	51.594	21,8	19,9
Machinery for g.p.	899	144.708	24,1	19,5	313	24.168	10,8	10,5
Total	2.294	500.515	30,7	21,1	807	93.268	16,4	14,0
			IMPORT					
Other textiles	320	54.897	34,1	17,0	166	16.052	18,5	22,8
Footwear	186	54.986	36,1	27,8	109	20.238	26,8	32,3
Machinery for g.p.	437	100.337	30,8	17,0	157	13.853	15,7	24,9
Total	943	210.220	32,9	19,4	432	50.143	20,0	26,9

As expected, the NEW series appear in general smoother than official UVIs (Table 4), the latter showing in fact a stronger upward trend as well as a wider infra-annual variability.

In all the CPA groups under scrutiny the annual inflation measured by UVIs is several points greater than that measured by NEWs. The spread between the two indicators (which can be considered as a proxy of the entity of the residual component R in par. 2.1) is always positive despite some distinctions across industries and areas.

Other textile NEW indices registered a very sluggish growth both at import and export whilst UVIs revealed a faster growth especially toward non-EU countries. A similar outcome is found for machinery.

[51] Almost an identical picture was found with reference to the rest of the period (years 2006 and 2007).

[52] In fact only 3.316 domains, out of the 20 thousands found in the universe of export transactions in year 2005, respected the condition of having at least 200 observations; moreover, a group of 200 domains were also excluded due to the impossibility to clearly identify a high density cluster.

The case of Footwear appears emblematic: EU and non-EU price dynamics as measured by NEW series appear similar and regular, whilst UVIs record a faster growth for extra-community trade, especially at export. In general, there is evidence of a residual component *R* being more impressive for trade outside EU, where competition is usually based on wider product offers and firms' pricing policies are exposed to currency fluctuations. Table 4 also reports two variability indicators confirming (if necessary) the higher volatility of UVIs.

Figures 1 and 2 in the Appendix show the peculiarity of Footwear: UVIs in both flows have a cyclical behaviour associated to a strong upward trend whilst, on the contrary, NEW indices seem to reflect cyclic movements only partially; the average yearly rate of change registered by UVIs, nearly 10% in the whole period, is probably the result of a change in the relative weight of market segments, seasonal effects not controlled for and the evidence of the role of the component *R*.

Table 4 - Official UVIs and NEW indices, by flow, product and area. Years 2005-2007 *(% average annual change of monthly indexes)*

	Other textiles				Footwear				Machinery for g.p.			
	EU27		Rest of the World		EU27		Rest of the World		EU27		Rest of the World	
	NEW	UVI	NEW	UVI	NEW	UVI	NEW	UVI	NEW	UVI	NEW	UVI
EXPORT												
% change (a)												
2006/2005	0,7	1,1	0,1	4,1	1,2	6,2	1,2	9,2	0,6	4,7	2,2	5,8
2007/2006	0,3	1,7	-0,5	1,3	1,2	5,7	1,3	10,0	1,6	3,1	2,6	4,9
2007/2005	0,9	2,8	-0,4	5,5	2,4	12,3	2,4	20,0	2,2	7,9	4,8	10,9
Variabilit indicators (b)												
CV (%)	0,6	1,6	0,7	2,7	1,6	6,3	1,7	9,9	1,2	3,6	2,6	6,2
Quartile range	0,8	2,3	0,8	4,8	2,4	9,8	2,6	14,9	2,3	6,0	4,7	8,9
IMPORT												
% change (a)												
2006/2005	0,7	2,1	0,1	3,5	1,6	6,4	1,2	6,1	1,0	4,7	-1,2	8,1
2007/2006	0,0	2,2	-0,6	2,2	-1,0	2,7	-0,6	4,1	1,1	1,5	0,0	0,1
2007/2005	0,7	4,4	-0,6	5,8	0,6	9,3	0,6	10,4	2,1	6,2	-1,2	8,2
Variabilit indicators (b)												
CV (%)	0,7	2,1	1,0	2,9	1,3	5,9	1,6	5,1	1,1	3,2	1,0	4,4
Quartile range	0,7	3,5	1,4	4,5	1,8	11,5	2,2	7,6	1,5	4,7	0,9	7,0

Source: ISTAT for UVIX series

(a) Compiled on the yearly averages of the indices.

(b) Compiled on the series of the monthly indexes in the whole period.

A similar behaviour can be found for Machinery with reference to extra-EU trade in both flows; in this case, more attention is probably needed to define market segments since the sector is characterised by a strong heterogeneity. The comparison between the official XPIs based on a sample survey and the NEW index at export (see Table 5 below and Figure 3 in the Appendix) shows, as expected, smaller and non-systematic differences in the dynamics of the two indicators in almost all the CPA groups of products, as compared to the previous comparison between NEW indices and the official UVIs. Note that the best approximation between the two series is shown by Other textiles, especially at the Euro-zone level where the indices dynamics looks quite similar. As opposed, XPIs referred to Machinery but regarding the same area show a much stronger and growing trend, whilst greater similarity is found with regard to the Rest of world.

Table 5 - XPI and NEW indices by product and area of destination. Years 2005-2007 *(% average annual change of monthly indexes; values)*

INDICATOR	Other textiles				Footwear				Machinery for g.p.			
	EMU12		Rest of the World		EMU 12		Rest of the World		EMU12		Rest of the World	
	NEW	XPI	NEW	XPI	NEW	XPI	NEW	XPI	NEW	XPI	NEW	XPI
Indexes												
2005-2006	0,8	0,6	0,2	0,7	1,2	2,5	1,0	3,1	0,3	4,0	2,1	2,7
2006-2007	0,3	-0,3	-0,2	1,3	1,5	2,0	1,5	0,7	1,8	5,0	2,2	1,8
2005-2007	1,1	0,3	0,0	2,0	2,8	4,5	2,4	3,7	2,1	9,2	4,2	4,5
CV (a) (b)	0,6	0,6	0,5	0,9	1,8	2,0	1,5	1,8	1,1	3,8	2,1	1,9
Quartile range	0,9	0,7	0,6	1,3	2,6	4,3	2,0	3,6	2,2	8,6	4,1	3,7

Source: ISTAT for XPIs series
Note: (a) Indicator compiled on the series of the index levels (36 observations)
(b) Percentage

5. Concluding remarks

This work discusses the possibility of producing, on a regular basis, a set of price indicators of the goods traded on international markets by means of a more intensive and efficient use of the available administrative data. Two perspectives, in particular, have been discussed. The first one relates to an integrated use of custom data and other statistical sources - such as business registers - as a cost effective way to widen the current offer of indicators of unit values dynamics. The possibility of partitioning transactions into segments according to some criterion defined on the basis of auxiliary variables relating to firms characteristics and behaviour, can help to split UVI dynamics in a “within-segment” component and in a residual “between-segment” component. Useful information for competition analysis and strategic positioning on international markets can be thus derived in a relatively easy and low cost way.

The second perspective extends this logic to the case of partitions based on the degree of homogeneity of the goods. Here, though, more relevant methodological issues are raised: firstly, there is the need to tackle the impossibility to appreciate homogeneity in the available data due to the lack of detailed information on items’ characteristics even at the finest levels of the classification; secondly, the huge amount of information provided by custom records compels to tailor a suitable conceptual and methodological framework for the estimation of any price measure of an economic significance. In this respect, the challenge would concern the definition of the reference population, the statistical design and the target parameter.

With respect to this latter point, the most recent achievements in HICP methodology appear particularly suitable for this task, in particular with regard to the role played by the concept of “consumption segments” in the statistical design: in an external trade context, market segments are thus identified as the statistical objects to be followed over time in order to provide the reference population of transactions with a structure. The proposed approach to define market segments is based on univariate density functions’ estimation and clustering so to single out, within each domain, groups of transactions homogeneous with regard to elementary UVs behaviour.

The empirical results shown in par. 4 and in the Appendix appear quite encouraging, although obtained by means of a basic version of the method outlined in par. 3. The estimated dynamics appears reasonably similar to the official XPI and the long term differences among the two indicators are relatively small, even if the estimated series show a higher short term variability due to the entropy deriving from the larger amount of data used for their estimation. On the contrary, the new series often diverge from UVIs as a consequence of the composition effects taking place among market segments.

