

NON-RESPONSE - THE CURSE OF THE PRE-ELECTORAL AND ELECTORAL SURVEYS

KRASSIMIR KALINOV,
NEW BULGARIAN UNIVERSITY

Abstract

The problem of non-response or missing data is central for present-day social surveys. Especially important it is for pre-electoral and electoral surveys where the response rate usually is very low. The aims of this survey presentation are: i) to describe the sources of non-response; ii) to discover the reasons of non-response; iii) to summarize the methods for overcoming the problem, and iv) to give real examples from Bulgarian recent elections. Additionally, the methods for calculating the response rate are summarised.

The comparison is made between the response rates (RR) for different types of electoral and pre-electoral surveys, and the impact of missing data on the estimates is investigated.

Special attention is given to exit-polls where the response rate is extraordinary low. A comparative analysis is given how the non-response rates change over the past 10 years. Real data from Bulgarian municipal elections are used for this analysis.

The conclusion is that the problems with low RR in electoral and pre-electoral surveys remain unresolved and investigations are necessary to discover the determinants of non-response.

INTRODUCTION

Incomplete data are a common problem of survey research. Missing data is a widespread phenomenon in survey research especially in electoral and pre-electoral researches. On the basis of empirical articles published in American political science journals, King et al. (2001) estimate that on average about a third of the original sample used in these papers is excluded from the analyses due to item non-response in any one of the underlying variables.

The most common way to handle incomplete data is the listwise deletion of such cases. Deleting in the event of missing data confronts analysts with a trade-off: either they lose numerous cases, or they exclude covariates of interest when their share of item non-response appear too high. Both alternatives are unsatisfactory as they invalidate costly information and, more importantly, limit the statistical inferences that can be drawn from survey data. Yet the listwise deletion of missing data may not only obstruct practical restrictions on the number of analyzable units and variables; it may also generate selection bias in the models estimated. The biasing effect of incomplete data depends on the marginal distribution of non-response. Missing data for statistics like fractured extremity for medicine – it may be cured but the pain remains.

SOURCES AND TYPES OF NONRESPONSE

All sample surveys are subject to various types of errors in consequence of non-response. Table 1 summarizes these errors and their sources.

Table 1. Type of errors

Error	Source
Coverage error	Failure to give any chance of sample selection to some persons in the population
Non-response error	Failure to collect data on all persons in the sample. <ul style="list-style-type: none">- Item non-response- Unit or complete non-response
Sampling error	Caused by observing a sample instead of the whole population. Can be split into a selection error and an estimation error.
Measurement error	Inaccuracies in responses recorded on the survey instruments arise from: <ul style="list-style-type: none">- influence of interviewers on the respondents' answers to survey questions;- error due to respondents, from inability to answer questions, lack of requisite effort to obtain the correct answer, or other psychological factors;- error due to the weaknesses in the wording of survey questionnaires;- error due to effects of the mode of data collection, the use of face to face or telephone communication.

REASONS FOR NON-RESPONSE

Unit Non-Response

- Failure of the interviewer to locate/identify the respondent;
- Failure to make contact with the respondent;
- Refusal of the respondent to participate;
- Inability of the respondent to participate (e.g. illness, absence, etc);
- Inability of the interviewer and respondent to communicate (e.g. language barriers);
- Accidental loss of the data/ questionnaire.

Unit non-response is often divided into three components: non-contact, inability to respond, lack of co-operation (refusal).

Item non-response

A sample unit participates but data for some survey items are not available for analysis.

Reasons could include:

- Refusal to provide an answer
- Inability to provide an answer
- Other failure to answer (e.g. by accident)
- Provided answer being of inadequate quality (e.g. incomplete, implausible, failing an edit/consistency check, etc.)

Item non-response can be caused by:

- the action of responder (e.g. refusal to answer);
- the action of an interviewer (e.g. failure to ask a question that should have been asked, or failure to record the answer adequately);
- the survey design (e.g. poor routing instruction).

