WORKING PAPER

# *Statistical methods for evaluating satisfaction with public services*

Giancarlo MANZI & Pier Alda FERRARI

**CIRIEC N° 2014/04**

**CIRIEC activities, publications and researches are realised with the support of the Belgian Federal Government - Scientific Policy and with the support of the Belgian French Speaking Community - Scientific Research.**

**Les activités, publications et recherches du CIRIEC sont réalisées avec le soutien du Gouvernement fédéral belge - Politique scientifique et avec celui de la Communauté française de Belgique - Recherche scientifique.**

**This working paper is indexed and available in SSRN and RePEC**

**Ce working paper est indexé et disponible dans SSRN et RePEC**

**ISSN 2070-8289**

# *Statistical methods for evaluating satisfaction with public services*[*]

Giancarlo Manzi[**] & Pier Alda Ferrari[***]

**Working paper CIRIEC N° 2014/04**

[*] This case study was presented at the Seminar "Public enterprises in the 21st century: Recent advances on public missions and performance – Theory contributions and literature review", VKU Forum, Berlin, February 14-15, 2013, Research Project of CIRIEC International Scientific Commission on Public Services/Public Enterprises on "The Future of Public Enterprise - Mission, performance and governance: Learning from success and failures".

[**] Department of Economics, Management and Quantitative Methods, Università degli Studi di Milano, Italia (Email: giancarlo.manzi@unimi.it).

[***] Department of Economics, Management and Quantitative Methods, Università degli Studi di Milano, Italia (Email: pieralda.ferrari@unimi.it).

## Abstract

*Contrary to private enterprises, public enterprises can be unaware of the impact of their performance when providing services to the public. This is often the case when a small array of choices is in citizens' hands or coercion is the only possibility and a public service must be received and accepted as it is. In these situations where citizens cannot switch to other providers, customer churn cannot occur, or the use of the service cannot be reduced, the assessment of users' satisfaction for public services becomes a very important topic. At the same time, this is also a tricky task, since satisfaction may vary among citizens according to their personal needs and expectations. Using proper statistical methods to assess and explain the level of satisfaction for services provided by public enterprises can be useful to face these issues. In this paper we analyse some of these statistical methods and suggest how to use them to improve citizens' satisfaction.*

**Keywords:** evaluation, services of general interest, statistical indicators, statistical models, meta-analysis, consumers.

**JEL-codes:** C43, D18, H40, L88.

**Acknowledgements:** An earlier version of this paper was presented at the CIRIEC workshop "The Future of Public Enterprise", 14-16 February 2013, Berlin, and we thank those who were present for helpful discussion. We are also grateful to Stefania Scuderi for competent research assistance.

## 1. Introduction

Over the last years the interest in the analysis of customer satisfaction has steadily increased in the private sector. Less interest has been paid to this aspect by public administrations, especially in the case of public services, even though it is precisely in this sector that investigations on customer satisfaction should be more frequently and more accurately performed. In fact, while private companies can be aware of customers' dissatisfaction with a product because, for example, its purchasing decreases, a public enterprise providing a service and operating in a monopoly might well be unaware of the lack of satisfaction if users cannot switch to other providers, refuse or reduce the consumption of the service. Furthermore, a good knowledge of satisfaction for different aspects of the service in connection with the characteristics of the users can suggest a multiple and more satisfactory provision of the service. Therefore a careful evaluation and monitoring of satisfaction through specific surveys and investigation should be particularly welcome in the public sector.

The European Union (EU) has gradually shifted its policy on public sector governance towards the so-called "Europeanization of public services" (Zatti, 2012). This shift has created a constant pressure on Member States to follow new general principles in this area: (i) a faster convergence towards a high-quality service provision; (ii) a standardized provision mode; (iii) a new financial system with limited and regulated transfers, and, wherever possible, (iv) an open encouragement for privatisation and liberalisation, with (v) a special focus on the spread of competition (Bognetti and Obermann, 2012).

Contemporary to the development of this new policy, the regulatory reform process on privatisation and liberalisation started in the 1990's has been viewed by the EU as the main way to improve citizens' well-being, as the liberalisation/privatisation process should imply an increased competition and a greater consumer choice for an improved welfare and a higher satisfaction (Clifton et al., 2012). Therefore, this satisfaction level should constantly be monitored to check if this process effectively works.

But why is monitoring customers' satisfaction with public services so important for the EU policy? Three main reasons may be provided. First, the EU needs to check the outcome of privatisation policies in terms of social welfare. Second, the EU needs to assess the efficiency of public institutions, especially in a context characterized by tight budget constraint for public administrations. Third, the EU needs definitive answers in terms of choosing different forms of organization in the provision of a public service (Fiorio and Florio, 2011; Roland, 2008). To this purpose, the EU has increasingly introduced monitoring instruments to evaluate citizens' and consumers' perception and satisfaction about services of general interest (SGI) in their home countries (Clifton and Diaz-Fuentes, 2010). Satisfaction monitoring tools adopted by the EU and other EU-related institutions are mainly in the form of opinion surveys or portals, such as the Eurobarometer (EB) Survey (European Opinion Research Group, 2002),

the European Quality of Life Survey (Anderson et al., 2009) and the "Your voice in Europe" portal (Sarikakis, 2011).

But an increasing interest towards customer satisfaction data analysis in the public sector would be perhaps beneficial beyond and independently of what the EU is already doing. In this paper, we will draw attention to the most recent statistical methods and models of satisfaction data analysis. In particular, we will focus on the objectives of these analyses, the interpretation of their results and their potential use in public administration. More specifically, after a brief discussion on the problems connected to customer satisfaction data collection, typology and related analysis (Section 2), some dependence models and reduction techniques for customer satisfaction analysis will be presented in Section 3. Section 4 is then devoted to methods for summarizing results from multiple data sets which are useful when information on satisfaction comes from different sources. Finally, Section 5 concludes the paper and outlines some possible future research developments.

## 2. Measuring satisfaction for public services: issues and sources of bias

Public service satisfaction data are usually collected through opinion survey questionnaires containing items on personal judgements or perceptions about public services (European Commission, 2006).

A scale perception bias is sometimes present in responses to this item, especially with sensitive topics (León, Araña, and León, 2013; Tourangeau, and Smith, 1996). Respondents may have different reactions to the same question according to their cultural background, education and environment. For instance, a 5-level Likert scale answer to a question about corruption equal to 3 may denote a high level of corruption in countries where corruption is not a problem, whereas may be a choice to denote the "normality" in countries where corruption is widespread. Therefore, issues of comparison arise in these cases, especially in cross-country analyses. Even if survey researchers have built considerable experience and knowledge on survey respondent behaviour, developing many approaches to solve problematic issues in choice surveys and experiments (McFadden et al., 2005), still there exist significant bias to be handled.

When dealing with satisfaction for public services, answers from the public are self-reported expressed opinions and can be affected by many bias sources: respondents might feel uncomfortable and distressed about revealing their opinion (especially with services like the police service, the prison service, the health service) when they feel that their views are in the minority (Noelle-Neumann, 1974; Ho et al., 2013). Sometimes it could be the case that respondents have a negative attitude towards public service satisfaction surveys and have an interest to under report their satisfaction, being influenced by "not-in-my-backyard" mentality or because they believe that in this way improvements can be obtained more easily, or have an incentive to strategically

