QUALITY ASSESSMENT OF STATISTICAL INFORMATION SERVICES. AN EXPLORATORY ANALYSIS WITH RASCH MODEL

Giuseppe Lecardane

Abstract. The quality has now assumed a primary position in the field of research and statistical information. Quality aspects of the statistical organization and production is essential to collect valuable information and identify service dimensions.
The exploratory analysis with the Rasch method concerns the qualitative and functionality assessment of service and territorial structures through satisfaction degree expressed by thematic referents.
Rasch model has proved to be useful for diagnostics of service quality and allows valid and reliable information on the production and organizational process with analysis of strengths and weaknesses.
Qualitative survey results show organizational skills of the statistical offices and referents to produce the data. In other words, it is argued that perceived quality depends on organizational variables of a soft type (organizational climate, human resources, knowledge) rather than a system one (structures and/or processes).
The objective of the study is exploratory to define specific measures aimed at improving the statistical function within the National Statistical System.

1. Introduction

Qualitative process evaluation of the services offered from a complex system, such as that of statistical information, is a key point towards improving the quality and effectiveness of the service provided to users.

The analysis of the process and the choice of an appropriate statistical methodology constitute excellent tools for evaluating the data production chain which guarantees continuous monitoring and the possibility of promptly intervening with corrective actions.

Istat has always maintained a policy for the quality of statistical data with a strong orientation towards the production process. Strengthening processes is one of the possible ways to improve the quality of the final data.

Based on these premises, this work is the result of some lines of intervention of the Istat research project "Experimentation for new methods of collecting environmental data" (Lecardane, G. 2022) [1] aimed at measuring and evaluating the perception of thematic managers/referees environmental data of the municipal environmental statistics offices on the organizational and production process of the statistical function within the territorial structures according to Istat directives on environmental data.

In the first part of the work, some main results of the quality survey will be analysed: the profile of the environmental thematic referent, the organizational and instrumental endowment of the service, the methods of acquisition, control, sending and archiving of statistical data.

We proceeded with the analysis of the evaluative perception of the environmental thematic referents regarding the organizational and production process of the statistical function within the territorial structures involved.

Due to the latent nature of the phenomenon, there is a need to identify appropriate statistical tools for an objective measurement of perception and to implement a synthesis of the answers through a questionnaire provided to the representatives of the public and private entities that collaborated in the Istat survey for environmental statistics.

Knowing the opinion of the representatives on the reliability of the data produced and transmitted to Istat is certainly important information, although not the only one, which can allow you to identify critical elements in the system and implement appropriate corrective actions.

The objective of the study is exploratory and proactive to identify specific measures aimed at improving the statistical function within the National Statistical System.

2. A survey on the production activity of statistical information

To study the functioning of the statistical production process, an evaluation questionnaire was prepared as a tool for critical reading of the data production.

The questionnaire, designed and used for the survey, was administered to officials and technicians of the public and private offices participating in the collection of urban environmental data, with the assignment of thematic referents and heads of the

[1] Research project Istat RD, Specific projects, ID 1095, Experimentation for new ways of collecting environmental data of cities - PPMO 2017/19 (Project ID: 1905 and 1400. ODS/164/2018 of 13/09/2018, ODS/141 /17, from 2017 to 2012). Project Manager and Coordinator: Lecardane Giuseppe. Project aimed at experimenting with new methods of data acquisition with the ARCAM online platform and analysis of the organizational processes of the municipal statistics offices for the archiving and transmission of environmental data to Istat.

statistics offices. The experimental phase involved the recognition and selection of nine municipal statistical offices (Table 1) among the regions of northern and southern Italy participating in the project (Abruzzo, Basilicata, Friuli-Venezia-Giulia, Molise, Sicily and Veneto) and the collaboration of almost seventy officials and technicians with the role of manager, referent and collaborator of the Istat environmental surveys.

Table 1 – *Statistical offices of the capital municipalities in the regions adhering to the project*

Regions	**Capital municipalities**
Abruzzo:	Chieti
Basilicata:	Matera
	Potenza
Friuli-Venezia-Giulia:	Udine
Molise:	Campobasso
Sicilia:	Caltanissetta
	Palermo
Veneto:	Padova
	Rovigo

Based on the topics structured in the evaluation questionnaire, the organizational and functional system of the service, the structure of the office, the human and instrumental resources involved, the production process of statistical information on environmental issues (acquisition, treatment, review, validation and transmission of data), as well as the evaluation of the quality perceived in the organizational aspects of the office/service.