As may be seen from Table 1, non-response error can be divided into two components:

- Error due to unit non-response
- Error due to item non-response

Each of these can in turn be decomposed into subsources as shown in Table 2.

Table 2. Reasons for Non-response

Type of non-response	Reason
Unit non-response	• Non-contact
	• Refusal to respond
	• Inability to respond
Item non-response	• Instrument error
	• Interviewer error
	• Refusal to respond
	• Inability to respond

The effect of non-response on a survey estimate can be defined (Groves, 1989) as:

$$\theta_k = \theta_n + \frac{n_k}{n} (\theta_k - \theta_{n_k}),$$

where

θ_k is statistic for the k responding units,

θ_n is statistic for all n sample units,

θ_{n_k} is statistic for the n_k non-responding units.

The term $\frac{n_k}{n} (\theta_k - \theta_{n_k})$ is the estimate of total non-response error. Thus, the error introduced to the survey estimate is a function of the percentage of the sample not responding to the survey and the differences on the statistic between respondents and non-respondents [i.e., non-response error = (non-response rate) x (difference between respondent and non-respondent values)]. In other words, the non-response error has two elements:

- i) the non-response rate, and
- ii) the difference between responding and non-responding units in terms of θ .

Both of these elements are important. Note that the second element will result from both non-response bias and variance. In practice, studies suggest that bias is usually the main component.

Thus bias is associated with both low response rates and strong differences in the estimates between respondents and non-respondents.

Any estimate from a study can be subject to bias due to non-response across one or more stages. Evaluation of the bias is not always possible as the true value of the population parameter is unknown. Wherever a true population value is known, the difference between the value computed from the survey data and the true population value can be considered an estimate of the bias related to the survey estimate.

A non-response bias analysis is the process that results in the quantification of estimated non-response bias, and identification of potential sources of non-response bias on estimates. Non-response bias analyses allow for the evaluation of survey statistics that are estimated using both base (only reflecting selection probabilities) and non-response adjusted weights.

There are different ways in which non-response bias analyses are useful. Non-response bias analyses serve as an indicator of the quality of the data collected, and help identify potentially biased estimates. Such analyses can help reassure data users, as well as the agency collecting and releasing data, of the quality of the data available.

Simultaneously, it warns users of data vulnerable to bias. Such analyses can also be used to evaluate the variables used in non-response weighting adjustments.

CALCULATION OF RESPONSE RATE (RR)

Response rates measure the proportion of eligible sample units that successfully provided data, i.e. the reduction due to non-response of cases available for analysis. Response rates are more likely to be reported than almost any other survey process quality indicator.

Response rates are often mistakenly used as a measure of quality of *survey statistics*. In fact, response rate is just one component of non-response error.

Unfortunately, there is no unique standard definition of a response rate. Here a definition is used in its simplest and the most popular form that is similar to the one introduced by Lynn et al. (2002): “The response rate is defined as the proportion of eligible elements in the sample for which a questionnaire has been completed”.

Let the initial sample size be n . It can be written as a sum of the following components:

$$n = \sum_{i=1}^5 n_i ,$$

where n_1 is the number of non-contacts, n_2 is of noneligible elements among the contacts (i.e., cases of overcoverage), n_3 is the number of refusers, n_4 is the number of not-able elements, and n_5 the number of respondents.

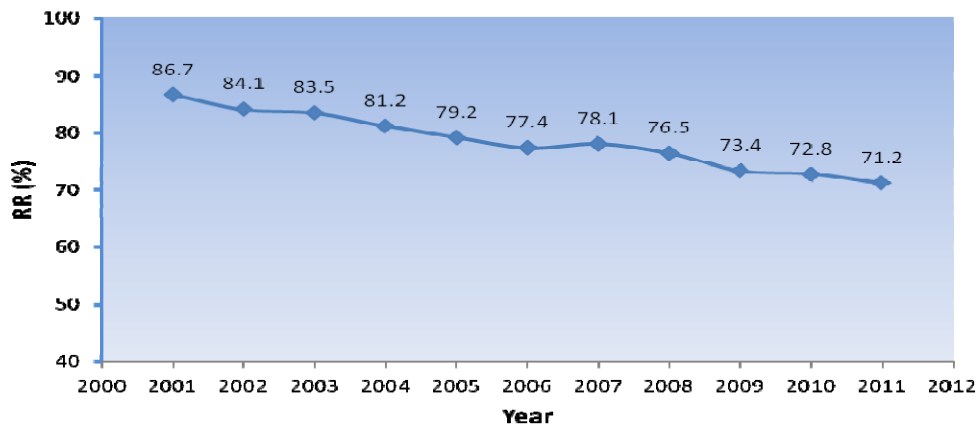
The response rate is defined as the number of respondents over the number of eligible elements in the sample:

$$\text{Response Rate} = \frac{n_5}{n_{\text{eligible}}} .$$

There is a problem in computing the number of eligible elements. In the case when the method of the nearest birthday (Kish, 1965) is applied then the formula above may be used directly. It is the Bulgarian case. Other computational formulas one may find in Bethlehem (2009).

Fig.1 shows the change over time of RR for Bulgarian population since 2001.

Trend in Average Response Rate*



* Data from National Center for Public Opinion

Fig. 1 Change in average response rate over time for pre-electoral surveys on the example of Bulgaria

Response rates have declined over time in many countries, including Bulgaria (Bethlehem (2009), Table 9.4). Fig. 1 shows the trends in declination based on data from National Center of Public Opinion. Within 11 years RR has reduced more than 15 percents or declination of approximately 18% is observed. Thus the problem with survey non-response becomes serious from practical as well from scientific point of view.

The main problem of nonresponse is that estimates of population characteristics may be biased. This situation occurs if, due to nonresponse, some groups in the population are over- or underrepresented in the sample, and these groups behave differently with respect to the characteristics to be investigated. Then, nonresponse is said to be selective. Moreover, if non-response is related to unobserved determinants of the variable of interest, and if this is ignored, then in general the estimation results are inconsistent. Empirical studies based on pre-electoral survey data do not pay much attention to such non-response, for the reason that it is felt that there is nothing one can do about it.

As Horowitz and Manski (1998) state, "... the only way to identify population statistics is to make assumptions that determine the distribution of the missing data. A fundamental problem of empirical analysis is that such assumptions are untestable."

METHODS FOR OVERCOMING THE PROBLEM

1. Unit non-response. To avoid the realized sample of being too small, the initial sample size should be taken larger. For example, in omnibus pre-electoral surveys usual required sample size is of at least 1000 face-to-face filled questionnaires. As it may be seen from Fig. 1 the response rate is about 70%. Then the initial sample size should be enlarged approximately by 30% so it becomes 1300.

Studying Non-Response

- Special studies of non-respondents
- Using information on the sampling frame
- Asking others about non-respondents or having interviewer provide information

- about them
- Comparison of respondent characteristics by call number
- Comparison of respondent characteristics to census or other external information
- Studying persons who drop out of a panel survey after an initial interview

More about this topic can be found in Groves, Cialdini and Couper (1992).

Formal methods for modeling of unit non-response

A relatively new technique to adjust for an unit non-response bias is the propensity score method described by Rosenbaum and Rubin (1983).

Let $\{Y_1, Y_2, \dots, Y_N\}$ be the population of interest, and let $\{y_1, y_2, \dots, y_n\}$ be the probability sample from this population. Let δ_i be a selection index, equal to 1 if the i -th element is selected, and 0 otherwise.

For every sample element i we have a vector of covariates (auxiliary variables), denoted by $X'_i = (X_{1i}, X_{2i}, \dots, X_{ki})$. This information usually includes such respondent's characteristics as sex, age, social status etc. In addition, we have a survey item y that is only observed for respondents. We assume that every element in the population has a nonzero, unknown response probability, denoted by ρ_i . This corresponds to the so called Random Response Model.

The response probabilities can be estimated based on the sample data. By using an appropriate model based on covariates, one can compute sample-based estimates of the response probabilities, i.e.

$$p_i = \hat{\rho}(\vec{X}_i) = \mathbf{Prob}(d_i = 1, \vec{X}_i), \quad i=1, 2, \dots, n,$$

where d_i is the observed binary response indicator which is either 1 for response or 0 for non-response.