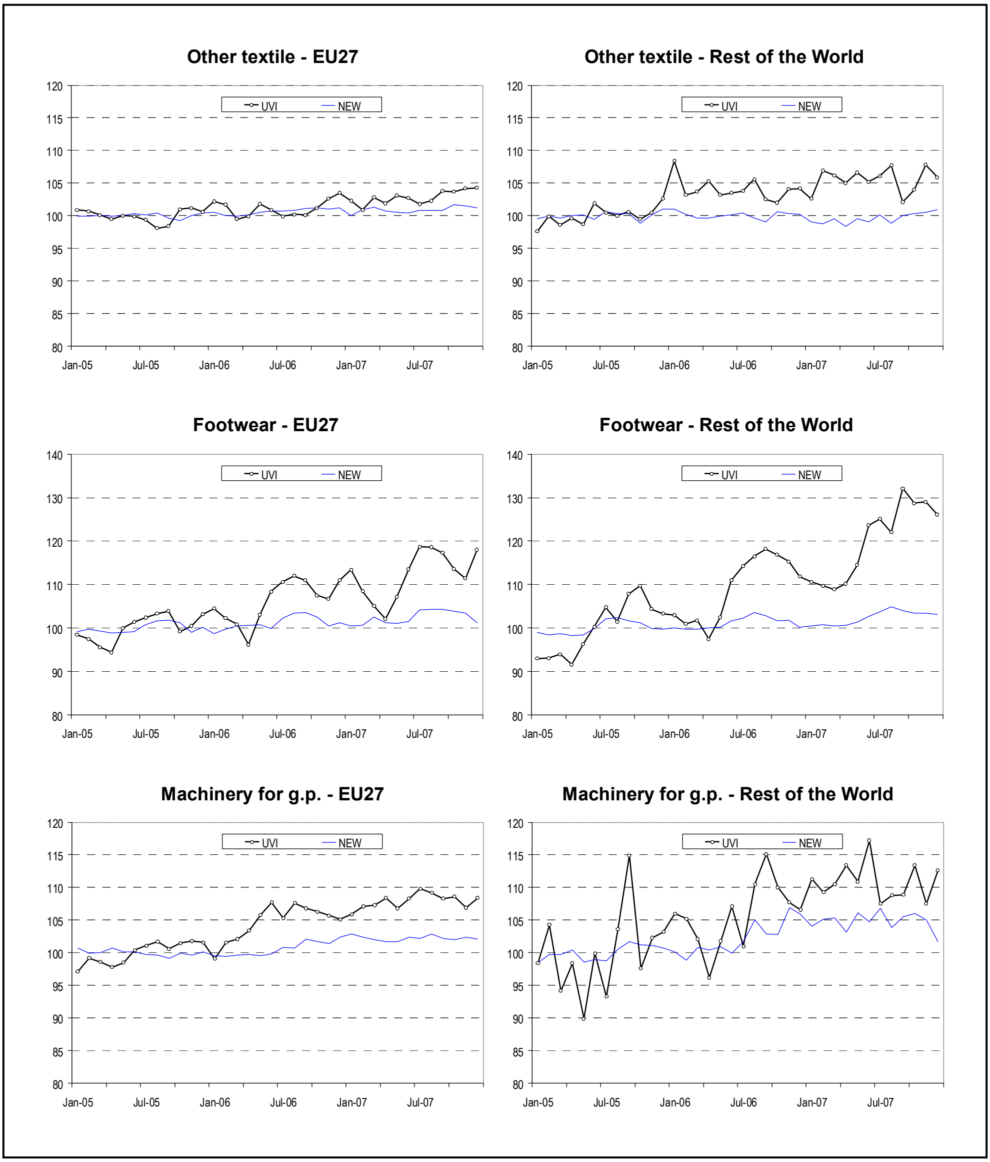
Further research is obviously needed to improve the general statistical frame and, in particular, the method adopted for market segments partitioning. The choice of the target time period used to estimate market segments is one of the key issues: the annual segmentation adopted in par. 4 might result vulnerable should inflationary pressures or high short term UV variability emerge, since it could induce smoothing and consequently a zero-change bias into the estimates. Alternative solutions are actually under scrutiny and could be found in the adoption of a shorter-term rolling target periods (such as for instance a moving quarterly period).

The estimation of density functions and the clustering algorithm also deserve further studies; the generalised method in par.3 is promising but more sophisticated methods can be found to optimize and stabilise the resulting partition of the reference population; the main challenge seems to lay in non-parametric density estimation techniques, since the choice of the degree of smoothing is a key element also for the successive application of the clustering method. Given the extremely large number of domains and observations, one more issue has to do with the research of automatic solutions able to select for each domain the best smoothing parameter. The adoption of automatic clustering algorithm also needs further investigations in two respects: the choice of a time saving method and the control of the borders of the clusters. As for this latter point, the adoption of probabilistic clustering appears quite promising.

Finally, sensitivity analysis can be applied to test the robustness of the estimates by means of simulation techniques. Robustness in fact is a key element to assess the feasibility of the methodology especially for dissemination purposes.

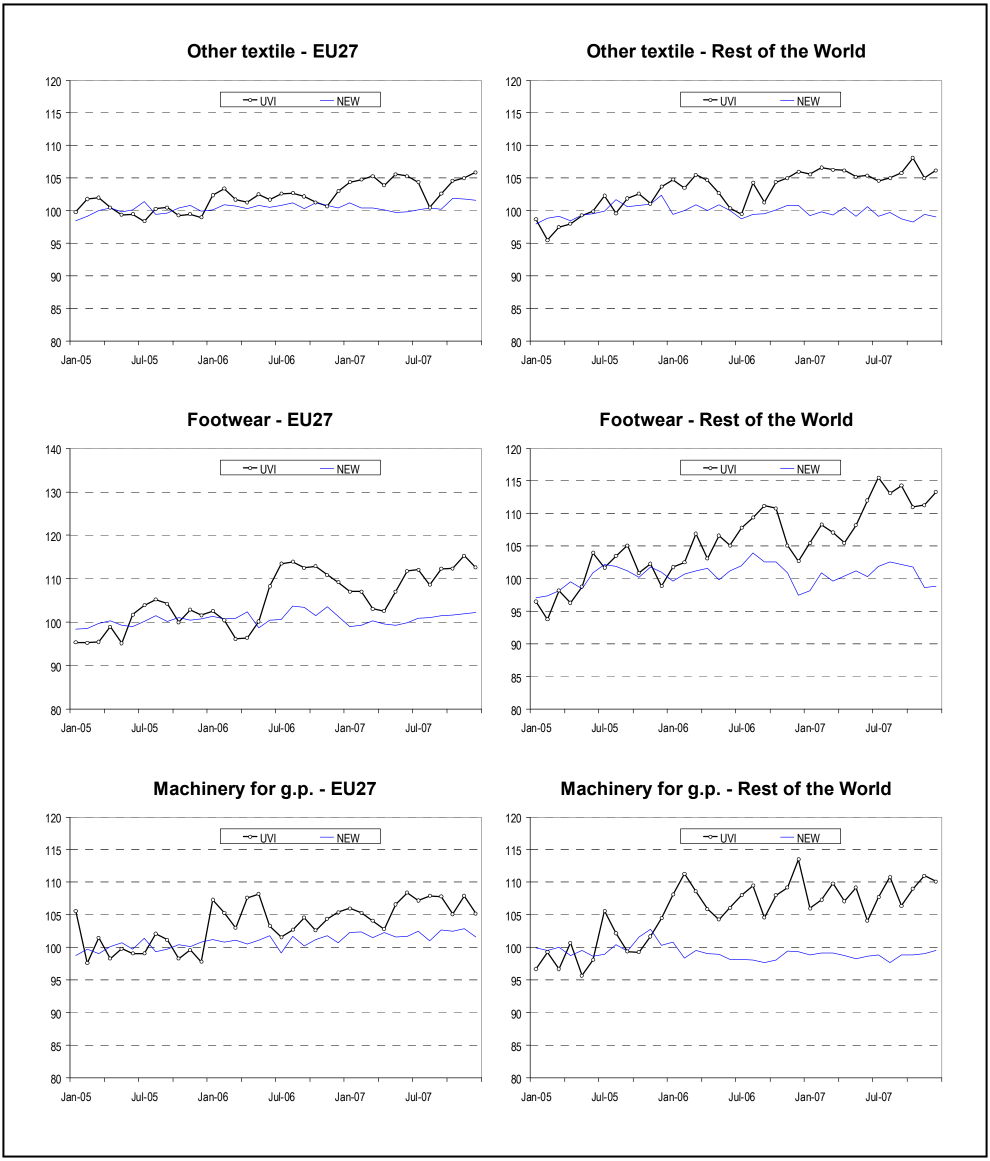
Appendix

Figure 1 - UVI and NEW monthly indices by product and area of destination. Exports. Years 2005-2007 *(base year: 2005=100)*



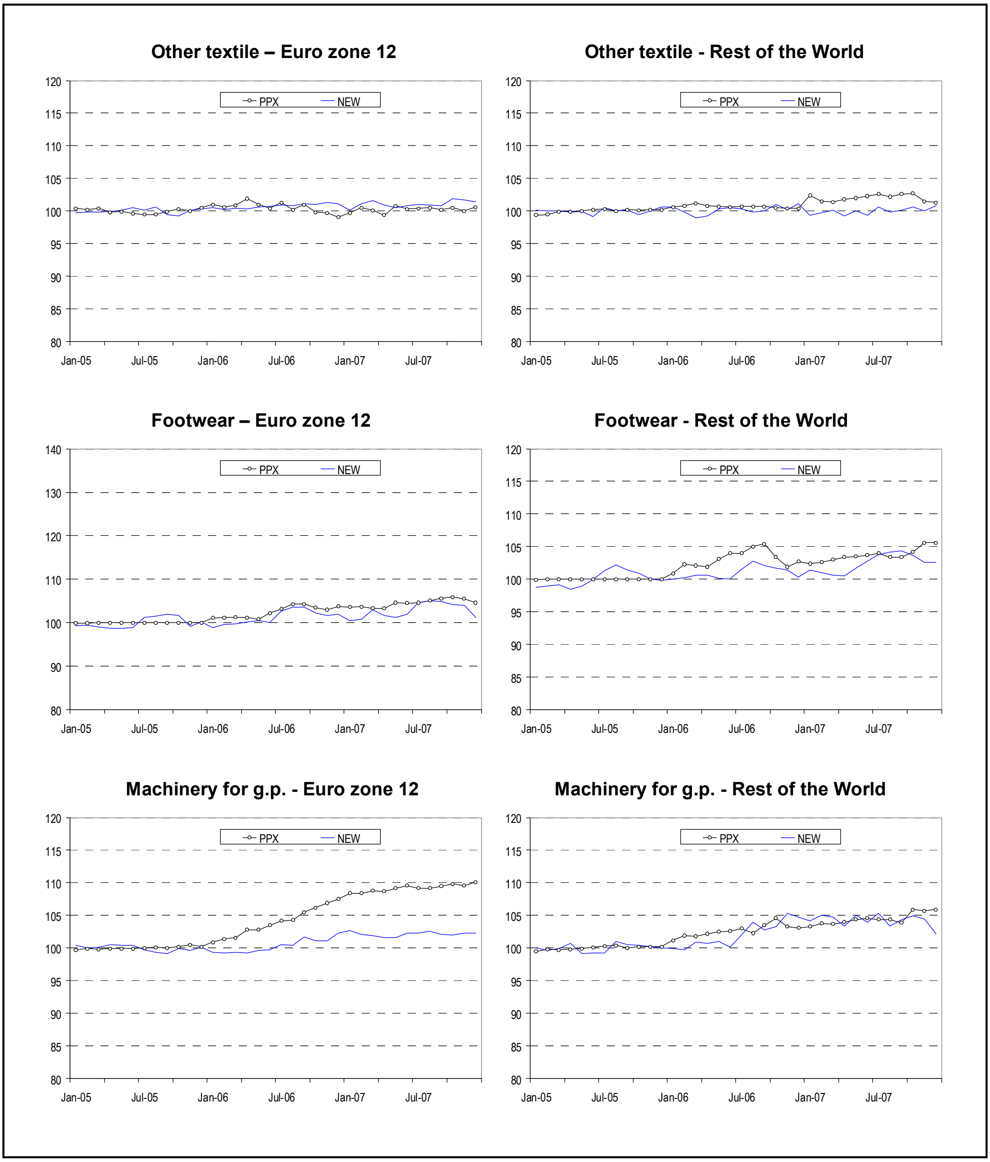
Source: ISTAT for UVIs series

Figure 2 - UVI and NEW monthly indices by product and area of origin. Imports. Years 2005-2007
(base year 2005=100)