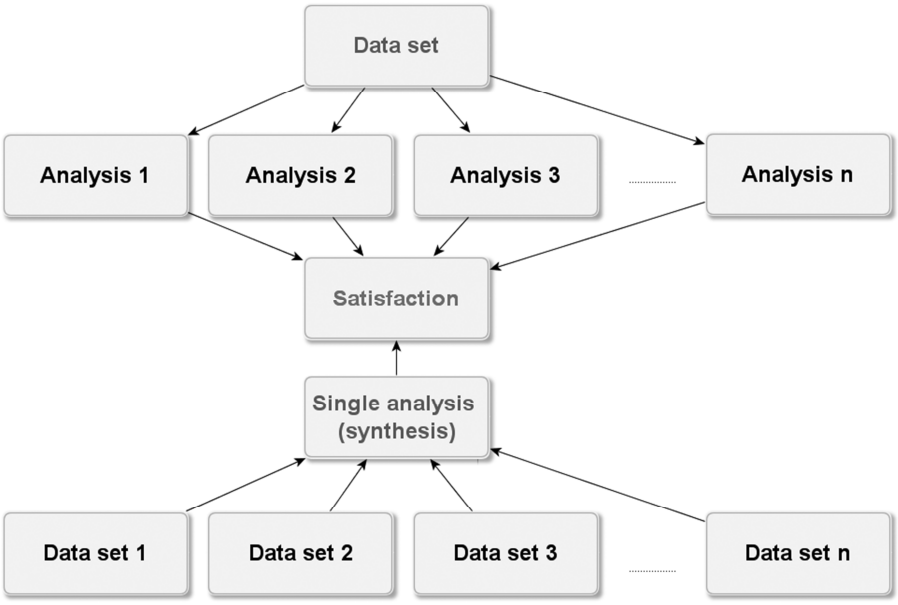
misrepresent their preferences in survey studies, so to influence the policy decision (Wardman, 1988; Ansolabehere & Koniski, 2009). Furthermore, non-response rates can also be high and post-survey validation of results is always hard to perform. A huge stream of literature has investigated the factors linking the refusal to answer the questionnaire to individual characteristics, being the understanding of this relationship a key factor to obtain real preferences in public opinion (Gray et al., 1996; Mannetje et al., 2011; Riphahn and Serfling, 2005).

Nevertheless, the collection of these data and a proper statistical analysis, as we will try to show in what follows, may be useful to measure and compare the level of satisfaction, and to improve the quality of the service, as perceived by users.

One of the most important issues is the nature of the items, and the nature (ordinal/categorical, not numerical) of the resulting variables. In many cases, survey respondents are given a Likert scale (see, Likert [1932]) or a list of ordered categories (see Agresti [2013]) to choose from. In both these cases, labels are used to assess the order of the categories (from the lower to the higher or vice versa), but not their real values, even if numerical as in a Likert scale, and distances between subsequent numerical labels do not reflect a numerical scale. With all these problems in public service opinion surveys, a well-managed statistical analysis is needed. A new stream of successful statistical methods developed to solve these problems has indeed flourished in the last years.

In the next sections we present some statistical methods for the evaluation of satisfaction for public services – focusing our investigation to the most promising proposals of the last years – and sketch some possible outlines of the usefulness of these tools for policy makers. We will categorize these methods into two groups: methods applicable to the same dataset for obtaining different kinds of information, and methods which give insights on specific information on a service by pooling the results from the analysis of different data sets. The first group can be subdivided in model-based methods, composite synthesising indicators, or a mix of them. The second group is essentially formed by meta-analytic procedures. In the first case, the analyses assume the availability of appropriate datasets (large, reliable, and consolidated over time) and appropriate methods of analysis, and stress the interpretative advantages in improving the knowledge on the levels of satisfaction provided by different methods with different objects of analysis applied to the same dataset; the focus is therefore on methods, findings and interpretation. In the second case, the information comes from several but possibly “weak” datasets using a unique method of analysis, and a “good” value about the strength of a specific hypothesized relationship is reached by pooling analyses even coming from not very reliable datasets. This dichotomy among methods to evaluate satisfaction is shown in Figure 1.

**Figure 1 – Approaches to the evaluation of users' satisfaction**




## 3. Some methods to analyse user satisfaction data

There is a huge literature on methods for extracting useful information on satisfaction from survey datasets. Some of these methods are presented in the following subsections. More specifically, subsection 3.1 is dedicated to models which consider satisfaction as a dependent variable, explained by some covariates (model-based methods). In subsection 3.2 methods addressed to obtain a synthetic measure for satisfaction are described, while in subsection 3.3 mixed methods are presented. For each method, its potential and successful application given the available data and objectives of the analysis are critically discussed.

### *3.1. Model-based methods*

These methods rely on statistical models to explain the dependence of the level of satisfaction with regard to a specific service item (for example, the electricity price) on individual (features of survey respondents: gender, level of education, etc.) or contextual variables (features of countries: GDP level, national education expenditure, regulatory index, etc.). Among these models, probit, logit, and linear regression are the most used.

Formally, for all these models individual satisfaction for the *i*-th individual, *i*=1,…,*n*, is intended as a variable $S_i^*$ which is assumed dependent on *k* explicative variables (regressors) according to the following equation:

$$S_i^* = \beta_0 + x_i'\beta + e_i \qquad (1),$$

where $x_i'$ is the k-dimensional vector of the explicative variable values, $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ is the *k*x1 vector of the coefficients of the *k* explicative variables and expresses the influence of the single variables on the level of satisfaction. Statistical models then differ depending on the distribution of $e_i$. If a linear regression is performed, $S_i^*$ is the observed quantified variable and $e_i$ is given a normal distribution with mean 0 and variance $\sigma^2$. Even if this is the most popular model, it does not seem suitable, due to the categorical nature of the dependent variable (Cameron & Trivedi, 2005).

If a probit model is adopted, $e_i$ is assumed distributed according to a standard normal distribution $N(0,1)$ and $S_i^*$ is a latent continuous variable related to a dichotomous observed variable $S_i$, having two categories: $S_i = 1$ (satisfied) and $S_i = 0$ (unsatisfied), so that the following relationship holds:

$$P(S_i = 1|x) = P(S_i^* > 0|x) = \Phi(x_i'\beta - \beta_0) \qquad (2),$$

where $P$ denotes the probability that $S_i^* > 0$ and $\Phi$ is the cumulative distribution function of $N(0,1)$. If an ordered logit model is adopted, the individual satisfaction $S_i^*$ in model (1) is still a latent variable, but $\beta_0 = 0$ and $e_i$ is given a logistic distribution with mean 0 and variance $\pi/\sqrt{3}$. The relationship between the ordered level of the variable $S_i$ with $j$+1 ordered categories and $S_i^*$ is expressed in the following way:

$$\begin{array}{lll} S_i = 0 & \text{if} & S_i^* \le \alpha_1, \\ S_i = 1 & \text{if} & \alpha_1 \le S_i^* \le \alpha_2, \\ \dots. & \dots & \dots \\ S_i = J & \text{if} & S_i^* \ge \alpha_J, \end{array} \qquad (3),$$

where $\alpha_1 < \alpha_2 < \cdots < \alpha_J$ are unknown thresholds. Whatever is the statistical model: probit, logit, or linear regression, the β coefficient vector is usually estimated by maximum likelihood method based on the $e_i$ specific distribution.

The objective of all these methods of analysis is to detect the influence of some variables (regressors) on the level of satisfaction. They are usually based on responses to a specific item (observed variable). They are applied also in the context of SGI, and more specifically for utilities. For example, Jilke, and Van de Walle (2013) explicitly model the responses to questions on complaints about some aspects of the provision of public utilities against age and education through a binary logistic regression on EB survey data. They focus on citizens' complaints about SGI as a surrogate for dissatisfaction and as a mean for

amplifying citizens' voice. Their work is therefore aimed at finding socio-economic determinants of a two-track complaint behaviour among citizens: those who are assertive and best informed and receive high quality services and those who, by virtue of their socio-economic weakness, are marginalized and made vulnerable. Fiorio and Florio (2011) aim at addressing the question "are European consumers happy with the price they pay for electricity supply services after two decades of reforms?", being interested in the correlation between satisfaction and regulatory reforms in the EU, and using the random-effects probit model recalled in (2) where the latent level of individual satisfaction for each aspect of the services has to be explained by a set of socio-economic variables (i.e., gender, occupation, etc.), country macro-economic variables (i.e. GDP level, population density, etc.) and the level of privatisation and market regulation (Conway and Nicoletti, 2006). A similar analysis for telecommunications is carried out in Bacchiocchi, Florio, and Gambaro (2011). Fiorio et al. (2007) adopt an ordered logit model defined with thresholds (3) and find different level of satisfaction for gas supply, fixed telephone and electricity services in each country, depending on many explicative variables.