$\rho(\vec{X}_i)$ is so called *response propensity*. It is obvious that the response propensity is the estimated probability to receive a response conditional on the sample and the individual covariates \vec{X}_i . The response propensity can be used to adjust for nonresponse in different ways. Two approaches are used most frequently.

The first approach uses the response propensities directly. Two variations are known: *response propensity weighting* and *response propensity stratification*. In this direct approach, the available auxiliary information is used to model the relationship between the response and the covariates.

The second approach involves the relationship between the survey item and the covariates as well. This approach may be considered as an extension of the GREG-estimator (Generalized Regression).

i) Reponse propensity weighting estimator

The response propensity can be inserted in the modified Horvitz-Thompson estimator (Sarndal, 1981).

Let the target of interest be the population mean of survey item. The population consists of N elements, from which a sample of size n is drawn. The Horvitz-Thompson estimator is an unbiased estimator for the population mean (Horvitz and Thompson (1952)) which may be written in this case as:

$$\bar{y}_{HT} = \frac{1}{N} \sum_{i=1}^n \frac{d_i y_i}{p_i}.$$

ii) Response propensity stratification estimator

Suppose the sample is stratified in k strata according given covariate. Let s_1, s_2, \dots, s_k denote the strata, and let n_1, n_2, \dots, n_k be the subsample sizes. All elements from given stratum have the same weight. This weight may be written as:

$$\omega_i = \frac{n_i}{r_i},$$

where r_i is the number of respondents in stratum i . Then the estimator can be expressed as:

$$\bar{y}_{st} = \frac{1}{N} \left(\sum_{i=1}^{n_1} \omega_1 y_i^{s_1} + \sum_{i=1}^{n_2} \omega_2 y_i^{s_2} + \dots + \sum_{i=1}^{n_k} \omega_k y_i^{s_k} \right),$$

where $y_i^{s_k}$ is the i -th observation from k stratum.

iii) The response propensity GREG-estimator

Bethlehem (1988) describes this approach in details. The conclusions Bethlehem has made are that good stratification, i.e. strata are homogeneous according the target variable, perform well in reducing bias. Stratifications that are homogeneous with respect to the response probabilities work well too.

2. Item non-response. In statistical literature there are a lot of proposals how to cope with item non-response, the most of them based on the following classification (Little and Rubin (2002)):

Table 3. Item Non-Response Classification

Item Non-Response Pattern	Explanation
Missing Completely At Random (MCAR)	The probability of missingness is independent of the data
Missing at random (MAR)	The probability of missingness depends only on the observed data
Missing Not At Random (MNAR)	The probability of missingness may also depend on the unobservable part of the data

When incomplete data are MCAR, analyses will not be biased, because there are no systematic differences between respondents who completed the question and respondents who have a missing value for that question.

Extra information is needed to test the MAR hypothesis and help to determine the causes of item non-response. This information may be available in the data set, but often additional information (information from other sources than the actual sample) is needed, such as theory, logic, or auxiliary data from registers, sampling frames, reinterviews, or other studies.

MNAR is clearly the most problematic. Since missingness may depend on data that are unobserved nonidentifiability problem appears. Because of these obstacles, there are no reliable methods that try to model the missingness mechanism for MNAR pattern since the correctness of the model cannot be verified using the observed data.

In fact, there is no formal way to distinguish whether the missing data were MCAR or MNAR from the observed data. That is, there is an inherent nonidentifiability problem here. More on this topic can be found in Lavrakas (2008).