Source: ISTAT for UVIs series

Figure 3 - XPI and NEW monthly indices by activity and area of destination. Years 2005-2007. *(base year 2005=100)*



Source: ISTAT for PPIX series

References

Anitori P., Causo M.S. (2008), “La metodologia di calcolo dei nuovi indici dei valori medi unitari del commercio con l’estero”, Atti del Convegno ”*L'informazione statistica ufficiale per l'analisi economica dell'internazionalizzazione delle imprese*”. Istat, Roma 12 giugno 2008. http://www.istat.it/istat/eventi/2008/internazionalizzazione_imprese/relazioni/anitori_causo.pdf.

Anitori P., Causo M.S., de Santis G. (2010), “Alcune considerazioni sulla qualità degli indici dei valori medi unitari del commercio estero”. *Rivista di statistica ufficiale*, n. 2-3, pag. 5-29.

Balk B.M. (1998), “On the use of unit value indices as consumer price subindices”. Proceedings of the *Fourth Meeting of the International Working Group on Price Indices*.

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

Banca d’Italia. (2008), *Bollettino Economico*, n°52, Aprile, Roma.

Bean S.J., Tsokos C.P. (1980), “Developments in Nonparametric Density Estimation”, *International Statistical Review*, Vol. 48, No. 3 (Dec., 1980), pp. 267-287

Bradley R. (2005), “Pitfalls of using unit values as a price measure or price index”. *Journal of Economic and Social Measurement,* 30, 39–61

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

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

EFC - Economic and Financial Committee (2010), *2010 EFC Status Report on Information Requirements in EMU*. Bruxelles, draft.

European Commission. (2006), *Communication from the Commission to the Council, the European Parliament, the European economic and social committee and the Committee of the regions global Europe: competing in the world. A Contribution to the EU's Growth and Jobs Strategy*. Brussels, 4 october, COM(2006) 567 final

European Commission. (2009), *Communication from the Commission to the European Parliament and the Council on the production method of EU statistics: a vision for the next decade*. Brussels, 10 august, COM(2009) 404 final

EUROSTAT. (2001a), *Handbook on price and volume measures in national accounts*. Luxembourg, p. 145.

EUROSTAT. (2001b), *Compendium of HICP Reference Documents. European Communities*. Luxembourg, p. 354.

EUROSTAT. (2005), *European Statistics code of practice*. Luxembourg.

EUROSTAT. (2006), *Methodology of short-term business statistics- Interpretation and guidelines*. Luxembourg.

Fenwick D. (2001), "Sampling in Consumer Price Indices: what role for scanner data?" Sixth Meeting of the *International Working Group on Price Indices*, Canberra, 2-6 April.

Francois J., Manchin M., Norberg H., Spinanger D. (2007), *Impacts of textiles and clothing sectors liberalisation on prices*. Kiel Institute for the world economy and the Commission of the European Union - Directorate-General for Trade.

Giovannetti G. (2007), "Strategie di prezzo e qualità: i settori tessile-abbigliamento, calzature e gioielleria". In: Lanza A., Quintieri B., *Eppur si muove*. Rubettino, Roma.

Institute for International and Development Economics. (2007), *Passing on of the benefits of trade openness to consumers*. Directorate General for Trade, European Commission.

IMF (International Monetary Found). (2009), *Export and Import price index manual: theory and practice*. Washington, US.

ISAE. (2009), *Rapporto ISAE: le previsioni per l'economia italiana - Luglio 2009*. Roma.

ISTAT (2003), I nuovi indici del commercio con l'estero (base 200=100), Roma.

Izenman A.J. (1991), "Recent Developments in Nonparametric Density Estimation". *Journal of the American Statistical Association*, Vol. 86, No. 413 (Mar., 1991), pp. 205-224

Jain M., Caddy J. (2001), "Using Scanner Data to Explore Unit Value Indexes". Sixth Meeting of the *International Working Group on Price Indices*, Canberra, Australia 2-6 April

Jones M.C., Marron J. S., Sheather S. J. (1996), "A Brief Survey of Bandwidth Selection for Density Estimation". *Journal of the American Statistical Association*, Vol. 91, No. 433 (Mar., 1996), pp. 401-407

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

OECD. (2008), *Use of administrative data: Short-Term Economic Statistics (STES) Administrative Data: Two Frameworks of Papers*. Paris.

Parniczky G. (1980), "On the Inconsistency of World Trade Statistics". *International Statistical Review*, Vol. 48, No. 1 (Apr., 1980), pp. 43-48

Pisani A. (1993), "Non-Parametric and Scale-Independent Method for Cluster Analysis - Part One - the Univariate Case". *Royal Astronomical Society Monthly Notices*, V.265, no. 3/Dec1, p. 706-726.

Ramsey P.H., Ramsey P.P. (2007), "Optimal Trimming and Outlier Elimination". *Journal of Modern Applied Statistical Methods*, Vol.6, No. 2, 355-360

Ribe M. (2000), "Price Index, Statistical Inference and Harmonisation". Fourth International Conference on *Methodological Issues in Official Statistics*. Stockholm,12-13 October.

Siegel D. (1991), "Errors in output deflators revisited: unit values and the PPI". *NBER Working Papers Series*, n. 3935

Silver M. (2007), "Do Unit Value Export, Import, and Terms of Trade Indices Represent or Misrepresent Price Indices?" *IMF Working Paper*, n.121

Silver M., Webb B. (2002), "The measurement of inflation: Aggregation at the basic level". *Journal of Economic and Social Measurement*, 28 (2002) 21–35

Silverman B.W. (1981), "Using Kernel Density Estimates to Investigate Multimodality". *Journal of the Royal Statistical Society*. Series B (Methodological), Vol. 43, No. 1 (1981), pp. 97-99

Silverman B.W. (1986), *Density estimation for statistics and data analysis*. London, Chapman & Hall, IX, p.175.

Sprent P., Seeton N.P. (2001), *Applied non parametric statistical method*, 3rd edition Chapman & Hall.

Stigler S.M. (1977), "Do Robust Estimators Work with Real Data?" *The Annals of Statistics*, vol.5 no.6, 1055-1098

Terrell G.R., Scott D.W. (1992), "Variable Kernel Density Estimation". *The Annals of Statistics*, vol. 20, no. 3, pp. 1236-1265

Triplett J.E. (2001), "Using Scanner Data in the Consumer Price Index: Some Neglected Conceptual Considerations". In Feenstra R., Shapiro M. (eds.), *Scanner data and price indices*. Chicago: University of Chicago Press.

United Nations. (1981) *Strategies for price and quantity measurement in external trade.*, Serie M, n. 69, New York.

United Nations. (2003), *Handbook of statistical organisation*. Third edition, Series F, n°88, New York.

United Nations. (2008), *International Merchandise Trade Statistics: supplement to the Compilers Manual.* ST/ESA/STAT/SER.F87/Add.1, New York.

Von der Lippe P. (2001), "Chain indices: a study in price index theory". *Publication Series*, Vol.16, Metzler Poeschel, Wiesbaden.

Von der Lippe P. (2007), "Price indices and unit value indices in German foreign trade statistics". *MPRA Paper* n.5525, Munich.

Wand M. P., Marron J. S., Ruppert D. (1991), Transformations in Density Estimation. *Journal of the American Statistical Association*, Vol. 86, No. 414 (Jun., 1991), pp. 343-353

Wilcox Rand R. (2006), *Robust estimation and hypothesis testing.* Elsevier Academic Press, UK.

Un indice sintetico non compensativo per la misura della dotazione infrastrutturale: un'applicazione in ambito sanitario[1]

Matteo Mazziotta[2] *e Adriano Pareto*[3]

Sommario

Il presente lavoro si propone di fornire uno strumento innovativo per la misura della dotazione infrastrutturale sanitaria, nell'ipotesi che ciascuna componente della dotazione non sia sostituibile con le altre (i valori non possono compensarsi). Il criterio proposto, denominato "metodo delle penalità per coefficiente di variazione", standardizza gli indicatori in modo da depurarli sia dall'unità di misura che dalla loro variabilità e utilizza, come funzione di aggregazione, una media aritmetica corretta mediante un coefficiente di penalità che dipende, per ciascuna unità, dalla variabilità degli indicatori rispetto al valor medio. L'indice ottenuto è di semplice determinazione ed è facilmente interpretabile e comparabile nel tempo. La metodologia di sintesi è stata applicata, a titolo di esempio, agli indicatori di dotazione delle apparecchiature ospedaliere ed extraospedaliere ad alta specialità ed elevata intensità organizzativa.

Abstract

The aim of this work is to provide a tool for the measurement of the health infrastructure endowment based on the assumption of "non-substitutability" of the indicators (a compensation among them is not allowed). The proposed approach, called "Method of Penalties by Coefficient of Variation", normalizes the indicators by a specific criterion that deletes the unit of measurement and the variability effect and uses, as aggregation function, an adjusted mean by a penalty coefficient that is function, for each unit, of the indicators' variability in relation to the mean. The obtained index is easily computable, interpretable and comparable over time. As an example of application, we consider a set of indicators of endowment of advanced medical equipment.