Even if the outcome of the survey is not exhaustive of the entire structural and organizational of the local public system, the results nonetheless provide some important indications on the production process of the local statistical structures involved in the production of the statistical data and on possible corrective actions to be realize.

3. Exploratory analysis of service quality and the Rasch model

The analysis allows the quality and functionality assessments of territorial structure services through the satisfaction degree expressed by thematic respondents of the municipal statistics offices. Theme of quality has now assumed a primary position in the field of research and statistical information. Having a good knowledge

of functional and statistical quality in organizational and production terms is essential for gathering valuable information and identifying critical issues on which to intercede. The choice of the methodology for the estimation mainly depends on the categorical-ordinal nature of the results or *outcomes* (taken from questionnaires, tests, etc.). Attitude of the subjects to provide answers/judgments towards aspects of reality can be summarized with the term of *latent* results (latent trait). In the literature, the estimation of latent results is performed to evaluate skills, perceptions, abilities or more generally attitudes on different aspects of reality often not directly observable but which can be inferred from a set of items that measure *latent traits*. Among the many multidimensional methodologies used for the estimation of latent variables, Rasch model allows a particularly valid analysis, satisfying all the ideal properties of a quantitative and objective measure (Bergmann, Foster, et al.).

Rasch methodology (Rasch, 1980 and Wright 2000) is part of the *item response theory* models (Pearlin et al., 1990) for the estimation of *outcome* or *latent trait* measures. *Latent trait* is measured through the subjects responses to a series of items that represent manifestations of different agreement or disagreement degree: an "ability" towards the *latent trait* is defined as any "propensity" β_i of an individual *i* (in terms of satisfaction, ability, aptitude, etc.) respect on a *latent trait* and with δ_j the difficulty of the subjects to answer on a item *j* (j=1,…,m).

The model estimate, for each respondent, a parameter of ability and, for each level of response, a parameter of difficulty. Furthermore, it is possible to state that a subject is *able* not in absolute sense but in relative to the other subjects and to the questionnaire examined. Same criterion is re-proposed in terms of difficulty in the answer. Rasch is a logistic model and assumes that the probability for a subject to answer a question (items of the questionnaire or the test, etc.) depends on the difference of two parameters: ability/characteristics of the respondent β (direct function) and the difficulty of the question (item) δ (test or questionnaire) answering the items in increasingly difficult ways (inverse function), with a non-linear and logistic model. The measure is not proportional to the probability of response but to its transformation (*logit*):

Logit(p) = log(p/(1-p)) = β_i - δ_j p = probability of correct answer to an item, β_i = answering ability of the interviewee *i*, δ_j = item difficulty *j*.

The comparison between ability of an individual and difficulty of an item allows us to predict the probability of choosing an answer mode. Items with two categories and more the answer x on item j (x = 0, 1, …, k_j) implies a positive answer to each previous modality 0, 1, 2, ..., x-1; the difficulty of answering to category x in item j (δ_{jx}) is considered as overcoming the difficulties of previous steps (δ_{jw} with w = 0, …, x); the parameter that expresses the difficulty of item j step w (δ_{jw}) can be expressed as the sum of a component defined as *location* of the item (δj) with a component is the differential effect of step w respect on *location* (τ_{jw}): $\delta_{jw} = \delta_j + \tau_w$;

the classic *Dichotomous Rasch Model* (Rasch, 1960) is extended to items with *kj* possible answers in the *Rating Scale* (Andrich, 1988) in which the differential effect is identical for each item ($\delta_{jw} = \delta_j + \tau_w$). In the *Partial Credit* version, difficulties of each step are different between items ($\delta_{jw} = \delta_j + \tau_w$) and the probability in a subject *i* to answer in *x* mode in item *j* depends only on the difference between two parameters β_i e δ_j:

$$\pi_{ijx} = \frac{\exp\sum_{w=0}^{x}(\beta_i - \delta_{jw})}{\sum_{q=0}^{kj}\exp\sum_{w=0}^{x}(\beta_i - \delta_{jw})} \quad x = 0, \ldots, k_j \text{ with } \sum_{x=0}^{kj}\pi_{ijx} = 1 \text{ and } \delta_{j0} = 0 \qquad (1)$$