Formal methods for modeling of item non-response

- i) **Complete-case analysis.** A common default approach is complete-case analysis, also known as *listwise* deletion, where incomplete cases are discarded and standard analysis methods applied to the complete cases (Little and Rubin (2002)). When the missing data are MCAR, the complete cases are a random subsample of the original sample, and analysis results is valid (but often inefficient) inferences. If the data are not MCAR then the complete cases are a biased sample, and analysis is often (though not always) biased. The bias depends on the degree of departure from MCAR, the amount of missing data, and the specifics of the analysis. The potential for bias is why sample surveys with high rates of unit nonresponse (say 30% or more) are often considered unreliable for making inferences to the whole population. A modification of complete-case analysis, commonly used to handle item non-response in pre-electoral surveys, is to weight respondents by the inverse of an estimate of the probability of response.
A simple approach to estimation is to form adjustment cells (or subclasses) on the basis of covariates and weight respondents in an adjustment cell by the inverse of the response rate in that cell. This method eliminates bias if respondents within each adjustment cell can be regarded as a random subsample of the original sample in that cell (i.e., the data are MAR given indicators for the adjustment cells). A useful extension with extensive background information is **response propensity** stratification, where adjustment cells are based on similar values of estimates of probability of response, computed by a logistic regression of the indicator for missingness on the covariates. Although weighting methods can be useful for reducing nonresponse bias, they are potentially inefficient.
- ii) **Single imputation.** Methods that impute or fill in the missing values have the advantage that, observed values in the incomplete cases are retained. A simple version imputes missing values by their unconditional sample means or medians based on the observed data, but this method often leads to biased inferences and cannot be recommended in any generality (see Little and Rubin (2002)).
Variants of this approach are known, for instance, *conditional mean imputation*, *stochastic regression imputation*, *hot-deck imputation*, *predictive mean matching* and so on.
- iii) **Multiple imputation.** A serious drawback with imputation is that a single imputed value cannot represent the uncertainty about, which value to impute, so analyses that treat imputed values just like observed values generally underestimate uncertainty, even if nonresponse is modeled correctly. Multiple imputation (Rubin, (1987)) fixes this problem by creating a set of Q (say $Q = 5$ or 10) draws for each missing value from a predictive distribution. Then Q datasets are available, each containing different sets of draws of the missing values.
- iv) **Maximum likelihood. Maximum likelihood (ML)** avoids imputation by formulating a statistical model and basing inference on the likelihood function of the incomplete data (Little and Rubin (2002)).

v) **Bayesian Simulation Methods.** (Schafer (1997)).

How well these methods work on electoral and pre-electoral empirical data, and how useful they are? At the moment we have no answer. The practical utility of these formal methods for electoral surveys is under investigation. An empirical comparison of these three approaches may be found in Cobben (2009).

CONCLUSIONS

Every potential solution to the missing data problem rests on assumptions about the data. The use of different methods therefore hinges upon the correctness of assumptions, something that has to be judged for each variable in the dataset separately. Deleting all missing data without considering the consequences can produce biased results. But the idea of cleaning datasets by universally applying multiple imputation techniques is also problematic.

If one encounters data missing at random (MAR) and is confident that measurement problems are at the root of item non-response, i.e. that a true score for each missing score exists, obtainable multiple imputation techniques are the proper choice. They provide estimates of the unobserved scores. Such methods apply in all cases in which respondents are asked to report facts like demographics or past behavior. Irrespective of whether respondents answer questions on their age or education, they nonetheless have a certain age and education. Missing values on this question most likely reflect measurement problems.

In many situations one cannot know with certainty what caused certain scores to be missing. Research on survey response may indicate how likely measurement problems are for different variables. If one does not have such information to reject or support either the application of existing imputation techniques or the application of a multiple complete random imputation, the latter may be regarded as the more conservative method.

The best way to avoid bias is to improve response rates by using methods such as intensive refusal conversion techniques, incentives, multiple modes of data collection, flexible scheduling, and interviewer training. However, despite best efforts, non-response does occur. In such cases, surveys adjust probability-based weights to compensate for non-response. However, despite adjusting weights for non-response, bias can still persist in estimates. Possible paths to solution: more comprehensive survey theory relating response mechanism to nonresponse bias; more comprehensive statistical theory to adjustment to deal with different statistics; more multivariate approach to both of these theories.