Parole chiave: indici sintetici, infrastrutture, sanità

1. Introduzione

La misura del grado di infrastrutturazione di un generico territorio è un obiettivo complesso e ambizioso che pone rilevanti problemi di carattere metodologico; oltre alle difficoltà relative al reperimento dei dati, si presentano, infatti, diversi problemi di aggregazione e interpretazione dei risultati. La complessità principale risiede nella

[1] Sebbene il lavoro sia frutto dell'opera di entrambi gli autori, i paragrafi 1, 2, 4.1, 4.3 e 6 vanno attribuiti a Matteo Mazziotta e i paragrafi 3, 4.2, 4.4 e 5 vanno attribuiti ad Adriano Pareto.

[2] Ricercatore (Istat), e-mail: mazziott@istat.it.
[3] Ricercatore (Istat), e-mail: pareto@istat.it.

multidimensionalità del fenomeno, la misurazione del quale richiede, inizialmente, il superamento di ostacoli di natura concettuale e definitoria legati alla peculiarità del fenomeno stesso e, successivamente, la scelta tra l'adozione di una misura analitica, rappresentata da un sistema di indicatori elementari, e la costruzione di una misura sintetica che, mediante un'opportuna funzione di aggregazione, sia capace di raccogliere i molteplici aspetti del fenomeno oggetto di studio.

Considerato che il sistema di indicatori elementari fornisce sempre una informazione completa ed esaustiva, è convinzione degli autori che la multidimensionalità possa rendere eccessivamente gravosa la lettura e l'analisi del fenomeno in osservazione. L'obiettivo di poter disporre, per ciascuna area geografica, di una quantificazione univoca (unidimensionale) che raccolga in sé "tutte" le informazioni, in modo da renderle immediatamente visibili e interpretabili, può semplificare notevolmente l'analisi territoriale dei dati, completando e non sostituendo quanto già emerso dall'analisi dei singoli indicatori. Inoltre, una singola misurazione può costituire un valido ausilio per il *policy maker* che, dovendo trasformare le informazioni in decisioni, può risultare particolarmente favorito dalla immediata fruibilità propria di tali indici.

L'ampio confronto esistente in letteratura relativamente alla validità e affidabilità degli indici sintetici deve indurre tutti gli addetti ai lavori, *in primis* il *policy maker*, alla massima cautela. E' pertanto opportuno sottolineare le notevoli limitazioni di queste misure: da una parte spiccano le numerose componenti di arbitrarietà che necessariamente si introducono, in particolare per ciò che riguarda la selezione degli indicatori elementari, dall'altra emergono tutti gli aspetti metodologici connessi ai criteri di normalizzazione, standardizzazione e sintesi dei dati (Brunini *et al.*, 2002). Il rischio di un utilizzo "non consapevole" di queste misure può portare a trarre conclusioni eccessivamente semplicistiche in quanto, tali misure, sono frequentemente utilizzate dai decisori, allo scopo di formare la graduatoria delle unità territoriali oggetto di studio. Rischiosa può essere anche la realizzazione di confronti temporali: in tal senso è necessario utilizzare con estrema attenzione i vettori "obiettivo" e, più in generale, l'analisi multivariata, cercando, comunque, di mantenere il controllo delle numerose variabili coinvolte.

In sostanza, implementare un indice sintetico significa percorrere una strada con numerosi ostacoli, il superamento dei quali può richiedere decisioni difficili e arbitrarie, con il rischio di perdere lungo il tragitto informazioni preziose che evidentemente caratterizzano le aree geografiche. Si pensi, da un lato, agli ostacoli riguardanti la disponibilità dei dati, la scelta degli indicatori più rappresentativi e il loro trattamento per renderli confrontabili (standardizzazione) e, dall'altro, alla definizione della funzione di sintesi; in quest'ultima fase, forse la più critica, l'arbitrio del ricercatore assume un ruolo fondamentale; infatti, le scelte possibili per giungere a un indice sintetico sono numerose e spaziano dagli strumenti di statistica descrittiva alle tecniche di analisi multivariata, dall'adozione di misure di distanza fino all'applicazione di funzioni lineari e non.

Nonostante i limiti metodologici accennati, gli indici sintetici sono ampiamente utilizzati da numerosi organismi internazionali per misurare fenomeni di natura economica, ambientale e sociale (Saisana e Tarantola, 2002) e per questo costituiscono uno strumento quanto mai attuale e in corso di evoluzione.

In questo lavoro, si propone un metodo per il calcolo di un indice sintetico non compensativo, basato sulla proprietà di "non sostituibilità" delle componenti.

2. Il quadro concettuale di riferimento

Gli autori che, nel passato, hanno fatto ricorso a indici sintetici per misurare la dotazione infrastrutturale sono numerosi e autorevoli. In Europa, l'esperienza apripista di questo approccio risale ai primi anni Ottanta con la pubblicazione del documento della Commissione Europea che proponeva uno studio sul contributo delle infrastrutture allo sviluppo regionale. A tale studio si legano, direttamente e indirettamente, tutti i lavori realizzati successivamente in Italia dal mondo accademico e da qualificati istituti di ricerca. Questo approccio, descritto da D. Biehl (1991) in numerosi lavori, fornisce una misura della dotazione infrastrutturale di un'area in termini fisici, utilizzando quindi una versione semplificata del metodo dell'inventario comune, del quale si evita la fase di determinazione del prezzo dei singoli beni capitali (Mazziotta, 2005). Mediante tale approccio[4] si giunge alla formulazione di un set di indicatori che esprimono la dotazione di infrastrutture nel territorio. Il gruppo della Commissione Europea opta per una tecnica di sintesi articolata in due fasi che vedono dapprima l'applicazione della media aritmetica e successivamente della media geometrica. Questa stessa tecnica è stata riproposta in numerosi studi italiani per più di un decennio; solo successivamente sono stati pubblicati, da fonti ufficiali, indici di infrastrutturazione calcolati mediante un approccio diverso che vede anche l'applicazione dell'analisi multivariata (Istituto Tagliacarne e Unioncamere, 2001).

In tale contesto, si inserisce il "Sistema di indicatori di dotazione e funzionalità delle infrastrutture", pubblicato dall'Istat nel corso del 2006, nato nell'ambito del Progetto "Informazione statistica territoriale e settoriale per le politiche strutturali 2001-2008" cofinanziato dal Dipartimento per le Politiche di Sviluppo e Coesione del Ministero dello Sviluppo Economico. Il prodotto, supportato da un *software* per la navigazione dei dati e contenente circa 500 variabili e 450 indicatori, calcolati dal 1996 e articolati secondo i diversi livelli territoriali, a partire da quello provinciale, ha il pregio di fornire agli studiosi uno strumento ufficiale e completo per la determinazione di misure analitiche di infrastrutturazione. L'esperienza critica maturata in Istat in materia di tecniche di standardizzazione e di sintesi è stata oggetto, negli ultimi anni, da parte dei ricercatori impegnati nella realizzazione del "Sistema di indicatori di dotazione e funzionalità delle infrastrutture", di ampi studi che hanno messo in evidenza le caratteristiche e gli effetti che la loro applicazione produce sull'output finale (Brunini *et al.*, 2002; Paradisi e Brunini, 2006). Negli ultimi anni, la letteratura sull'argomento si è popolata di numerosi studi e ricerche applicate a diversi contesti scientifici (tra gli altri, Palazzi, 2004). Queste analisi hanno persuaso i ricercatori dell'inesistenza di un metodo che produca risultati oggettivamente e universalmente validi e che, al contrario, debbano essere i dati e gli obiettivi contingenti a condurre, di volta in volta, all'individuazione di un criterio in grado di produrre soluzioni robuste, attendibili e coerenti con gli obiettivi proposti. Per tale motivo, nel processo di creazione di una misura sintetica della dotazione di infrastrutture sanitarie, vengono ripercorse in modo analitico tutte le fasi, ponderando con attenzione le implicazioni di ciascuna rispetto ai risultati finali.

[4] E' il caso di notare che questo approccio non è il solo presente in letteratura per misurare la dotazione infrastrutturale. Numerosi studiosi preferiscono, infatti, ricorrere al metodo dell'inventario permanente che fornisce una stima dello stock di capitale pubblico. Il dibattito sulla contrapposta validità dei due metodi è tuttora molto vivace, tuttavia non mancano autori che considerano i due approcci complementari (Picci, 2001).

3. La sintesi degli indicatori

Com'è noto, la metodologia per la costruzione di un indice sintetico prevede le seguenti fasi: a) definizione del fenomeno oggetto di studio; b) selezione degli indicatori elementari; c) standardizzazione degli indicatori elementari; d) ponderazione e aggregazione degli indicatori standardizzati in uno o più indici sintetici (OECD, 2008).