[1] probability that the *i-th* subject answers *x* and not *x-1* on the item (*odds*) is a function of the β_i person's *ability and difficult parameter* in according to a function logistic:

$$\frac{\pi_{ijx}}{\pi_{ij(x-1)} + \pi_{ijx}} = \frac{\exp\sum_{w=0}^{x}(\beta_i - \delta_{jx})}{1 + \exp(\beta_i - \delta_{jx})} \qquad x = 1, \ldots, k_j \qquad (2)$$

From [2] we deduce the relationship between the *logit* scores and the subject's *ability* and *difficulty* parameters of the item in real terms

$$\ln\left(\frac{\pi_{ijx}}{\pi_{ij(x-1)}}\right) = \beta_i - \delta_{jx} \qquad (3)$$

[3] shows that the *ability* and *difficulty* of the items are estimated in the same measurement unit (*logit*) and that subject *i* has a greater probability of answering *x* rather than *x-1* on item *j* only if the his *ability* is higher than the *difficulty* of category *x* for item *j*; if for this subject the difference were zero, he would have the same probability of answering *x* and *x-1* on item *j*. The estimation of the parameters with Rasch model in the *Partial Credit* version is obtained through the maximization of *likelihood function* based on the response vector of the individuals on each item, conditionally on the overall score obtained on all the *m* items (Wright, Masters, 1982).

4. Results of Rasch Analysis

The study provides interesting aspects on the activities, processes, resources and information products of structures involved to verify qualitative levels of territorial

statistical function. Table 2 lists items examined and the evaluation in terms of degree of judgment expressed by respondents useful for applying the model. Tables 3, 4 and 5 show the percentages of the ten items evaluated by thematic referents regarding the organizational aspects and collaborative degree with internal and external personnel to the service. Evaluations of interviewees on the functional and operational structure of the office are concentrated to a positive degree of judgment. However, some critical issues, judged below sufficient, are reported in staff training (34%), in assignment of personnel with statistical skills (32 %) and in organizational structure (18%). When asked whether there has been a change in the structure and processes compared to the previous three years, almost all subjects (over 70%) agree that the condition has remained almost unchanged. Finally, on the collaboration received from colleagues or other personnel in the statistical activities, the results are concentrated on good levels of cooperation, even if some difficulties for external personnel in other administrations (16% of respondents judge it not positive). The estimate carried out using the Rasch model and elaborated using the statistical package Winsteps & Facets Rasch Software - Rasch measurement computer program. Tables 6a and 6b show the statistics of the items: parameter estimate of the average *difficulty* of the items or *measure*, variability degree of the estimate or *standard error*, *infit* and *outfit* indices for measuring the goodness of fit to the model (*average and standardized*). Values are arranged in descending order of the estimate *measure*: from items with greater *difficulty* in answering and which require greater *ability* of the subjects (*positive values*) to more "*simple*" items with less *ability* to answer (*negative values*). Model sets *logit*=0 the *average difficulty* of the items.

Table 2 – E*valuation of perceived quality in the organizational aspects of the statistics office*

ITEM CODE	Item	Score	Judgement
66	IT and technological equipment	5	Excellent
67	Statistical staff assignment	4	Good
68	Staff training	3	Sufficient
69	Relationships with staff	2	Insufficient
70	Processes and workloads	1	Poor
71	Organizational structure		
72	Circulation, clarity of information		
73	Quality of statistical products and services offered		
74	Timeliness of response to users		
89	Overall office quality		

Table 3 – *Evaluation of the quality perceived in the organizational aspects of the office/service (percentage values)*

Organization of the office/service in terms of	Poor	Insufficient	Sufficient	Good	Excellent
IT and technological equipment	0,0	8,9	26,8	55,4	8,9
Statistical staff assignment	23,2	8,9	39,3	25,0	3,6
Staff training	10,7	23,2	37,5	28,6	0,0
Relationships with staff	5,4	1,8	28,6	53,6	10,7
Processes and workloads	10,7	8,9	26,8	53,6	0,0
Organizational structure	3,6	14,5	34,5	45,5	1,8
Circulation, clarity of information	3,6	5,4	37,5	50,0	3,6
Quality of statistical products and services offered	5,4	3,6	37,5	50,0	3,6
Timeliness of response to users	1,8	3,6	35,7	55,4	3,6
Overall office quality	1,8	1,8	21,4	69,6	5,4

Table 4 – *Evaluation of the change in the organizational aspects of the office/service compared to three years ago (percentage values)*