Clifton, Díaz-Fuentes, and Fernández-Gutierrez (2014) highlight the strong connection between consumers' stated and revealed preferences by contrasting them through five empirical probit models aiming at revealing the relationship between satisfaction for public utilities and socio-economic variables. As their approach is inspired by Behavioural Economics, their focus is intended to find the reasons why certain categories of vulnerable consumers behave differently from their peers. They use data from EB for stated preferences on satisfaction and from Household Budget Surveys collected by Eurostat for revealed preferences on expenditure.

Among linear regression analyses on satisfaction for SGI, an example is in Rahmqvist and Bara (2010) who deal with the relation of respondents' characteristics, and perceived quality dimensions of health care to overall patient satisfaction in out-patient hospital care.

### *3.2. Synthetic measures and composite indicators*

These methods address the problem of building up a synthetic measure of satisfaction by combining different aspects of a service or different services and, at the same time, by providing suitable weights for every single aspect or service. For example, Ferrari, Annoni, and Manzi (2010) proposed a synthetic indicator for consumer satisfaction based on Nonlinear Principal Component Analysis (NPCA), a method introduced by Gifi (1990) and Michailidis and De Leeuw (1998) for dimensional reduction. This approach expresses the level of satisfaction as a linear combination of observed ordinal variables, whose categories are optimally quantified and the variable coefficients of the combination are also optimally determined. Formally, the measurement of

satisfaction is obtained through a latent variable, whose scores for *n* individuals and *m* ordinal variables are given by:

$$\mathrm{x} = \frac{1}{m}\sum_j \mathrm{G}_j\, \mathrm{q}_j \beta_j = \frac{1}{m}\sum_j \mathrm{t}_j\, \beta_j, \qquad (4)$$

where x is the $n \times 1$ vector of the scores, $\mathrm{G}_j$ an indicator matrix of categories of the *j*-th variable, $\mathrm{t}_j$ the $n \times 1$ vector of the quantified variable *j*, obtained by substituting the observed ordinal categories of the variable with an $n \times 1$ vector of optimal quantifications $\mathrm{q}_j$, $\beta_\mathrm{j}$ the coefficient of $\mathrm{t}_j$ and can be read as the importance of the variable in determining the level of satisfaction. This approach allows, like the standard Principal Component Analysis (PCA), to find the level of satisfaction for each individual, passing from *m* variables to a synthetic univariate indicator through a combination of the original *m* variables. Contrary to PCA, where category values are predetermined *a priori* with fixed distances as in Likert scale, NPCA category quantification allows for an optimal assignment of the category values taking into proper account measurement levels and nonlinearity (Ferrari and Barbiero, 2011).

Optimal $\mathrm{q}_j$ and $\beta_\mathrm{j}$ in (4) are obtained by minimizing the loss of information due to dimensional reduction, given by the following sum of squared distances:

$$\sigma = \frac{1}{m}\sum_{j=1}^{m}\left(\mathrm{x} - \mathrm{G}_j \mathrm{q}_j \beta_j\right)^{\mathrm{T}}\left(\mathrm{x} - \mathrm{G}_j \mathrm{q}_j \beta_j\right).$$

Another important method in this area is the Rasch model (RM). It was introduced by Rasch (1960) to analyse tests on a subject's ability. These tests are based on a set of items and the model describes the probability of giving a specific answer to an item as depending on two factors: the *relative* subject *ability* and the item's intrinsic *difficulty*. In its simplest formulation, RM expresses the probability of having an answer $x_{ij} = 0$ (wrong) or $x_{ij} = 1$ (correct) from subject *i* having ability $\theta_i$ when he meets item *j* of difficulty $\beta_j$. This probability is:

$$P\left[X_{ij} = 1 \middle| \theta_i, \beta_j\right] = \frac{\exp\{\theta_i - \beta_j\}}{1 + \exp\{\theta_i - \beta_j\}}.$$

The greater is the ability or lower is the difficulty, the greater is the probability of $x_{ij} = 1$.

In the context of user satisfaction, the two factors *ability* and *difficulty* become *customer satisfaction* (ability) and *item lack of quality* (difficulty),

respectively. Note that low ability ($\theta_i$) corresponds to low level of satisfaction, while low difficulty ($\beta_j$) means high quality of the item.

RM can be extended to the polythomous case in different ways with similar meaning and interpretation (De Battisti, Nicolini, and Salini, 2012). For example, in the case of item *j* having $m_j$ ordered categories $(0,1 \dots, (m_j - 1))$, we can set $\dots, (m_j - 1)$ thresholds $\beta_{jk} = \beta_j + \tau_{ik}$, intended as the points where two adjacent answer categories (for example '*Good*' and '*Very good*') have the same probability to be chosen, being $\tau_{ik}$ the deviations of $\beta_{jk}$ about the mean $\beta_j$, $\sum_{k=1}^{(m_j-1)} \tau_{jk} = 0$. The probability of subject *i* responding to item *j* through answer categories $x = 0,1 \dots, (m_j - 1)$ is:

$$P[X_{ij} = x] = \frac{\exp\{\kappa_{jx} + x(\theta_i - \beta_j)\}}{\sum_{x=0}^{(m_j-1)} \exp\{\kappa_{jx} + x(\theta_i - \beta_j)\}},$$

where $\kappa_{jx} = -\sum_{k=1}^{x} \tau_{ik}$, $\kappa_{j0} = \kappa_{j(m_j-1)} = 0$ (Ferrari and Salini, 2011).

With this approach the objective is double: we can compare the level of satisfaction among individuals similarly to the scores in NLPCA, but, at the same time, evaluate the quality of different services or aspects of a service.

The complementary use of RM and NPCA thus allows for obtaining two rankings of the items of the service, one based on its perceived quality (via RM), and the other based on its importance (via NPCA), significantly enriching the interpretation of the results.

### *3.3. Other methods*

With regard to users' satisfaction, there are other proposals that combine the methods described above or are not included in the previous classification. For example, Ferrari et al. (2011) adopted a two-step procedure to analyse user satisfaction data which can be categorized as a mix of synthetic measure and model-based approach. With regard to a model of the type (1), in the first step they built a synthetic indicator via NPCA in order to obtain quantitative values for the continuous latent variable "satisfaction" (measurement of the variable $S_i^*$ on the left-hand side in equation (1)). In the second step a dependence model is formulated to explain the level of satisfaction in relation to covariates (detection of explicative variables in the terms on the right-hand side of equation (1)). More specifically a Multilevel Model (ML) is used for detecting personal and environmental characteristics[1].

[1] Multilevel models have been extensively used in public service satisfaction analysis, especially in medical care (see, for example, Sixma et al., 1998). But in general they are used only with satisfaction scores as dependent variables, not with synthetic measures of satisfaction.

Salini and Kennet's method (2009) considers Bayesian network models to analyse the connection between different aspects of a service in the case of EB data. Furthermore, several comparative and ranking analyses have been performed in the last years in order to evaluate the opinion of Europeans on public services. For example, Annoni (2007) uses Hasse diagrams and POSAC methods to perform a ranking analysis on the EU15 countries through EB data; Annoni and Brüggemann (2009) show a way to rank EU countries through partial order theory; Clifton and Díaz-Fuentes (2010) analyse EB survey data by averaging responses on satisfaction for country comparison.

Other methods of analysis focus on the control of the response bias which may arise in satisfaction surveys and, more in general, in opinion surveys, either depending or not on the number of questions in the questionnaire. If the questionnaire is long and the cognitive task is hard, then problems such as learning effects, boredom and anchoring to earlier tasks may occur. But even if the questionnaire is short and well-conceived, individual fixed effects may emerge, as explained in McFadden (1986). Therefore, to avoid measurement error problems and to circumvent this issue, it is possible to exploit the fact that similar questions regarding different aspects of services are repeated over time in different survey waves, and a specific-individual fixed effect introduced in the model could adjust for this source of bias. A possible solution to the problem is given by Grassi and Puglisi (2008) who stack the observation about reported satisfaction over these different dimensions and run a regression model with individual-specific fixed effects. They assume that many of the biases affecting the respondents in satisfaction surveys, although specific to the individual, are common across similar questions being asked, and then can be captured by an individual fixed effect.