Nel caso della misura della dotazione infrastrutturale, i passi costitutivi di tale procedura possono sintetizzarsi come segue:

a) *definizione del fenomeno oggetto di studio*. Il concetto da misurare deve essere chiaro e univoco; in particolare, occorre stabilire se s'intende misurare la dotazione sulla base di dati fisici (si veda l'applicazione del par. 5) o valutarla in termini monetari. Tutti gli aspetti costitutivi devono essere esplicitati;

b) *selezione degli indicatori elementari*. Generalmente, nasce dal compromesso tra l'esigenza di non sostituibilità e rappresentatività di ogni aspetto essenziale del concetto e la disponibilità dei dati che può condizionare sia il dettaglio delle categorie infrastrutturali da quantificare, sia il livello delle aree territoriali da prendere in considerazione;

c) *standardizzazione degli indicatori elementari*. Spesso, gli indicatori elementari non sono comparabili tra loro, poiché risultano espressi in unità di misura diverse (per esempio, "Percentuale di Asl con dipartimento di salute mentale" e "Camere iperbariche extraospedaliere per 100 mila abitanti"). La standardizzazione è il procedimento che consente di convertire gli indicatori elementari in numeri puri o adimensionali. I principali metodi di standardizzazione consistono nel trasformare gli indicatori elementari in:
   - ranghi (si sostituiscono i valori assoluti con le posizioni);
   - numeri indici (si dividono i valori assoluti per una "base" di riferimento, per esempio la media o il massimo);
   - indici relativi rispetto al campo di variazione (si riportano i valori assoluti a un intervallo comune, per esempio 0-1);
   - scarti standardizzati (si riportano i valori assoluti a una scala con media e varianza prefissate, per esempio media 0 e varianza pari a 1);

d) *ponderazione e aggregazione degli indicatori standardizzati*. Consiste nella definizione di un sistema di "pesi" con cui ponderare gli indicatori a seconda della loro importanza nel descrivere il fenomeno e nella scelta della funzione o metodologia di sintesi. Esistono, infatti, procedure tecnicamente semplici, di natura prevalentemente soggettiva, e metodologie più complesse, che richiedono l'uso dell'analisi multivariata. Per quanto riguarda le funzioni di aggregazione, il procedimento più consueto consiste nel ricorso a una qualche forma di media, spesso identificata nella media aritmetica per l'aggregazione degli indicatori elementari all'interno di categorie e nella media geometrica per la sintesi delle categorie.[5] Tra le procedure più complesse ricordiamo il metodo tassonomico di Wroclaw e l'Analisi in Componenti Principali.

4. Il metodo delle penalità per coefficiente di variazione

4.1 Generalità

Il *metodo delle penalità per coefficiente di variazione* consente di costruire una misura sintetica della dotazione infrastrutturale di un insieme di unità territoriali, nell'ipotesi che ciascuna componente della dotazione non sia sostituibile con le altre o lo sia solo in parte.[6] Tale approccio, detto anche non compensativo, richiede una dotazione bilanciata di tutte le componenti elementari (Mazziotta e Pareto, 2007).

L'indice ottenuto si basa sui seguenti requisiti:

1) standardizzazione degli indicatori mediante un criterio di trasformazione che consenta di liberarli sia dall'unità di misura che dalla loro variabilità (Delvecchio, 1995);
2) sintesi indipendente da un'unità "ideale", in quanto la definizione di un insieme di valori obiettivo è soggettiva, non è univoca e può variare nel tempo (Aureli Cutillo, 1996);
3) semplicità di calcolo;
4) facilità di interpretazione.

Tali requisiti possono essere soddisfatti sulla base delle seguenti considerazioni.

Com'è noto, distribuzioni di indicatori diversi, misurati in modo diverso, possono essere confrontate, al netto della ponderazione implicita che la diversa variabilità induce sull'indice sintetico, mediante la trasformazione in scarti standardizzati. Pertanto è possibile riproporzionare gli indicatori elementari, in modo che oscillino tutti entro la medesima scala, trasformando ciascun indicatore in una variabile standardizzata con media 100 e scostamento quadratico medio pari a 10: i valori così ottenuti saranno compresi, all'incirca, nell'intervallo (70; 130).[7]

La standardizzazione rispetto alla media e allo scostamento quadratico medio, inoltre, non richiede la definizione di un vettore di valori obiettivo, in quanto sostituisce tale vettore con l'insieme dei valori medi. In tal modo, risulta agevole individuare le unità territoriali che hanno un livello di dotazione delle infrastrutture superiore a quello medio (valori maggiori di 100) e le unità che hanno un livello di dotazione inferiore (valori minori di 100).

In questo contesto, è possibile "correggere" la funzione di aggregazione (media aritmetica delle variabili standardizzate) mediante un coefficiente di penalità che dipende, per ciascuna unità territoriale, dalla variabilità degli indicatori rispetto al valor medio

---

[5] Il ricorso a due diverse medie assume, per implicito, che vi sia una certa compensazione tra le componenti elementari appartenenti a una stessa categoria e una minore sostituibilità tra le diverse categorie che concorrono a determinare l'indice sintetico complessivo (Mazziotta *et al.*, 2010).

[6] Generalmente, nell'ipotesi di non sostituibilità delle componenti elementari, si utilizza la media geometrica (Biehl, 1991). Tuttavia, la media geometrica presuppone che la grandezza da sintetizzare sia di natura moltiplicativa, anziché additiva, e assegna un peso maggiore ai valori più bassi. Inoltre, non può essere calcolata in presenza di valori negativi o nulli.

[7] In base al teorema di Bienaymé-Cebycev, i termini della distribuzione interni all'intervallo (70; 130) costituiscono almeno l'89 per cento del totale dei termini della distribuzione.

("variabilità orizzontale"). Tale variabilità, misurata attraverso il coefficiente di variazione, consente di penalizzare il punteggio delle unità che presentano un maggiore squilibrio tra i valori degli indicatori e, quindi, una dotazione sbilanciata.

L'uso degli scarti standardizzati nel calcolo dell'indice sintetico, infine, permette di costruire una misura robusta e poco sensibile all'eliminazione di un singolo indicatore elementare (Mazziotta *et al.*, 2010).

Nella tavola 1 è riportato un esempio di standardizzazione degli indicatori mediante: a) trasformazione in numeri indici con base uguale alla media e b) calcolo degli scarti standardizzati con media 100 e scostamento quadratico medio pari a 10.

Si noti che il calcolo dei numeri indici consente di liberare gli indicatori dall'unità di misura, ma non di svincolarli dalla loro variabilità. Ciò comporta un maggiore peso sulla media aritmetica degli indicatori che, in termini di numeri indici, hanno una variabilità più grande. Infatti, utilizzando i numeri indici, I3 ha un peso maggiore di I1 nel calcolo della media e l'unità A ottiene un punteggio maggiore dell'unità E (107,9 contro 101,6); mentre con le variabili standardizzate le due unità assumono lo stesso punteggio (103,7).

Quindi, volendo attribuire uguale importanza a ogni variabile, è necessario ricorrere a un criterio di trasformazione degli indicatori che consenta di depurarli, oltre che dall'unità di misura, anche dalla loro variabilità.

Tavola 1 - Confronto tra metodi di standardizzazione degli indicatori

UNITA'	Indicatori			Numeri indici (base=media)				Var. standardizzate			
	X1	X2	X3	I1	I2	I3	Media	Z1	Z2	Z3	Media
A	3	200	1.000	42,9	114,3	166,7	107,9	85,9	111,2	114,1	103,7
B	5	150	800	71,4	85,7	133,3	96,8	92,9	88,8	107,1	96,3
C	7	175	600	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
D	9	150	400	128,6	85,7	66,7	93,7	107,1	88,8	92,9	96,3
E	11	200	200	157,1	114,3	33,3	101,6	114,1	111,2	85,9	103,7
Media	**7**	**175**	**600**	**100**	**100**	**100**		**100**	**100**	**100**	
S.q.m.	**2,8**	**22,4**	**282,8**	**40,4**	**12,8**	**47,1**		**10**	**10**	**10**	

4.2 Il calcolo dell'indice sintetico

Il metodo per il calcolo dell'indice sintetico prevede i seguenti passi.

1) *Standardizzazione degli indicatori*

Sia $\mathbf{X}=\{x_{ij}\}$ una matrice di n righe (unità territoriali) ed m colonne (indicatori elementari). Indicando con:

$$\mathrm{M}_{x_j} = \frac{\sum_{i=1}^{n} x_{ij}}{n} \quad \text{e} \quad \mathrm{S}_{x_j} = \sqrt{\frac{\sum_{i=1}^{n} (x_{ij} - \mathrm{M}_{x_j})^2}{n}}$$

si costruisce la matrice $\mathbf{Z}=\{z_{ij}\}$ in cui:

$$z_{ij} = 100 \pm \frac{(x_{ij} - \mathrm{M}_{x_j})}{\mathrm{S}_{x_j}} 10 \qquad (1)$$

dove x_{ij} è il valore del j-esimo indicatore nell'i-esima unità e ± è il segno della relazione tra il j-esimo indicatore e il fenomeno da misurare (nel nostro caso, tutti gli indicatori elementari sono concordanti con il fenomeno della dotazione infrastrutturale e quindi si assume il segno +).

2) *Calcolo della "variabilità orizzontale"*

Data la matrice $\mathbf{Z}=\{z_{ij}\}$, si calcola il vettore dei coefficienti di variazione $\mathbf{CV}=\{\mathrm{cv}_i\}$ in cui:

$$\mathrm{cv}_i = \frac{\mathrm{S}_{z_i}}{\mathrm{M}_{z_i}}$$

dove:

$$\mathrm{M}_{z_i} = \frac{\sum_{j=1}^{m} z_{ij}}{m} \quad \text{e} \quad \mathrm{S}_{z_i} = \sqrt{\frac{\sum_{j=1}^{m} (z_{ij} - \mathrm{M}_{z_i})^2}{m}}$$

3) *Costruzione dell'indice sintetico*

L'indice sintetico dell'i-esima unità MPcv_i si ottiene mediante la formula:

$$\mathrm{MPcv}_i = \mathrm{M}_{z_i}(1 - \mathrm{cv}_i^2) = \mathrm{M}_{z_i} - \mathrm{S}_{z_i}\mathrm{cv}_i$$

in cui si corregge la media aritmetica degli indicatori standardizzati sottraendo una quantità (il prodotto $\mathrm{S}_{z_i}\mathrm{cv}_i$) proporzionale allo scostamento quadratico medio e funzione diretta del coefficiente di variazione.[8] In tal modo, le unità con valori standardizzati simili tra loro, ossia in analoga proporzione rispetto al vettore delle medie, sono meno penalizzate.

Il metodo delle penalità per coefficiente di variazione si basa su un modello additivo e non richiede, come la media geometrica , che l'intensità totale del fenomeno (la dotazione infrastrutturale) sia uguale al prodotto delle singole componenti.