Organization of the office/service in terms of	Much worsened	Worsened	Unchanged	Improved	Much improved
IT and technological equipment	0,0	0,0	55,4	39,3	5,4
Statistical staff assignment	1,8	10,7	64,3	21,4	1,8
Staff training	1,8	8,9	64,3	25,0	0,0
Relationships with staff	1,8	1,8	69,6	23,2	3,6
Processes and workloads	3,6	19,6	53,6	21,4	1,8
Organizational structure	1,8	12,5	57,1	25,0	3,6
Circulation, clarity of information	1,8	3,6	69,6	21,4	3,6
Quality of statistical products and services offered	1,8	5,4	62,5	30,4	0,0
Timeliness of response to users	1,8	1,8	67,9	28,6	0,0

Table 5 – *Evaluation of degree of collaboration (percentage values)*

Evaluation of degree of collaboration with	Poor	Insufficient	Sufficient	Good	Excellent
Staff	3,6	1,8	21,4	58,9	14,3
Executive	3,6	1,8	14,3	67,9	12,5
Other operating units	5,4	5,4	35,7	50,0	3,6
Other public and/or private administrations	5,4	10,7	33,9	48,2	1,8
ISTAT	0,0	0,0	17,9	57,1	25,0

Table 6a – *Item fit statistics in the Rasch model*

INPUT: 66 Person 10 Item MEASURED: 66 Person 10 Item 5 CATS MINISTEP 3.69.1.4
Item statistics: MEASURE ORDER Person Reliability: 0,89 Item Reliability: 0,92

Entry	Total	Meas	Stand.	INFIT		OUTFIT		Item	Description
Num	Score	ure	Error	Mnsq	Zstd	Mnsq	Zstd		
2	155	1,63	0,21	0,87	-0,6	0,81	-0,8	67	Statistical staff assignment
3	159	1,46	0,21	0,81	-1	0,95	-0,2	68	Staff training
5	181	0,45	0,22	0,88	-0,5	0,66	-1,5	70	Processes and workloads
6	180	0,29	0,23	0,68	-1,6	0,7	-1,3	71	Organizational structure
8	192	-0,14	0,24	0,78	-1	0,91	-0,3	73	Qual. stat. prod. serv. offered
7	193	-0,2	0,24	1,07	0,4	0,7	-1,3	72	Circul., clarity of inform.
9	199	-0,55	0,25	0,66	-1,7	0,64	-1,6	74	Tim. of response to users
4	203	-0,8	0,25	1,34	1,5	1,14	0,7	69	Relationships with staff
1	204	-0,87	0,26	2,09	3,9	1,92	3,2	66	IT and tech. equipment
10	210	-1,28	0,27	0,89	-0,4	0,66	-1,6	89	Overall office quality
MEAN	*187,6*	*0,00*	*0,24*	*1,01*	*-0,1*	*0,91*	*-0,5*		
S.D.	*17,8*	*0,92*	*0,02*	*0,41*	*1,6*	*0,37*	*1,4*		

Table 6b – *Item fit statistics in the Rasch model*

INPUT: 66 Person 10 Item MEASURED: 66 Person 10 Item 5 CATS MINISTEP 3.69.1.4
Item statistics: MEASURE ORDER Person Reliability: 0,89 Item Reliability: 0,92

PT-MEASURE		EXACT MATCH		Item	Description
CORR.	EXP.	OBS%	EXP%		
0,85	0,79	62,5	59,0	67	Statistical staff assignment
0,79	0,78	66,1	60,1	68	Staff training
0,76	0,74	64,3	63,8	70	Processes and workloads
0,78	0,73	74,5	63,7	71	Organizational structure
0,73	0,71	69,6	66,8	73	Qual. stat. prod. serv. offered
0,69	0,71	78,6	67,2	72	Circul., clarity of inform.
0,75	0,69	78,6	67,1	74	Timeliness of response to users
0,68	0,68	60,7	68,0	69	Relationships with staff
0,46	0,67	58,9	69,0	66	IT and tech. equipment
0,64	0,66	82,1	70,4	89	Overall office quality
MEAN		69,6	65,5		
S.D.		7,9	3,6		

Results of the Rasch analysis show good model reliability indices (*Reliability index*: ratio between the *variance explained* by the model and the *total variance*. For values >0.8 index is *good*) with *item reliability* indices equal to 0.92 and *person reliability* equal to 0.89 (*good* reproducibility characteristics of the model) as well as a *good extension* of the *logit measure* of the item *difficulty* (from -1.28 to 1.63).