Finally, it is worth mentioning other possible methods still not applied to EU data which use statistical models to estimate the relationship between the latent variable and the manifest variables through structured equation models obtained via the Partial Least Squares method - PLS (Wold, 1982, see also the extensive presentation in Tenenhaus et al., 2005) or the LISREL method - LInear Structured RELationship (Jöreskog 1970; O'Brien and Homer, 1987).

A summary of some of the methodologies most used in the last few years to evaluate the user satisfaction and their characteristics are reported in Table 1. For a comparison of their applications to EU data see Ferrari and Manzi (2014).

**Table 1 – Summary of the characteristics of some of the main methodologies available to measure satisfaction for public services**

| Methodology | Category | Characteristics |
|---|---|---|
| Logit, probit and linear regression | Model-based | Dependence analysis. Satisfaction can be explained by some hypothesized determinants. Only one response variable (one item at a time) is considered. |
| Multilevel models (ML) | Model-based | Dependence analysis. Satisfaction is still explained by some hypothesized determinants as above, but at different levels, e.g. at individual and country levels. Useful for hierarchical data. Only one response variable (one item at a time) is considered. |
| Nonlinear Principal Component Analysis (NPCA) | Synthetic measures & composite indicators | The focus is on measurement. More items (aspects) of satisfaction can be taken into consideration and weighted accordingly. Level of satisfaction, importance of items, and optimal quantifications of answers are determined. |
| Rasch Model (RM) | Synthetic measures & composite indicators | The focus is on measurement. Level of satisfaction and quality of items (aspects) of satisfaction can be assessed. |
| RA + NPCA | Synthetic measures & composite indicators | The complementary use of RM and NPCA allows for the joint representation of quality and importance of items in order to provide a set of indicators to decision makers. |
| NPCA + ML | Synthetic measures & composite indicators + models | Both synthesis and explanatory analyses are considered. The ML model is applied on a synthetic measurement of satisfaction obtained via NPCA. |
| POSAC, Partial order | Ranking analyses | Allows a ranking of countries according to satisfaction. |
| Bayesian networks | Model-based | Models of cause and effect. Only one response variable is processed at a time. |
| Averaging | Synthetic and comparative tools | Immediate synthetic indicator. Comparative analysis based on conditional mean values of observations. |

## 4. Methods to simultaneously analyse multiple datasets

In the previous section we considered methods to analyse datasets with individual information on satisfaction coming from surveys carried out by major pollsters. In most cases, only aggregate findings from multiple studies/datasets are available with no access to individual data, and analysis can be developed only through a systematic review of the literature, aiming at summarizing the empirical evidence emerging out of these studies. Meta-analysis (MA) can therefore be useful in that it allows the pooling of the different satisfaction findings with regard to the variety of techniques used to get them. Sometimes

judgments from experts should also be considered to correct for bias in the studies selected for MA[2].

MA, originated in the social sciences but soon became a fundamental technique in other fields, could produce interesting results in the satisfaction analysis context. Hunt (1997), and Rosenthal and Dimatteo (2001) point out some criticism that has always been raised in many research fields about the meta-analytic approach. They state that "MA is like to taking apples and oranges and averaging over their weights, sizes flavour and shelf life. Too different are apples and oranges to be put together". Paradoxically, the real strength of MA lies exactly in this criticism: if one wants "to generalize about fruit rather than about apples or oranges, then it is correct to consider both apples and oranges altogether" (Rosenthal and Dimatteo, 2001).

In the case of satisfaction, this "*generalization about fruit*" applies naturally: apple or oranges can be items of satisfaction or different measurements of satisfaction, but finally, MA helps to put everything together, giving a definitive evaluation of the *fruit-satisfaction*. MA can correct adequately the possible bias emerged in the studies, giving them more or less importance in the pooling exercise. As such, MA can be considered a valid alternative to the methods used to model satisfaction on a single dataset when individual or information is not available or is fragmented.

In the subsections below we first briefly recall the standard meta-analytic approach to summarize information on public satisfaction from multiple studies or datasets in the particular case of the correlation coefficient, and then describe a non-standard meta-analytic approach recently introduced, aimed at adjusting for multiple sources of bias (subsection 4.1). An example of application of the latter method to assess satisfaction for SGI is provided in subsection 4.2.

### *4.1. Standard and non-standard methods of MA*

Very few MAs having satisfaction for public services as the main objective have been performed so far. One of these examples is in Hall and Dornan (1990) where the authors performed a standard meta-analysis of product-moment correlation coefficients with the aim of examining the relationship of patients' socio demographic characteristics to their satisfaction with public medical care. Meta-analysis of correlation coefficients *r*s is commonly performed by first converting them into Fisher-transformed *z* scores:

$$z = \frac{1}{2} log\left(\frac{1+r}{1-r}\right),$$

[2] This bias could concern, for example, the choice and misspecification of a model for the determinants of satisfaction, possible omitted variables, the quality of the sample, the particular technique used, etc.

which have an approximate normal distribution with standard error $\sqrt{\frac{1}{(n-3)}}$, where *n* is the study sample size. In that analysis, *z*s are pooled across studies with weights proportional to the standard errors and are back-transformed to *r*, and Hall and Dornan found that a higher satisfaction is significantly associated with greater age and less education, and marginally associated with being married and having higher social status.

In primary studies (i.e. in studies selected for the MA) satisfaction can have been measured in many different ways. This introduces other types of biases in the analysis. First, studies can differ in design as well as in participants, exposures or outcomes. We refer to this bias as "statistical bias" or "*internal bias*" (i.e. lack of statistical rigour in performing the study). Second, if a given policy target protocol has been agreed, studies may be more or less relevant to this target and can or cannot agree with this protocol. We refer to this bias as "relevance bias" or "*external bias*" (i.e. lack of relevance of the study in view of the policy targets). To explain this latter point let's suppose that an agreed protocol on "better standards" to evaluate satisfaction states, for example, that satisfaction is better evaluated if an item on the perception of quality of the service is present. Studies may or may not have such information: studies without the information on quality are not excluded from the MA, but are considered more "relevance biased" than the others. This setting requires a different approach where statisticians and policy experts are called together to evaluate these biases, and a model to incorporate these elicitations is thus required.

A proposal in this direction is introduced in Turner et al. (2009). They assume that potential internal and external biases act additively and proportionally on the estimates. To elicit these biases a panel of expert meets and discusses first jointly and then assess independently the amount of bias they believe is present in the studies. Each assessor completes checklists for sources of biases evaluating them on a shared scale. The number of different types of bias is usually defined in advance.

More formally, this method works as follows. Let's have *k* selected studies forming a MA, in each of which the study-specific quantity $\theta_i$ has to be estimated, and the statistic $y_i$ (*i*=1,…,*k*) is used as its estimator. Let $\tau^2$ be an additional variance parameter to allow for unexplained between-study heterogeneity. Let's have $J^I$ sources of internal bias in the estimation process. Then, for each study *i*, an *additive model* to incorporate internal sources of bias $j = 1, \dots, J^I$ is formulated. A bias component $\delta_{ij}^I$, having distribution $\delta_{ij}^I \sim f(\mu_{ij}^I; {\sigma_{ij}^I}^2)$, is supposed to have an additive effect on $y_i$ in the following way:

$$y_i \sim f(\theta_i + \mu_i^I; \left[s_i^2 + {\sigma_i^I}^2\right]), \quad (5)$$

where $\mu_i^I = \sum_{j=1}^{J} \mu_{ij}^I$ is the total internal bias for study *i* elicited by the expert, $s_i^2$ is the sampling variance assumed known, and ${\sigma_i^I}^2 = \sum_{j=1}^{J} {\sigma_{ij}^I}^2$

Similarly, an *additive* model to incorporate external sources of bias $j = 1, \dots, J^E$ is assumed and expressed as follows:

$$y_i \sim f(\theta_i + \mu_i^E, s_i^2 + \tau^2 + {\sigma_i^E}^2), \quad (6)$$

where $\mu_i^E = \sum_{j=1}^{J} \mu_{ij}^E$ is the total external bias for study *i* elicited by the expert, and ${\sigma_i^I}^2 = \sum_{j=1}^{J} {\sigma_{ij}^E}^2$, being the external bias $\delta_{ij}^E$ distributed as $\delta_{ij}^E \sim f(\mu_{ij}^E, {\sigma_{ij}^E}^2)$.