EXIT-POLLS AND NON-RESPONSE



Frank and Ernest @Thaves/Dist. by Newspaper Enterprise Association, Inc.

Non-response is a big problem in exit polls, with less than half the sampled voters responding. The results of the survey are thus very sensitive to how the non-response is handled. The frequency of non-responders is an important indicator about quality of the gathered information. The measure of the non-response level usually is **Rejection rate** (RR) computed as a ratio of non-responders and total number of contacts. Formal:

$$RR = \frac{n_R}{n_R + n_A} \cdot 100\% ,$$

where n_R is the absolute number of non-responders and n_A is the number of responders .

Table 4 summarizes the rejection rates for eight Bulgarian municipalities computed on the basis of data from recent local elections. Parallel to RR, a root mean square error (RMSE) was calculated as a measure for discrepancy from the official vote. More about how to measure the exit-poll errors may be found in Scheuren and Alvey ((2008), ch.7).

$$RMSE = \sqrt{\frac{1}{k} \sum_{i=1}^k (f_i - o_i)^2} ,$$

where k is the number of the participants (parties and/or independent nominees); f_i is exit-poll result (%) for i -th party or nominee, and o_i is the official vote.

Almost all the values of RR are less than 40% which is considered as acceptable (cf. Bautista et al. (2008), Table 1). The only exception is municipality of Pestera where the rejection rate is very high - more than 55% of the contacted voters refused to fill the questionnaire. This is an indication of poor quality of the data and of unreliable prognosis. This fact is well observed by inspection of the RMSE which is more than four times greater than any other RMSE.

The relation between RR and RMSE is well seen in Fig.2. The Pearson's correlation coefficient even for this very small sample is statistically significant and very high, i.e. the relation may be thought as almost linear¹.

¹ Obviously, more data are necessary to *proof* the hypothesis about linear relationship between RR and RMSE.

Table 4. Recent local Bulgarian election results*

	Rejection Rate	Root MSE	RMSE/RR
Municipality			
Blagoevgrad	38.51	1.298	0.0337
Dobrich	35.71	0.886	0.0248
Kjustendil	34.85	0.657	0.0189
Sliven	37.68	1.219	0.0324
Pestera	55.46	5.757	0.1038
Veliko Tarnovo	35.1	0.855	0.0244
Pleven	27.96	0.752	0.0269
Varna	38.63	1.212	0.0314
Blagoevgrad	38.51	1.298	0.0337

* Data from National Center for Public Opinion

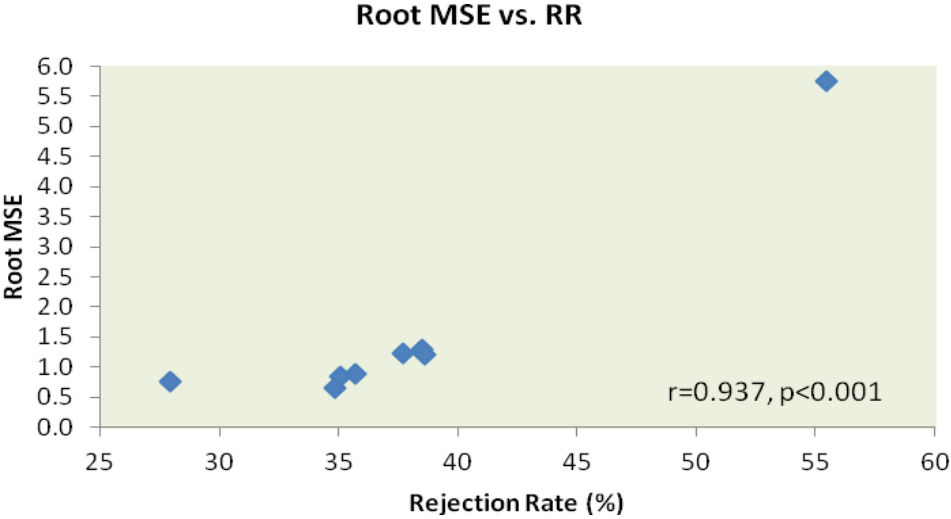


Fig.2 Scatter plot of Root MSE vs. Rejection rate

CONCLUSIONS

Exit polls were developed in the 1970s in the U.S. and quickly disseminated and adopted by numerous countries over the world in the following three decades. Due to the scientific and statistical principles behind them, exit polls became a helpful way to gather electoral results before the last vote was cast. Exit polling has served different purposes in well-established and emerging democracies, ranging from mass media information, scholarly analyses on voting behavior, political tool to assess tampering with electoral results, argument to claim electoral victories by political parties, up to validation of official results when elections are believed to be fraudulent.