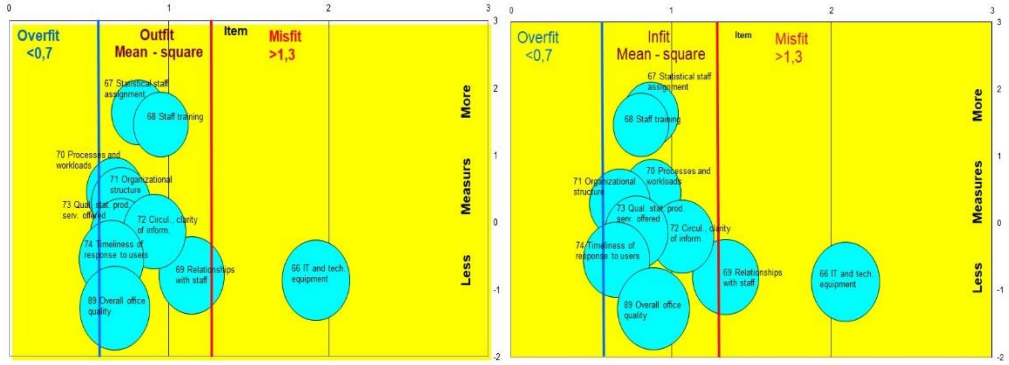
The *goodness of fit* to the *Rasch model* is evaluated in terms of *infit* (Unweighted Mean Square statistic) and *outfit* (Weighted Mean Square statistic) with *expected value* equal to 1 and defined on the comparison between *observed responses* (for

individual to each *item*) and *expected responses* based on the model. In practice, it is necessary to verify whether and to what extent the estimates obtained (*units* and *items*) distort the initial methods of *response*, a direct expression of the individual's responses to the items. Therefore, it is necessary to verify the existence of *systematic errors* that invalidate the good approximation of the starting data in the model. Two types of errors can be identified: *Misfitting*, unpredictable responses in contrast with the individual's *ability level* and the item's *difficulty* degree, i.e. *too positive responses* to *too difficult items* or *unsatisfactory responses* to *too difficult items*. easy, given the *ability* of the subject; *Overfitting*, answers too systematically linked to the *individual's ability* and the *item difficulty* without any effect of the *random component*. In essence, responses that are *too predictable* and therefore "*suspicious*" of being determined by *unknown external factors*. There is a *good fit* for values between 0.7 (*overfit*) <*Good Model*<1.3 (*misfit*). If *Overfit*, item responses are *too predictable* therefore "*suspicious*" and statistically *without any effect* of the *random component*; if *Misfit* responses are *not predictable*, in contrast to the *ability* of the *individual* and the *difficulty degree* of the *item*.

Figure 1 shows the results of the diagnostic tests for the *goodness of fit* (*infit* and *outfit*) to the model according to the items examined. It is observed that almost all the items are concentrated within the range of *good adaptability* to the *Rasch model* conditionally on the subject's *ability levels* and on the probability of choosing response methods. Only item 66 " *IT and technological equipment* " is *out of range* with a value greater than 1.3 (*misfit*), therefore, based on the model, item could be poorly worded or extraneous to the construct. Ultimately, the items for which respondents evaluate a general *dissatisfaction* with the *organizational aspects* of the office/service are: item 67 " *Statistical staff assignment* " (1.63) and item 68 "S*taff training*" (1.46), confirming the previously obtained results. Instead, items for which the respondents consider themselves *most satisfied* with the *quality of the service* are the following: item 89 "*Overall office quality*" (-1.28), item 66 " *IT and technological equipment* " (-0.87) even if in the presence of *misfits* and item 69 " *Relationships with staff*". *Item-Subject* map (Fig. 3) shows on a *logit scale* the joint distribution of the *subjects ability* (left side) who respond according to a decreasing scale of required *ability* (*more–less ability*) and positioning of the *items* (right side) on a decreasing scale of *difficulty*. Most of the *subjects* (*X* is represented by a single subject) and the *items* are concentrated in the central part of the graph, showing a substantial balance in *ability* of the respondents and in *difficulty level* of items.