Combining together models (5) and (6) a general additive model is obtained:

$$y_i \sim f(\theta_i + \mu_i^E + \mu_i^E, s_i^2 + \tau^2 + {\sigma_i^I}^2 + {\sigma_i^E}^2). \quad (7)$$

*Proportional models* corresponding to models (5), (6) and (7) are obtained in a similar way. Additional and proportional models are finally combined together to get the final bias-adjusted meta-analysis results. See Turner et al. (2009, pp. 37-41) for further details.

We believe that this approach can be used in MA on satisfaction *in general*, and on satisfaction with public services in particular.

### *4.2. An application to patients' satisfaction data*

Suppose policy makers are interested in knowing what is the relationship between the level of satisfaction of patients and the competence/efficacy/performance of medical and nursing staff in public hospitals and only aggregate findings are available. An MA with the methodology by Turner et al. can be conducted in the specific framework of the evaluation of satisfaction. The steps to be performed can be summarized as follows:

(i) Select the primary studies containing measures of relationship between medical/nurse competence/efficacy/performance and satisfaction with standard MA selection techniques; (ii) list all the sources of (statistical or relevance) bias possibly present in the studies; (iii) ask experts to quantitatively elicit such biases for each studies and how they affect study results; (iv) incorporate such elicitations to adjust study estimates with the methodology used to get model (7) and the following generalization.

To illustrate this procedure, we conducted a pilot *adjusted* meta-analysis in the area of satisfaction for medical services in hospitals. We were interested in finding studies where the relationship between patient satisfaction and efficiency/competence of medical and nursing staff was evaluated through the Pearson's linear correlation coefficient. Our strategy to obtain the final set of studies to be included in the meta-analysis aimed at first browsing the most important internet bibliographic databases on public health management (Medline, Scopus, Web of Science, Google Scholar, etc.), and then manually searching in the most important journals in health management and health economics from 2009 to 2011. We also searched in the reference sections of the identified studies and among other publications of the authors for further work on the subject. Examples of key-words used for this search are the following or a combination of them: "patient satisfaction", "satisfaction for health care", "hospital". The inclusion criterion comprises the presence of a correlation coefficient between patient satisfaction and variables related to hospital staff competence. Eligible studies were those with survey data containing self-assessments on satisfaction for staff performance and services in hospitals. Studies containing regression models with satisfaction measures as dependent variables and variables on efficiency/competence of hospital staff as independent variables were excluded. This search strategy produced a collection of 24 studies reported in Table 2.

## Table 2 – Studies in the meta-analysis

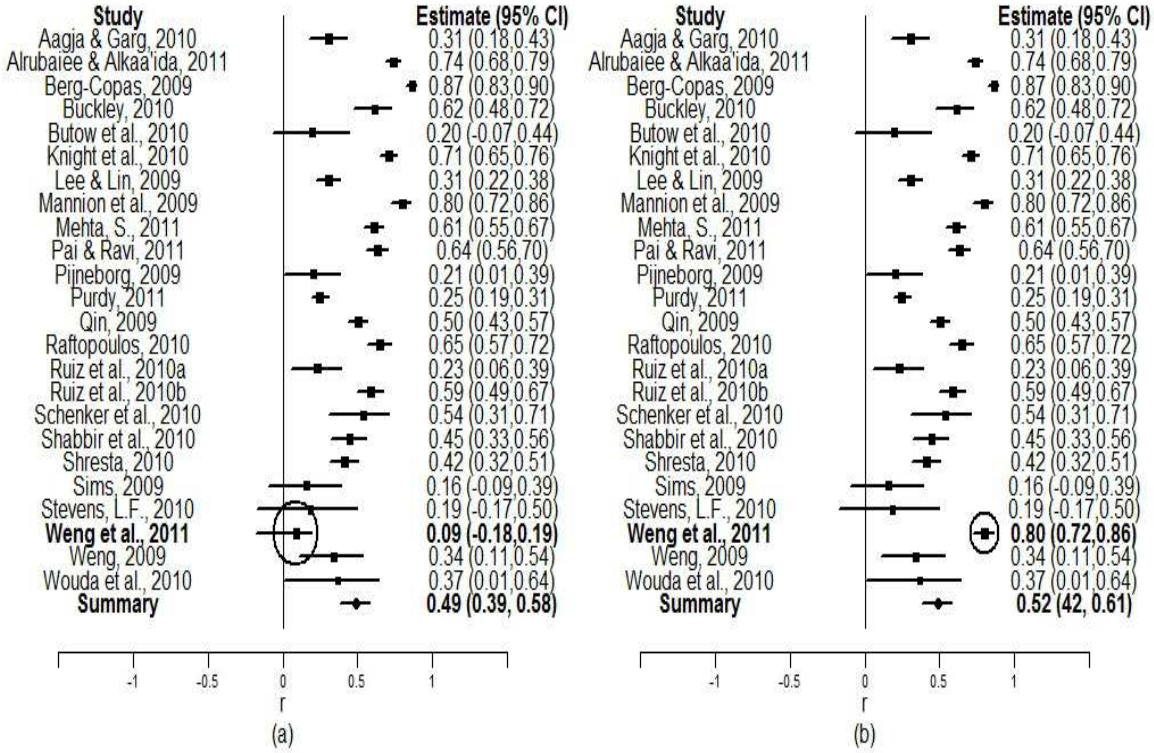
| Study | Year | Geography | Satisfaction variable | Related variable | Sample size | Rho value | Rho type[3] |
|---|---|---|---|---|---|---|---|
| Aagja and Garg | 2010 | India | Overall patient satisfaction in a public hospital | Competence and efficacy in medical services | 200 | 0.31 | direct |
| Alrubaiee and Alkaa'ida 2011 | 2011 | Jordan | General satisfaction in two private hospitals | Reliability-Empathy-Assurance (SERVQUAL) | 290 | 0.74 | averaged |
| Berg-Copas 2009 | 2009 | U.S.A. | Overall trauma patient satisfaction in the Wesley Medical Center, Wichita, Kansas, U.S.A. | Medical, nursing and staff care and efficacy | 209 | 0.87 | averaged |
| Buckley 2009 | 2010 | New Zealand | Overall patient satisfaction with the emergency department nursing in a regional hospital (postal survey) | Nurse caring performance | 100 | 0.62 | direct |
| Butow et al., 2010 | 2010 | NZ & Australia | Satisfaction with doctor consultation and skills | Doctors' care and concern | 55 | 0.20 | averaged |
| Knight, Cheng, and Lee 2010 | 2010 | Australia | Overall client satisfaction with outpatient physiotherapy care in 25 clinics | Therapist's behaviour in discussing negative and positive aspects of treatment | 312 | 0.71 | direct |
| Lee and Lin 2009 | 2009 | Taiwan | Diabetes patient trust in the physician assessed in one regional hospital and one district hospital | Adherence to physician's prescriptions and outcome expectations | 480 | 0.31 | averaged |
| Mannion et al. 2009 | 2009 | Switzerland | Overall satisfaction with surgery in a clinic and a hospital (patients after knee arthroplasty surgery) | Global treatment outcome | 120 | 0.80 | direct |
| Mehta 2011 | 2011 | India | Patient satisfaction in various hospitals | Medical, nursing and staff care and efficacy | 400 | 0.61 | averaged |
| Pai and Ravi 2011 | 2011 | India | General satisfaction in a private hospital | Doctors', nurses' and dieticians' care and concern | 257 | 0.64 | averaged |
| Pijnenborg 2009 | 2009 | Netherlands | Overall satisfaction with hospital, doctors and nurses | Expectations, medical empathy and quality | 102 | 0.21 | averaged |
| Purdy 2011 | 2011 | Canada | Overall in-patient satisfaction for nursing care in medical and surgical units from 21 big hospitals (more than 70 beds) | Nurse-assessed quality of care | 1005 | 0.25 | direct |
| Qin 2009 | 2009 | U.S.A. | Overall patient satisfaction with care providers (web survey) | Quality & expectations | 485 | 0.50 | averaged |
| Raftopoulos 2010 | 2010 | Greece | Overall satisfaction with primary care in a variety of primary health care settings (outpatient setting, home care) service) | Medical and nursing care and efficacy | 212 | 0.65 | averaged |
| Ruiz et al. 2010a | 2010 | Spain | Patient satisfaction for glaucoma therapy in three locations (patients recruited by specialized clinicians in hospitals) | Competence and efficacy in medical services | 124 | 0.23 | averaged |
| Ruiz et al. 2010b | 2010 | Spain | Patient satisfaction with dermatological treatment of hand eczema in 18 hospitals | Medical care | 213 | 0.59 | direct |
| Schenker et al. 2012 | 2011 | U.S.A. | Patient surrogates trust after 7 days in physicians in an intensive unit in a University Hospital | General trust in health care systems | 50 | 0.54 | direct |
| Shabbir, Kaufmann, and Shehzad 2011 | 2010 | Pakistan | Patient satisfaction in public and private hospitals | Trust & service quality (SERVQUAL) | 186 | 0.45 | averaged |
| Shrestha 2010 | 2010 | Indonesia | Out-patient satisfaction for access to primary care in three primary health care centers | personal treatment and medical ability | 300 | 0.42 | averaged |
| Sims 2009 | 2009 | U.S.A. | Trust in physician (chronic pain patients from 4 chiropractic offices in two different southwestern towns | Physician's empathy | 62 | 0.16 | averaged |
| Stevens 2010 | 2010 | U.S.A. | HIV-positive racial/ethnic minority patient satisfaction for doctors, nurses and administrative staff of two urban, community medical clinics | Trust in doctors | 33 | 0.19 | averaged |
| Weng et al. 2011 | 2011 | Taiwan | Relationship between patient satisfaction and doctor emotional intelligence, burn out and job satisfaction | Total emotional intelligence by doctors | 110 | 0.09 | direct |
| Weng 2009 | 2009 | Taiwan | Patient satisfaction at initial visit and after two weeks with surgeons and internists | Medical efficacy and empathy | 67 | 0.34 | averaged |
| Wouda et al. 2011 | 2010 | Netherlands | Patient satisfaction with physicians' working in a university hospital | Medical care | 30 | 0.37 | direct |