Tale proprietà rende l'indice MPcv facilmente interpretabile, in quanto è possibile scomporre il punteggio di ciascuna unità in due componenti:

- l'effetto medio (ammontare della dotazione rispetto alle altre unità);
- l'effetto penalità ("variabilità orizzontale" o della dotazione).

[8] L'uso del quadrato del coefficiente di variazione nel calcolo dell'indice sintetico consente di limitare l'effetto "scavalcamento" tra due unità con medie aritmetiche diverse solo ai casi in cui l'unità con media aritmetica più alta ha una variabilità sensibilmente maggiore dell'altra.

La sintesi degli indicatori mediante il metodo proposto consente di realizzare, in modo semplice e immediato, analisi descrittive finalizzate a confronti temporali, oltre che spaziali, dello stato di fenomeni complessi.

Infatti, l'incremento dell'indice MPcv di una generica unità corrisponde all'aumento della dotazione infrastrutturale e/o alla diminuzione della "variabilità orizzontale" rispetto al contesto territoriale in cui essa si colloca. Ciò significa che se la dotazione complessiva dell'unità rimane costante nel tempo, ma il livello medio delle unità diminuisce, l'indice risulterà crescente, al netto della "variabilità orizzontale".

Nella tavola 2 è riportato un esempio di sintesi degli indicatori mediante: a) media aritmetica e b) indice MPcv. Come si può notare, le unità B e D, pur registrando una dotazione complessiva pari a quella dell'unità C, hanno una dotazione più sbilanciata e, quindi, nella graduatoria secondo l'indice MPcv occupano una posizione inferiore (il rango passa dalla seconda alla terza posizione), a causa della maggiore "variabilità orizzontale".

Tavola 2 - Confronto tra metodi di sintesi degli indicatori

UNITA'	Indicatori			Var. standardizzate			Media aritmetica		Indice MPcv	
	X1	X2	X3	Z1	Z2	Z3	Valore	Rango	Valore	Rango
A	3	100	1.000	85,9	84,2	114,1	94,7	5	92,7	5
B	5	300	800	92,9	100,0	107,1	100,0	2	99,7	3
C	7	300	600	100,0	100,0	100,0	100,0	2	100,0	2
D	9	300	400	107,1	100,0	92,9	100,0	2	99,7	3
E	11	500	200	114,1	115,8	85,9	105,3	1	103,5	1
Media	**7**	**300**	**600**	**100**	**100**	**100**				
S.q.m.	**2,8**	**126,5**	**282,8**	**10**	**10**	**10**				

4.3 Un indice sintetico generalizzato

L'indice sintetico basato sul metodo delle penalità per coefficiente di variazione può essere scritto, in forma generalizzata, nel seguente modo:[9]

$$\mathrm{MPI}_i^{+/-} = \mathrm{M}_{z_i} \pm \mathrm{S}_{z_i} \mathrm{cv}_i ,$$

dove il segno ± dipende dal tipo di fenomeno considerato e, quindi, dal verso degli indicatori elementari (Istat, 2009).

Se l'indicatore è di tipo *crescente* o *positivo*, ossia se a variazioni crescenti dell'indicatore corrispondono variazioni positive del fenomeno (per esempio, lo sviluppo di un'area geografica), si utilizza la versione con penalità negativa:

$$\mathrm{MPI}_i^{-} = \mathrm{M}_{z_i} - \mathrm{S}_{z_i} \mathrm{cv}_i$$

[9] MPI è l'acronimo di Mazziotta-Pareto Index.

Viceversa, se l'indicatore è di tipo *decrescente* o *negativo*, ossia se a variazioni crescenti dell'indicatore corrispondono variazioni negative del fenomeno (per esempio, la povertà di un'area geografica), si ricorre alla formula con penalità positiva:

$$\mathrm{MPI}_i^+ = \mathrm{M}_{z_i} + \mathrm{S}_{z_i}\,\mathrm{cv}_i\,.$$

Nel primo caso, il coefficiente di penalità corregge la media degli indicatori standardizzati "spingendola" verso il basso, mentre nel secondo la corregge "spingendola" verso l'alto.

In questo lavoro si assume:

$$\mathrm{MPcv}_i = \mathrm{MPI}_i^- = \mathrm{M}_{z_i} - \mathrm{S}_{z_i}\,\mathrm{cv}_i$$

essendo MPcv una misura sintetica della dotazione infrastrutturale nella sanità (indicatore *crescente* o *positivo*).

4.4 Proprietà e osservazioni

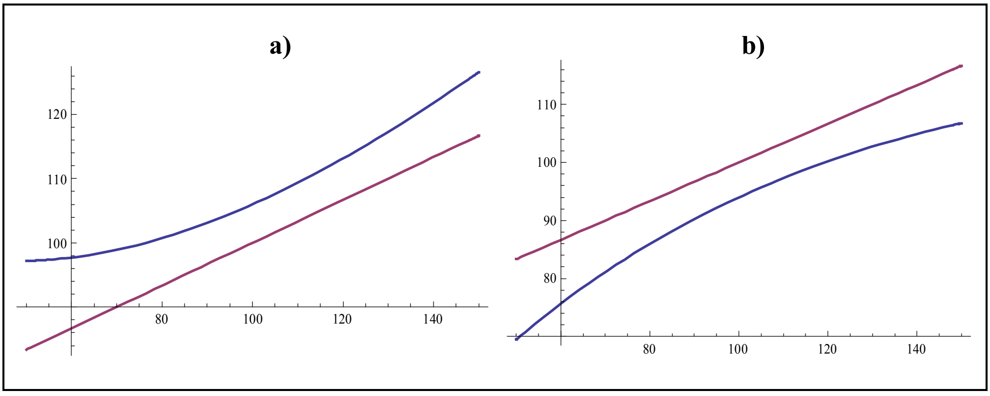
L'indice generalizzato può essere scritto, in forma compatta, nel seguente modo:

$$\mathrm{MPI}_i^+ = \frac{\sum_{j=1}^{m} z_{ij}^2}{\sum_{j=1}^{m} z_{ij}} \quad \text{e} \quad \mathrm{MPI}_i^- = \frac{2}{m}\sum_{j=1}^{m} z_{ij} - \frac{\sum_{j=1}^{m} z_{ij}^2}{\sum_{j=1}^{m} z_{ij}}$$

dove z_{ij} è dato dalla (1).

La figura 1 mostra un esempio dell'andamento degli indici MPI^+ (figura 1a) e MPI^- (figura 1b) al variare del generico valore z_{ik}. Nei due grafici è rappresentata anche la retta corrispondente alla media dei valori standardizzati in funzione di z_{ik}. Come si può notare, l'MPI^+ è una funzione *concava*, rispetto alla direzione positiva dell'asse delle ordinate, mentre l'MPI^- è una funzione *convessa*. Tuttavia, entrambe le funzioni possono essere considerate monotone crescenti nell'intervallo di riferimento.

Figura 1 - Esempio di funzione MPI^+ e MPI^-



Dall'esame dei grafici risulta chiaro l'effetto della penalità; infatti, maggiore è la distanza dal punto in cui la "variabilità orizzontale" è minima (nell'esempio z_{ik}=100), tanto più grande è la penalità e, quindi, la differenza tra l'indice MPI e la media aritmetica dei valori standardizzati.

Sulla base di quanto visto in precedenza, data una matrice $\mathbf{X}=\{x_{ij}\}$, è possibile ricavare le seguenti proprietà dell'indice MPI.

1) L'indice MPI$^+$ dell'i-ma unità è maggiore o uguale all'indice MPI$^-$ della medesima unità, ovvero:

$$\mathrm{MPI}_i^+ \geq \mathrm{MPI}_i^-$$

In particolare, si ha $\mathrm{MPI}_i^+ = \mathrm{MPI}_i^-$ se e solo se $\mathrm{S}_{z_i} = 0$

2) L'indice MPI$^+$ e l'indice MPI$^-$ dell'i-ma unità sono legati dalla seguente relazione:

$$\mathrm{MPI}_i^- = 2\mathrm{M}_{z_i} - \mathrm{MPI}_i^+$$

3) Date due unità i e h, con $\mathrm{M}_{z_i} = \mathrm{M}_{z_h}$, si ha:

$$\mathrm{MPI}_i^- > \mathrm{MPI}_h^- \text{ se e solo se } \mathrm{S}_{z_i} < \mathrm{S}_{z_h}$$

$$\mathrm{MPI}_i^+ > \mathrm{MPI}_h^+ \text{ se e solo se } \mathrm{S}_{z_i} > \mathrm{S}_{z_h}$$

4) Date due unità i e h, con $\mathrm{M}_{z_i} > \mathrm{M}_{z_h}$, si ha:

$$\mathrm{MPI}_i^- > \mathrm{MPI}_h^- \text{ se e solo se } \mathrm{M}_{z_i} - \mathrm{M}_{z_h} > \mathrm{S}_{z_i}\mathrm{cv}_i - \mathrm{S}_{z_h}\mathrm{cv}_h$$

$$\mathrm{MPI}_i^+ > \mathrm{MPI}_h^+ \text{ se e solo se } \mathrm{M}_{z_i} - \mathrm{M}_{z_h} > \mathrm{S}_{z_h}\mathrm{cv}_h - \mathrm{S}_{z_i}\mathrm{cv}_i$$

5) Sia $\mathrm{r}_{x_j x_k}$ il coefficiente di correlazione lineare tra il j-mo e il k-mo indicatore; se $\mathrm{r}_{x_j x_k} = 1$, per ogni j e k con $j \neq k$, allora:

$$\mathrm{MPI}_i^+ = \mathrm{MPI}_i^- = \mathrm{M}_{z_i}.$$

Tale risultato è dovuto al fatto che, per l'i-ma unità, si ottiene $z_{ij} = z_{ik}$ per $j \neq k$.

La proprietà 5 è molto interessante, in quanto mette in evidenza il legame tra il comportamento dell'indice MPI e la struttura delle correlazioni esistenti tra gli indicatori elementari.