Furthermore, the *Item-Subject* map reveals a *dissatisfaction* and *difficulty* level of the *respondents* in items "*Statistical staff assignment* (item 67)" and "*Staff training* (item 68)" as they differ strongly from the remaining group and higher levels close to *2 logits*. These items would require important corrective interventions.

Figure 1 – *Bubble chart, OUTFIT and INFIT of the items (percentage values)*




```
    xxxxxxxxxx   |
                S|
              xx |
 2              +
                |T 67 Statistical staff assignment
               x|  68 c) Staff training
                |
 1            xx+S
             xxxx|
                M|  70 Processes and workloads
              xx |  71 Organizational structure
 0           xxxx+M
             xxxx|  72 Circulation and clarity of informations
                 |  73 Quality of statistical products and services offered
             xxx |  74 Timeliness of response to users
             xxxx|  66 IT and tech. Equipment
                 |  69 d) Relationships with staff
-1           xx +S
             xx  |  89 Overall office quality
              x  |
             xxx S|T
-2            x +
             xx  |
                 |
                 |
              x  |
-3           xx +
                 |
                 |
                T|
-4            x +
Less ability                    Easy Item
             <less>|<frequ>
```

5. Conclusions

Knowing expectations and perception of statistical referent of one's own structure on the data production and dissemination methods is a prerequisite for improving the relationship between relationship between benefits provided and needs satisfied. The analysis and evaluation of statistical activity, within public and private structures, is an important tool to identify critical points to be corrected and strengths to be promoted and extended to other realities. The transversal reading of the present work offers numerous points of reflection both in terms of knowledge and proposals on

several fronts. The use of the Rasch model has proved useful in methodological terms for in-depth diagnostics of service quality. In fact, the model allows to obtain valid and reliable information on the production and organizational process highlighting any strengths and/or weaknesses. Results of the qualitative survey show appropriate organisational capacity of the statistical offices and statistical referents from data production to analysis. However, the analysis reveals the need to:

- strengthening of structures and services;
- substantial training investment in the staff with suitable qualifications for the production and control of statistical data;
- improvement in the ability to interact with other statistical data providers in terms of engagement and collaboration.

Essentially, the perception of the service quality depends partly on the structural endowment but, in a decisive way, on the human factor in terms of training (acquisition of technical, methodological and specialist knowledge) and interaction (collaboration and participation). In other words, it is argued that perceived quality increasingly depends on organizational *soft variables* (organizational climate, human resources, knowledge) rather than system (structures and/or processes).

References

ANDRICH D. 1988. *Rasch Models for Measurement*, Sage Publications. Newbury Park

BERGMANN K.B., FOSTER E.M., JUSTICE A.W., MATTHEWS V. 1978. *Management of the elderly demented patient in the community*. Brit J Psychiat, n. 132, pp. 441-449

LECARDANE G. 2022. *Sperimentazione per nuove modalità di raccolta dati ambientali nelle città*. Relazione del Progetto di Ricerca Istat RD, Progetti specifici ID 1095. ODS/141/17 e ODS/195/18 del DIRM. Prot. 1780683/22.

LOVAGLIO P.G. 2004. *La stima di risultati clinici con la Rasch Analisys*. STATISTICA, anno LXIV, pp. 127-144

MIPA – SISTAN. 2001. *Modelli organizzativi degli uffici di statistica*. Rapporto finale

PEARLIN L.I. et al. 1990. *Caregiving and the stress process: an overview of concepts and their measures*, Gerontologist, n. 30, pp. 583-594

RASCH G. 1980. *Probabilistic Models for Some Intelligence and Attainment Tests, Danish Institute for Educational Research*. Copenhagen. Expanded edition, Chicago: The University of Chicago Press

RASCH G. 1977. *On specific objectivity: an attempt at formalizing the request for generality and validity of scientific statements*. Danish Yearbook of Philosophy, 14, pp. 58-94

TRIVELLATO U. 2002. *Qualità dell'informazione statistica ufficiale e esigenze informative di Regioni e città*. Relazione alla Sesta Conferenza Nazionale di Statistica. Roma

VITTADINI G., LOVAGLIO P.G. 2001. *The estimate of latent variables in a structural model an alternative approach to PLS.* Second International Symposium on PLS and related methods. Cisia-Ceresta. Montreuil, France. pp. 423-434

WRIGHT B.D., MASTER G.N. 1982. *Rating Scale Analysis, Rasch Measurement*. MESA, Chicago.

Giuseppe LECARDANE, Istat, lecardan@istat.it