[3] "Averaged" means that the correlation coefficient entering the meta-analysis has been averaged across multiple correlation coefficients presented in the study.

Results of a standard MA performed on the selected studies are reported in Figure 2a where for each study the confidence interval for *r* is reported. It can be noted that all reported correlation coefficients are positive and the pooled Pearson's correlation coefficient is equal to 0.49, meaning a certain positive relationship between satisfaction and staff efficiency/competence. The overall *r* is also significantly different from 0.

Since we believed that one study (Weng et al., 2011) presented symptoms of internal and external bias, an elicitation process was performed. The reasons for adjusting this study for bias concerned the lack of statistical rigour (specifically about the sampling used), the measurement used to assess satisfaction, and the related variable (an "emotional intelligence"), all factors leading to the underestimation of the correlation coefficient. By applying the models from (5) to (7) for assessing elicitations (an *ad-hoc* software was used to this purpose[4]), a larger level of correlation was obtained in the study considered, affecting the overall effect sizes which increased from 0.49 to 0.52 (see Figure 2b).

**Figure 2 – (a) Meta-analysis results of the relationship between satisfaction and medical, nursing and staff competence (b) Adjusted meta-analysis results of the relationship between satisfaction and medical, nursing and staff competence: adjusted study is Weng et al. (2011)**




[4] A Java stand-alone version of this software is available on request.

## 5. Conclusion and future developments

This paper has presented some recent non-standard methodologies to evaluate users' satisfaction with public services and discussed their potential.

This has given us the opportunity to highlight that these methods do not only provide us with a measure of the level of satisfaction for public services, and show how it can vary in relation with different factors, but also allow us to evaluate and compare specific services or aspects of a service.

We have tried to highlight the practical utility of these statistical analyses both in presence of a suitable dataset and in the case when such availability is scarce but pieces of information can be extracted by pooling similar analyses whose results can be adjusted according to the different amount of bias they contain. All the presented methods are helpful to improve the efficacy of policy intervention/action, and can be adopted as tools for public managers to better understand citizens as public service recipients.

The methods proposed here are prone to further development both theoretically and in applications. Our future commitments will be centred on developing IT computational tools to apply some of these methods simultaneously or extend the application of these techniques to different datasets in multiple contexts. This will be addressed to give public managers the possibility to be provided with easy and convenient instruments for immediate decision implying a more central role for citizens in public enterprises' decision making.

Our hope is that in the next future increasing resources will be invested to reinforce citizen satisfaction surveys and to develop new methodologies to analyse these data, and methods to implement them in order to support the decision process of public enterprise.

## References


Aagja, P.G. and R. Garg, 2010, "Measuring Perceived Service Quality for Public Hospitals (PubHosQual) in the Indian Context", *International Journal of Pharmaceutical and Healthcare Marketing*, 4(1): 60-83.

Agresti, A., 2013, *Categorical Data Analysis*, 3rd Edition, New York: Wiley.

Alrubaiee, L. and F. Alkaa'ida, 2011, "The Mediating Effect of Patient Satisfaction in the Patients' Perceptions of Healthcare Quality-Patient Trust Relationship", *International Journal of Marketing Studies*, 3(1): 103-127.

Anderson, R., B. Mikuliç, G. Vermeylen, M. Lyly-Yrjanainen and V. Zigante, 2009, *Second European Quality of Life Survey*, Dublin: Eurofound, http://www.eurofound.europa.eu/publications/htmlfiles/ef0902.htm.

Annoni, P., 2007, "Different Ranking Methods: Potentialities and Pitfalls for the case of European Opinion Poll", *Environmental and Ecological Statistics*, 14: 453-471.

Annoni, P. and R. Brüggemann, 2009, "Exploring Partial Order of European Countries, *Social Indicators Research*", 92: 471-487.

Ansolabehere, S. and D.M. Koniski, 2009, "Public Attitudes Toward Construction of New Power Plants", *Public Opinion Quarterly*, 73(3): 566-577.

Bacchiocchi, E., M. Florio and M. Gambaro, 2011, "Telecom Reforms in the EU: Prices and Consumers' Satisfaction", *Telecommunications Policy*, 35: 382-396.

Berg-Copas, G. M., 2009, "Predicting Global Satisfaction Ratings of Quality Health Care among Trauma Patients: Testing a Structural Equation Model", PhD diss., Wichita State University.

Bognetti, G. and G. Obermann, 2012, "Local Public Services in European Countries: Main Results of a Research Project by CIRIEC International", *Annals of Public and Cooperative Economics*, 83(4): 485-503.

Buckley, C. A., 2009, "Consumer Satisfaction with Emergency Department Nursing: A Descriptive Correlational Study", Master diss., University of Wellington.

Butow, P., I. Juraskova, S. Chang, A. Lopez, R. Brown and J. Bernhard, 2010, "Shared Decision Making Coding Systems: How Do They Compare in the Oncology Context?", *Patient Education and Counseling*, 78: 261-268.

Cameron, A.C. and P.K. Trivedi, 2005, *Microeconometrics. Methods and Applications*, Cambridge: Cambridge University Press.

Clifton, J. and D.D. Díaz-Fuentes, 2010, "Evaluating EU Policies on Public Services: a Citizens' Perspective", *Annals of Public and Cooperative Economics*, 81(2): 281-311.

Clifton, J., D.D. Díaz-Fuentes, M. Fernández-Gutierrez, O. James, S. Jilke, and S. Van de Walle, 2012, "Regulatory Reform for Services of General Interest and Trends in Citizen Satisfaction", European Commission, European Research Area.

Clifton, J., D. Díaz-Fuentes, M. Fernández-Gutierrez, 2014, "The Impact of Socio-Economic Background on Satisfaction: Evidence for Policy-Makers", Unpublished manuscript.