Consideriamo il caso in cui $m = 2$.

Se tra gli indicatori c'è massima correlazione positiva, tutte le unità hanno scostamento quadratico medio S_{z_i} pari a 0 e l'MPI dipende esclusivamente dalla media M_{z_i}.

Se tra gli indicatori c'è massima correlazione negativa, tutte le unità hanno media M_{z_i} pari a 100 e l'MPI dipende esclusivamente dallo scostamento quadratico medio S_{z_i}.

Pertanto, nel primo caso, l'MPI ordina le unità secondo il valor medio; mentre nel secondo le ordina in base alla variabilità. In tutti gli altri casi, si ottiene una combinazione dell'effetto medio e dell'effetto penalità ("variabilità orizzontale").

In generale, tanto più gli indicatori elementari sono discordanti tra loro e maggiore è la variabilità orizzontale "indotta" in ciascuna unità, con conseguente aumento della penalità e, quindi, della differenza tra MPI e media aritmetica.

5. Un'applicazione a dati reali

In questo paragrafo sono illustrati i risultati relativi all'applicazione del metodo proposto agli indicatori di dotazione delle apparecchiature sanitarie ad alta specialità ed elevata intensità organizzativa, in Italia, a livello regionale.[10]

Gli indicatori utilizzati riguardano il numero di apparecchiature disponibili in relazione al bacino di utenza potenziale (popolazione residente) e si riferiscono all'anno 2005. I dati sono riportati nella tavola 3 e comprendono le apparecchiature disponibili presso strutture ospedaliere ed extraospedaliere.[11]

Tavola 3 - Apparecchiature ospedaliere ed extraospedaliere ad alta specialità ed elevata intensità organizzativa per regione - Anno 2005 *(indici per 100.000 abitanti)*

REGIONI	Camere iperbariche	Tac	Apparecchi per emodialisi	Tomografi a risonanza magnetica	Acceleratori lineari	Gamma camere computerizzate
Piemonte	0,02	2,47	26,75	1,87	0,51	0,71
Valle d'Aosta	0,00	2,43	29,98	1,62	0,00	2,43
Lombardia	0,15	2,54	25,80	1,46	0,56	0,82
Bolzano	0,21	1,88	23,34	1,88	0,42	0,83
Trento	0,00	1,80	41,60	0,60	0,80	0,20
Veneto	0,17	1,91	18,39	1,46	0,49	0,95
Friuli-Venezia Giulia	0,08	2,07	25,86	1,41	0,99	0,75
Liguria	0,06	2,19	28,79	1,50	0,56	1,12
Emilia-Romagna	0,10	2,18	21,90	1,22	0,41	0,74
Toscana	0,28	2,08	23,77	1,47	0,55	0,97
Umbria	0,00	2,90	38,34	1,27	0,81	0,69
Marche	0,07	2,49	30,78	2,10	0,59	1,31
Lazio	0,06	3,22	31,00	1,78	0,53	1,34
Abruzzo	0,00	2,23	25,34	1,38	0,54	1,00
Molise	0,31	4,67	32,04	2,49	0,62	2,18
Campania	0,40	3,89	9,79	1,26	0,31	1,90
Puglia	0,12	2,31	27,13	1,13	0,29	1,40
Basilicata	0,00	2,35	27,38	0,84	0,34	1,85
Calabria	0,15	3,69	22,47	1,25	0,40	1,54
Sicilia	0,28	3,07	28,87	1,22	0,52	2,33
Sardegna	0,61	2,36	28,80	1,39	0,12	1,21
Media	**0,146**	**2,606**	**27,053**	**1,457**	**0,493**	**1,251**
S.q.m.	**0,154**	**0,716**	**6,409**	**0,404**	**0,217**	**0,584**
C.V. (%)	**105,3**	**27,5**	**23,7**	**27,7**	**43,9**	**46,7**

Fonte: Istat, 2009

10 Si tratta di un esercizio di calcolo, per evidenziare le proprietà e il potenziale utilizzo dell'indice proposto.

11 In molte realtà regionali la dotazione di alcune apparecchiature risulta elevata nelle strutture extraospedaliere e contenuta in quelle ospedaliere e viceversa. Pertanto, nel computo degli indicatori, sono state considerate le apparecchiature di alta specialità ed elevata intensità organizzativa complessive (Istat, 2009).

Tra i dispositivi considerati, gli apparecchi per emodialisi sono i più numerosi, con una dotazione particolarmente elevata nella provincia autonoma di Trento, dove se ne registrano 41,6 ogni 100 mila abitanti, e in Umbria; mentre in Campania ne risultano 9,8 ogni 100 mila abitanti. Tuttavia, tale indicatore ha il coefficiente di variazione più basso (23,7%), contro una variabilità relativa quasi doppia per gli acceleratori lineari (43,9%) e le gamma camere computerizzate (46,7%).

Per quanto riguarda le camere iperbariche, invece, si passa dalla dotazione nulla di Valle d'Aosta, Trento, Umbria, Abruzzo e Basilicata ad una quota di 0,4 ogni 100 mila abitanti della Campania, con un coefficiente di variazione del 105,3%.

Gli indicatori elementari sono stati standardizzati mediante la formula (1) e la sintesi è stata condotta secondo i seguenti criteri:

a) *media aritmetica* degli indicatori standardizzati (Ma);
b) *media geometrica* degli indicatori standardizzati (Mg);
c) *mediana* degli indicatori standardizzati (Me);
d) *metodo delle penalità per coefficiente di variazione* (MPcv).

Nella tavola 4 sono riportati i quattro indici sintetici e le corrispondenti graduatorie delle regioni italiane, con le differenze di rango tra il metodo delle penalità per coefficiente di variazione e gli altri metodi.

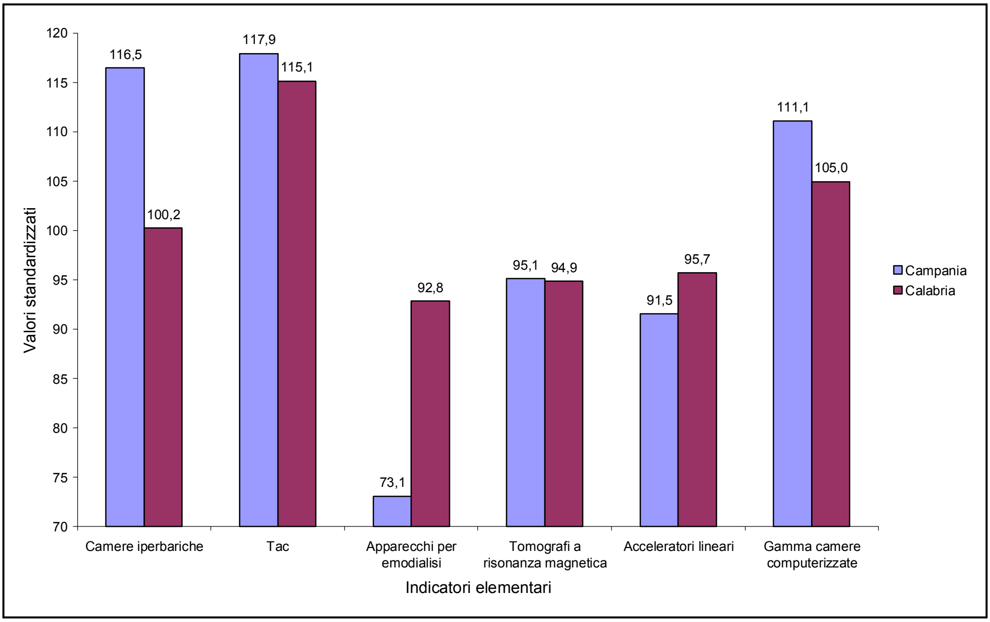
Tavola 4 - Graduatorie delle regioni italiane secondo la dotazione di apparecchiature ospedaliere ed extraospedaliere ad alta specialità ed elevata intensità organizzativa per metodo di sintesi - Anno 2005

Regioni	Ma		Mg		Me		MPcv		Differenza di rango		
	Valore	Rango	Valore	Rango	Valore	Rango	Valore	Rango	MPcv-Ma	MPcv-Mg	MPcv-Me
Piemonte	98,5	14	98,3	13	98,8	11	98,1	13	-1	0	2
Valle d'Aosta	99,0	11	98,1	14	100,8	6	97,2	15	4	1	9
Lombardia	98,9	13	98,8	11	99,6	8	98,8	10	-3	-1	2
Bolzano	98,0	15	97,8	15	95,4	18	97,5	14	-1	-1	-4
Trento	96,1	18	94,8	20	89,6	21	93,3	21	3	1	0
Veneto	95,5	20	95,3	19	97,3	14	95,2	19	-1	0	5
Friuli-Venezia Giulia	99,9	9	99,4	9	96,9	16	98,8	8	-1	-1	-8
Liguria	98,9	12	98,8	12	99,4	9	98,7	11	-1	-1	2
Emilia-Romagna	94,1	21	94,1	21	94,1	20	94,0	20	-1	-1	0
Toscana	99,1	10	98,9	10	97,8	13	98,8	9	-1	-1	-4
Umbria	102,1	5	101,5	5	99,7	7	100,9	5	0	0	-2
Marche	103,4	3	103,2	4	102,7	5	103,0	4	1	0	-1
Lazio	103,4	4	103,3	3	103,9	3	103,2	3	-1	0	0
Abruzzo	96,4	17	96,4	17	96,5	17	96,3	17	0	0	0
Molise	115,8	1	115,4	1	113,3	1	115,1	1	0	0	0
Campania	100,9	7	99,5	8	103,1	4	98,3	12	5	4	8
Puglia	96,6	16	96,5	16	97,1	15	96,4	16	0	0	1
Basilicata	95,9	19	95,6	18	94,7	19	95,2	18	-1	0	-1
Calabria	100,6	8	100,3	7	98,0	12	100,0	6	-2	-1	-6
Sicilia	105,3	2	105,0	2	104,7	2	104,8	2	0	0	0
Sardegna	101,6	6	100,7	6	98,8	10	99,7	7	1	1	-3
Differenza media assoluta									**1,3**	**0,7**	**2,8**
Indice di cograduazione									**0,952**	**0,983**	**0,786**