Thus far, there is little documentation of exit polls in terms of methodological problems from an international perspective; however, the scarce evidence suggests that nonsampling errors may have an important impact on exit polling estimates. Exit polls are often judged only by looking at the typical sampling error and how far the exit polling estimate is from the actual results. However, there is still a need for developing better ways to understand and measure other potential sources of error in exit polls and include them in the reporting.

REFERENCES

1. Bautista, R., Callegaro, M., Paniotto, V., Kharchenko, N. and Scheuren, F. (2008). Exit polling methodologies across nations: Some constraints and measurement problems. 3MC Conference, Berlin June 2008.
2. Bethlehem, J. (1988). Reduction of Non-response Bias through Regression Estimation, *Journal of Official Statistics*, 4(3), 251-260.
3. Bethlehem, J. (2009). *Applied Survey Methods: A Statistical Perspective*. John Wiley&Sons, New York.
4. Cobben, F. (2009). *New Developments in Non-response Adjustment Methods*, http://epp.eurostat.ec.europa.eu/portal/page/portal/research_methodology/documents/S9P2_NEW_DEVELOPMENTS_IN_NONRESPONSE_COBBEN.pdf (last visit 27 Nov 2011).
5. D'Agostino Jr., R.B. (2005) Propensity Scores. In *Encyclopedia of Statistics in Behavioral Science* (eds. Everitt, B.S. and Howell, D. C.), , v. 3, pp. 1617-1619, John Wiley & Sons, Chichester.
6. Groves, R. (2004). *Survey Errors and Survey Costs*. John Wiley&Sons, New York.
7. Groves, R., Cialdini R. and Couper, M. (1992). Understanding the Decision to Participate in a Survey, *Public Opinion Quarterly* 56: 475-495.
8. Horowitz, J.L. and Manski, C.F. (1998). Censoring of outcomes and regressors due to survey nonresponse: identification and estimation using weights and imputations, *Journal of Econometrics*, 84, 37-58.
9. Horvitz, D., and D. Thompson (1952). A Generalization of Sampling Without Replacement from a Finite Universe," *Journal of the American Statistical Association*, 47, 663-685.
10. King, G., Honacker, J., Joseph, A. & Scheve, K. (2001). Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation. *American Political Science Review* 95: 49-69.
11. Kish, L. (1965). *SURVEY SAMPLING*. John Wiley&Sons, New York.
12. Lavrakas, P., ed., (2008). *Encyclopedia of Survey Research Methods*, Volumes 1and 2, SAGE Publications, L.A.
13. Little, R. and Rubin, D.B. (2002). *Statistical Analysis with Missing Data*, 2nd Edition, John Wiley, New York.
14. Lynn, P., Beerten, R., Laiho, J.&Martin, J. (2002). Towards standardisation of survey outcome categories and response rate calculations. *Research in Official Statistics* 1, 63-86.
15. Rosenbaum, P., and D. Rubin (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika*, 70(1), 41-55.
16. Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*, John Wiley&Sons, New York.
17. Sarndal, C. (1981). Frameworks for Inference in Survey Sampling with Application to Small Area Estimation and Adjustment for Non-Response, *Bulletin of the International Statistical Institute*, 49, 494- 513.
18. Schafer, J.L. (1997). *Analysis of Incomplete Multivariate Data*, Chapman & Hall, London.
19. Scheuren, F.J. and Alvey, W. (2008). *Elections and Exit Polling*. John Wiley&Sons, New Jersey.