Conway, P. and G. Nicoletti, 2006, "Product Market Regulation in Non-manufacturing Sectors in OECD Countries: Measurement and Highlights", OECD Economics Department Working Paper.

De Battisti, F., G. Nicolini and S. Salini, 2012, "The Rasch Model", In Kenett, R.S. and S. Salini (eds.), *Modern Analysis of Customer Surveys: with Applications Using R.*, Chichester: Wiley.

European Commission, 2006, *Special Eurobarometer: Services of General Interest*, Brussels: European Commission.

European Opinion Research Group, 2002, *Eurobarometer 58 – Special Edition, Consumers' Opinions about Services of General Interest*, Brussels: European Opinion Research Group.

Ferrari, P. A., P. Annoni and G. Manzi, 2010, "Evaluation and comparison of European countries: public opinion on services", *Quality & Quantity*, 44: 1191-1205.

Ferrari, P. A. and A. Barbiero, 2011, "Nonlinear Principal Component Analysis", In Kenett, R.S. and S. Salini (eds.), *Modern Analysis of Customer Surveys: with Applications Using R.*, Chichester: Wiley.

Ferrari, P. A., P. Pagani and C. V. Fiorio, 2011, "A Two-Step Approach to Analyze Satisfaction Data", *Social Indicators Research*, 104: 545-554.

Ferrari, P. A. and S. Salini, 2011, "A Complementary Use of Rasch Models and Nonlinear Principal Components Analysis in the Assessment of the Opinion of Europeans About Utilities", *Journal of Classification*, 28: 53-69.

Ferrari, P. A. and G. Manzi, 2014, "Citizens evaluate public services: A critical overview of statistical methods for analysing user satisfaction", *Journal of Economic Policy Reform*, DOI: 10.1080/17487870.2014.909313. (Accepted).

Fiorio, C. V., M. Florio, S. Salini and P. A. Ferrari, 2007, European Consumers' Attitudes on Services of General Interest: Accessibility, Price and Quality, Working Paper 2007-04, DEAS University of Milan.

Fiorio, C. V. and M. Florio, 2011, "Would You Say that the Price You Pay for Electricity Is Fair? Consumers' Satisfaction and Utility Reforms in the EU15", *Energy Economics*, 33: 178-187.

Gifi, A., 1990, *Nonlinear Multivariate Analysis*, New York: Wiley.

Grassi, S. and R. Puglisi, 2008, "Regulation and Consumer Satisfaction from Public Services: An Individual Fixed Effects Approach", DEAS Working Paper 2008-21.

Gray, R., P. Campanelli, K. Deepchand and P. Prescott-Clarke, 1996, "Exploring Survey Non-Response: The Effect of Attrition on a Follow-Up of the 1984-85", *Journal of the Royal Statistical Society. Series D (The Statistician)*, 45(2): 163-183.

Hall, J.A. and M.C. Dornan, 1990, "Patient sociodemographic characteristics as predictors of satisfaction with medical care: a meta-analysis", *Social Science & Medicine*, 30(7): 811-818.

Ho, S. S, V. H.-H. Chen and C. C. Sim, 2013, "The spiral of silence: examining how cultural predispositions, news attention, and opinion congruency relate to opinion expression", *Asian Journal of Communication*, 23(2), 113-134.

Hunt, M., 1997, *How Science Takes Stock*, New York: Sage.

Jilke, S. and S. Van De Walle, 2013, "Two track public services? Citizens' voice behaviour towards liberalized services in the EU15", *Public Management Review*, 15(4), 465-476.

Jöreskog, K. G., 1970, "A General Method for Analysis of Covariance Structure. *Biometrika*, 57(2): 239-251.

Kenett, R. S. and S. Salini, 2012, *Modern Analysis of Customer Surveys: with Applications Using R*., Chichester: Wiley.

Knight, P. K., A. N. Cheng and G. M. Lee, 2010, "Result of a Survey of Client Satisfaction with Outpatient Physiotherapy Care", *Physiotherapy Theory and practice*, 26(5): 297-307.

Lee, Y. and J. L. Lin., 2009, "The Effects of Trust in Physician on Self-efficacy, Adherence and Diabetes Outcomes", *Social Science & Medicine*, 68: 1060-1068.

León, C. J., J. E. Araña and J. De León, 2013, "Correcting for Scale Perception Bias in Measuring Corruption: An Application to Chile and Spain", *Social Indicators Research*, 114: 977-995.

Likert, R., 1932, "A Technique for the Measurement of Attitudes", *Archives of Psychology*, 140: 595-639.

Mannetje, A., A. Eng, J. Douwes, L. Ellison-Loschmann, D. McLean, N. Pearce, 2011, "Determinants of non-response in an occupational exposure and health survey in New Zealand", *Australian and New Zealand Journal of Public Health*, 35(3): 256-263.

Mannion, A. F., S. Kämpfen, U. Munzinger and I. Kramers-de Quervain, 2009, "The Role of Patient Expectations in Predicting Outcome after Total Knee Arthroplasty", *Arthritis Research & Therapy*, 11(5): 1-13.

McFadden, D., 1986, "The choice Theory Approach to Market Research", *Marketing Science*, 5(4): 275-297.

McFadden, D., A.C. Bemmor, F.G. Caro, J. Dominitz, B. Jun, A. Lewbel, R.L. Matzkin, F. Molinari, N. Schwarz, R.J. Willis and J.K. Winter, 2005, "Statistical Analysis of Choice Experiments and Surveys", *Marketing Letters*, 16(3/4): 183-186.

Mehta, S., 2011, "Service Quality as Predicator of Patient Satisfaction: A Study of the Health Care Sector", *Journal of Health Management*, 13(2): 211-229.

Michailidis, G. and J. De Leeuw, 1998, "The Gifi System of Descriptive Multivariate Analysis", *Statistical Science,* 13: 307-336.

Noelle-Neumann, E., 1974, "The spiral of silence: a theory of public opinion", *Journal of Communication*, 24(2): 4351.

O'Brien, R. M. and P. Homer, 1987, "Corrections for Coarsely Categorized Measures: LISREL's Polyserial and Polychoric Correlations", *Quality & Quantity*, 21(4): 349-360.

Pai, Y. P. and G. Ravi, 2011, "Factors Affecting In-patient Satisfaction in Hospital - A Case Study", Paper presented at the International Conference on Technology and Business Management, Dubai, March 28-30.

Pijnenborg, A. A. J. M., 2009, "Anxiety and Overall Patient Satisfaction in Cataract Patients: Reducing the Impact of Anxiety by Special Adjustments of Care Delivery", Tilburg University working paper.

Purdy, N. M., 2011, "Effects of Work Environments on Nursing and Patient Outcome", PhD diss., University of Western Ontario.

Qin, H., 2009, "Links among Perceived Service Quality, Patient Satisfaction and Behavioral Intentions in the Urgent Care Industry: Empirical Evidence from College Students", PhD diss., University of North Texas.

Raftopoulos, V., 2010, "Assessment of Users' Expectations, Perceived Quality and Satisfaction with Primary Care in Greece", *International Journal of Caring Sciences,* 3(3): 110-120.

Rahmqvist, M. and A.C. Bara, 2010, "Patient characteristics and quality dimensions related to patient satisfaction", *International Journal for Quality in Health Care,* 22(2): 86-92.

Rasch, G., 1960, *Probabilistic Models for Some Intelligence and Attainment Tests,* Copenhagen: The University of Chicago Press.

Riphahn, R. T. and O. Serfling, 2005, "Item non-response on income and wealth questions", *Empirical Economics*, 30: 521-538.

Roland, G., 2008, *Privatization: Success and Failures*, Columbia University Press.

Rosenthal, R. and M. R. Dimatteo, 2001, "Meta-Analysis: Recent Developments in Quantitative Methods for Literature Reviews", *Annual Review of Psychology*,
52: 59-82.

Ruiz, M.A., A. Pardo, J. M. Martinez de la Casa, V. Polo, J. Esquiro and J. Soto, 2010[a], "Development of a Specific Questionnaire Measuring Patient Satisfaction with Glaucoma Treatment: Glausat", *Ophthalmic Epidemiology*, 17(3): 131-143.