I risultati più simili all'indice MPcv sono quelli ottenuti con la media geometrica, a cui corrisponde una differenza media assoluta di rango pari a 0,7 posizioni. Ciò è dovuto alla capacità intrinseca della media geometrica di "penalizzare" le distribuzioni con valori sbilanciati degli indicatori.[12] Nel complesso, infatti, 10 regioni mantengono la stessa posizione, 10 regioni si spostano di una posizione e solo una regione (la Campania) registra un incremento di 4 posizioni. Segue la media aritmetica che si pone, rispetto all'indice MPcv, in una posizione intermedia tra la media geometrica e la mediana (differenza media assoluta di rango pari a 1,3). In questo caso, solo 5 regioni mantengono la stessa posizione, mentre 4 presentano una differenza di rango di almeno 3 posizioni. Infine, la mediana fornisce risultati sensibilmente diversi (differenza media assoluta di rango pari a 2,8), con spostamenti, verso l'alto o verso il basso, anche di 8-9 posizioni. Infatti, se da una parte il calcolo della mediana consente di costruire un indice robusto e meno sensibile ai valori eccezionali, tale caratteristica non permette di valutare in maniera adeguata la qualità del bilanciamento della dotazione infrastrutturale. In generale, le maggiori differenze di rango tra l'indice MPcv e i metodi tradizionali tendono a essere positive, poiché tali metodi non penalizzano o penalizzano solo in parte eventuali squilibri tra i valori degli indicatori.

Consideriamo, a titolo di esempio, la dotazione infrastrutturale della Campania e della Calabria. Nella figura 2 è illustrato il diagramma a barre corrispondente ai valori standardizzati degli indicatori elementari di queste regioni.

Figura 2 - Indicatori di dotazione di apparecchiature ospedaliere ed extraospedaliere ad alta specialità ed elevata intensità organizzativa - Anno 2005



[12] Si noti che la media geometrica è minore o uguale alla media aritmetica. In particolare, la media geometrica e la media aritmetica coincidono se i valori degli indicatori sono uguali, mentre assumono valori diversi tanto più i valori degli indicatori differiscono tra di loro.

La Campania presenta un forte squilibrio nella dotazione di apparecchiature ospedaliere ed extraospedaliere, rispetto alla media, con un valore massimo di 117,9 per le tac e un valore minimo di 73,1 per gli apparecchi per emodialisi (campo di variazione uguale a 44,8). Al contrario, la Calabria ha una dotazione più equilibrata, con un valore leggermente inferiore per le tac, pari a 115,1, e un valore sensibilmente superiore per gli apparecchi per emodialisi, pari a 92,8 (campo di variazione uguale a 22,3).

In tal caso:

- la *media aritmetica* (Ma), basandosi sulla completa sostituibilità delle componenti, compensa la carenza di apparecchi per emodialisi con il surplus di tac, collocando la Campania al settimo posto e la Calabria all'ottavo;
- la *media geometrica* (Mg) penalizza, anche se in modo leggero, la presenza di un valore particolarmente basso per gli apparecchi di emodialisi, collocando la Campania all'ottavo posto e la Calabria al settimo;
- la *mediana* (Me) ignora i valori estremi (tac e apparecchi per emodialisi) e considera soltanto i valori centrali della distribuzione, particolarmente elevati nella Campania, collocando tale regione al quarto posto e la Calabria al dodicesimo;
- il *metodo delle penalità per coefficiente di variazione* (MPcv) penalizza il forte sbilanciamento della Campania, collocandola al dodicesimo posto, e al contempo valorizza il maggiore equilibrio della Calabria, classificandola in sesta posizione.

Pertanto, nell'ipotesi di non sostituibilità delle singole dotazioni, l'indice MPcv può costituire una valida soluzione, in grado di premiare le unità territoriali con una dotazione bilanciata delle infrastrutture.

6. Conclusioni

Valutare il grado di infrastrutturazione di un'area geografica, attraverso una misura unidimensionale, presenta numerosi rischi di errore; l'utilità di queste misurazioni in ambito sociale ed economico ha spinto, nel tempo, diversi studiosi a svilupparne di nuove e ad approfondirne gli aspetti strettamente metodologici, al fine di ottenere degli indici sintetici in grado di "conservare" il maggior numero di informazioni.

L'acceso dibattito all'interno della comunità scientifica, negli anni, sembra convergere verso l'idea che non esista un indice universalmente valido per tutti gli ambiti di applicazione e che, quindi, la sua validità dipenda dagli obiettivi strategici della ricerca.

Il *metodo delle penalità per coefficiente di variazione*, proposto dagli autori all'interno di un vasto programma di studi sull'argomento promosso dall'Istat, si propone di arricchire il panorama già esistente di un nuovo strumento di analisi che possa, da un lato, rispettare il più possibile le caratteristiche desiderabili di un indice dal punto di vista metodologico e, dall'altro, essere validamente applicato a più contesti scientifici.

Il difficile raggiungimento di tali obiettivi, durante la fase di implementazione dell'indice, ha spinto gli autori a limitare al massimo le fasi di arbitrarietà del ricercatore puntando, per la standardizzazione degli indicatori elementari, su strumenti statistici di semplice applicazione e comprensione che potessero eliminare unità di misura e variabilità. Inoltre, la funzione di aggregazione (media aritmetica) è stata corretta con una quantità proporzionale allo scostamento quadratico medio e funzione diretta del coefficiente di variazione, in modo da premiare le unità territoriali che presentano minore variabilità tra gli indicatori elementari ("variabilità orizzontale").

I risultati dell'applicazione alle apparecchiature ospedaliere (si veda, anche, Istat, 2009) mostrano una sostanziale uniformità di comportamento tra il metodo proposto, la media aritmetica e la media geometrica che, come noto, è la funzione di aggregazione maggiormente usata in letteratura per la misurazione della dotazione infrastrutturale. Tuttavia, il metodo delle penalità, basandosi su un modello additivo "corretto", supera i problemi di compensazione insiti nella media aritmetica e, rispetto alla media geometrica, non richiede che l'intensità totale del fenomeno (la dotazione infrastrutturale) sia pari al prodotto delle singole componenti.

Concludendo sembra emergere che, al di là dei considerevoli rischi di appiattimento dell'informazione di base, gli indici sintetici offrano un contributo di chiarezza insostituibile. A tal proposito, risulta fondamentale la selezione e l'interpretazione degli indicatori elementari e si ritiene, pertanto, "*…assolutamente indispensabile, per ottenere risultati validi e affidabili, sostenere la fase di scelta della batteria di indicatori di base con una impalcatura teorica che definisca la realtà sociale in ciascuna delle sue dimensioni*" (Delvecchio, 1995).

Riferimenti bibliografici

Aureli Cutillo E. (1996), *Lezioni di statistica sociale. Parte seconda, sintesi e graduatorie*, C.I.S.U., Roma.

Biehl D. (1991), “Il ruolo delle infrastrutture nello sviluppo regionale”, in: Boscacci F., Gorla G. (a cura di), *Economie locali in ambiente competitivo*, Franco Angeli, Milano.

Brunini C., Messina A., Paradisi F. (2002), “L’infrastrutturazione delle province italiane: metodi e sperimentazioni”, in: *VI Conferenza Nazionale di Statistica*, Istat, Roma.

Delvecchio F. (1995), *Scale di misura e indicatori sociali*, Cacucci editore, Bari.

Istat (2009), *La dotazione di infrastrutture e servizi nella sanità. Anni 2004-2005*, Informazioni, 8, Roma.

Istituto G. Tagliacarne, Unioncamere (2001), *La dotazione di infrastrutture nelle province italiane 1997-2000*, Roma.

Mazziotta C. (2005), “La stima del capitale pubblico a livello regionale: una riflessione di metodo”, in: Carlucci M., Esposito G. (a cura di), *Statistica economica e strumenti di analisi. Studi in memoria di Antonino Giannone*, ISCONA, Roma.

Mazziotta C., Mazziotta M., Pareto A., Vidoli F. (2010), “La sintesi di indicatori territoriali di dotazione infrastrutturale: metodi di costruzione e procedure di ponderazione a confronto”, *Rivista di Economia e Statistica del Territorio*, vol. 1, pp. 7-33.

Mazziotta M., Pareto A. (2007), “Un indicatore sintetico di dotazione infrastrutturale: il metodo delle penalità per coefficiente di variazione”, in: *Lo sviluppo regionale nell'Unione Europea - Obiettivi, strategie, politiche, Atti della XXVIII Conferenza Italiana di Scienze Regionali*, AISRe, Bolzano.

OECD (2008), *Handbook on Constructing Composite Indicators. Methodology and user guide*, OECD Publications, Paris.

Palazzi P. (2004), “Lo sviluppo come fenomeno multidimensionale. Confronto tra l’ISU e un indice di sviluppo sostenibile”, *Moneta e Credito*, vol. 227, pp. 279-309.

Paradisi F., Brunini C. (2006), “Una tecnica alternativa per la determinazione di una misura sintetica di infrastrutturazione”, *Rivista di Statistica Ufficiale*, vol. 2, pp. 33-65.

Picci L. (2001), “Le infrastrutture in Italia. Le differenze territoriali e l’efficienza della spesa”, in: Baldassarri M., Galli G., Piga G. (a cura di), *L’Italia nella competizione globale – Regole di mercato*, Edizioni Il Sole 24ore, Milano, 2001.

Saisana M., Tarantola, S. (2002), *State-of-the-art report on current methodologies and practices for composite indicator development*, EUR 20408 EN, Commissione Europea-JRC, Ispra.

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.

1B01201100400000

ISSN 1828-1982

€ 10,00