Ruiz, M.A., F. Heras, A. Alomar, L. Conde-Salazar, J. de la Cuadra, E. Serra, F. Regalado, F. Halback and GEIDAC Group, 2010b, "Development and Validation of a Questionnaire on 'Satisfaction with Dermatological Treatment of Hand Eczema' (DermaSat)", *Health and Quality of Life Outcomes*, 8:127-142.

Salini, S. and R. S. Kenett, 2009, "Bayesian Networks of Customer Satisfaction Survey Data", *Journal of Applied Statistics*, 36(11): 1177-1189.

Sarikakis, K., 2011, “Communicating Europe: Steps Towards the Public Sphere?”, In *Communication, Mediation and Culture in the Making of Europe*, edited by Lodge, J. and K. Sarikakis, Rome: Il Mulino.

Schenker, Y., D. B. White, D. A. Asch and J. M. Kahn, 2012, “Health-care System Distrust in the Intensive Care Unit”, *Journal of Critical Care*, 27(1): 3-10.

Shabbir, S., H. R. Kaufmann and M. Shehzad, 2010, “Service Quality, Word of Mouth and Trust: Drivers to Achieve Patient Satisfaction”, *Scientific Research and Essays*, 5(17): 2457-2462.

Shrestha, J., 2010, “Evaluation of Access to Primary Healthcare. A Case Study of Yogyakarta, Indonesia”, Master diss., International Institute for Geo-Information Science and Earth Observation.

Sims, W. D., 2009, “The Relationship of General Adult Attachment and Physician-Patient Attachment with Experience and Expression of Anger among Chronic Pain Patients”, Master diss., Oklahoma State University.

Sixma, H. J., P. M. M. Spreeuwenberg and M. A. A. van der Pasch, 1998, Patient Satisfaction with the General Practitioner: A Two-Level Analysis, *Medical Care*, 36(2): 212-229.

Stevens, L. F., 2010, “Cultural Factors and Communication During Medical Consultation with HIV-Positive Racial/Ethnic Minority Patients”, PhD diss., Virginia Commonwealth University.

Tenenhaus, M., E. V. Vinzi, Y. M. Chatelin and C. Lauro, 2005, “PLS Path Modelling”, *Computational Statistics & Data Analysis*, 48: 159-205.

Tourangeau, R. and T. W. Smith, 1996, “Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context”, *Public Opinion Quarterly*, 60: 275-304.

Turner, R. M., D. J. Spiegelhalter, G. C. S. Smith and S. G. Thompson, 2009, “Bias Modelling in Evidence Synthesis”, *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 172: 21-47.

Wardman, M., 1988, “A comparison of revealed preference and stated preference models of travel behaviour”, *Journal of Transport Economics and Policy*, 22(1): 71-91.

Weng, H., 2009, “A Multisource and Repeated Measure Approach to Assessing Patient-Physician Relationship and Patient Satisfaction”, *Evaluation & the Health Professions*, 32(2): 128-143.

Weng, H., C. Hung, T. Liu, Y. Cheng, C. Yen, C. Chang and C. Huang, 2011, “Associations between Emotional Intelligence and Doctor Burnout, Job Satisfaction and Patient Satisfaction”, *Medical Education*, 45: 835-842.

Wold, H., 1982, “Systems under indirect observation using PLS”, In *A Second Generation of Multivariate Analysis: Methods*, edited by C. Fornell, New York: Praeger.

Wouda, J. C., L. C. Zandbelt, E. M. A. Smets and H. B. M. Van De Wiel, 2011, “Assessment of Physician Competency Inpatient Education: Reliability and Validity of a Model-based Instrument”, *Patient Education and Counseling*, 85(1): 92-98.

Zatti, A., 2012, “New Organizational Models in European Local Public Transport: From Myth to Reality”, *Annals of Public and Cooperative Economics*, 83(4): 533-559.

This yearly series of working papers (WP) aims to publish essentially works in English or in French resulting from the scientific network of CIRIEC and more specifically its working groups. The WP are submitted to a review process and are published under the responsibility of the President of the International Scientific Council, the president of the scientific Commissions or the working groups coordinators and of the editor of the CIRIEC international scientific journal, the *Annals of Public and Cooperative Economics*.

These contributions may be published afterwards in a scientific journal or book.
The contents of the working papers do not involve CIRIEC's responsibility but solely the author(s') one.

The submissions are to be sent to CIRIEC, Université de Liège au Sart Tilman, Bât B33 (bte 6), BE-4000 Liège, Belgique.

Cette collection annuelle de Working Papers (WP) est destinée à accueillir essentiellement des travaux en français ou en anglais issus du réseau scientifique du CIRIEC et en particulier de ses groupes de travail. Les WP font l'objet d'une procédure d'évaluation et sont publiés sous la responsabilité du président du Conseil scientifique international, des présidents des Commissions scientifiques ou des coordinateurs des groupes de travail et de la rédactrice de la revue scientifique internationale du CIRIEC, les *Annales de l'économie publique, sociale et coopérative*.

Ces contributions peuvent faire l'objet d'une publication scientifique ultérieure.
Le contenu des WP n'engage en rien la responsabilité du CIRIEC mais uniquement celle du ou des auteurs.

Les soumissions sont à envoyer à l'adresse du CIRIEC, Université de Liège au Sart Tilman, Bât B33 (bte 6), BE-4000 Liège, Belgique.

## Publications

2014/01 Berliner Wasserbetriebe (BWB) – Water and sewage company in Berlin
Christina SCHAEFER & Stephanie WARM

2014/02 Berliner Wasserbetriebe (BWB)
Christina SCHAEFER & Stephanie WARM

2014/03 Publicization versus Privatization: Recent worldwide evidence
Stefano CLÒ, Chiara F. DEL BO, Matteo FERRARIS, Carlo FIORIO, Massimo FLORIO & Daniela VANDONE

2014/04 Statistical methods for evaluating satisfaction with public services
Giancarlo MANZI & Pier Alda FERRARI

CIRIEC (International Centre of Research and Information on the Public, Social and Cooperative Economy) is a non governmental international scientific organization.

Its **objectives** are to undertake and promote the collection of information, scientific research, and the publication of works on economic sectors and activities oriented towards the service of the general and collective interest: action by the State and the local and regional public authorities in economic fields (economic policy, regulation); public utilities; public and mixed enterprises at the national, regional and municipal levels; the so-called "social economy" (not-for-profit economy, cooperatives, mutuals, and non-profit organizations); etc.

In these fields CIRIEC seeks to offer information and opportunities for mutual enrichment to practitioners and academics and for promoting international action. It develops activities of interest for both managers and researchers.

Le CIRIEC (Centre International de Recherches et d'Information sur l'Economie Publique, Sociale et Coopérative) est une organisation scientifique internationale non gouvernementale.

Ses **objectifs** sont d'assurer et de promouvoir la collecte d'informations, la recherche scientifique et la publication de travaux concernant les secteurs économiques et les activités orientés vers le service de l'intérêt général et collectif : l'action de l'Etat et des pouvoirs publics régionaux et locaux dans les domaines économiques (politique économique, régulation) ; les services publics ; les entreprises publiques et mixtes aux niveaux national, régional et local ; l'économie sociale : coopératives, mutuelles et associations sans but lucratif ; etc.

Le CIRIEC a pour but de mettre à la disposition des praticiens et des scientifiques des informations concernant ces différents domaines, de leur fournir des occasions d'enrichissement mutuel et de promouvoir une action et une réflexion internationales. Il développe des activités qui intéressent tant les gestionnaires que les chercheurs scientifiques.

International Centre of Research and Information on the Public, Social and Cooperative Economy - aisbl
Centre international de Recherches et d'Information sur l'Economie Publique, Sociale et Coopérative - aisbl

Université de Liège au Sart-Tilman
Bât. B33 - bte 6
BE-4000 Liège (Belgium)

Tel. : +32 (0)4 366 27 46
Fax : +32 (0)4 366 29 58
E-mail : ciriec@ulg.ac.be
http://www.ciriec.ulg.ac.be