SYRTO

Systemic Risk Tomography
Signals, Measurements, Transmission channels and Policy Interventions

A proposal for the treatment of "don't know" responses

Marica Manisera, Paola Zuccolotto

SYRTO WORKING PAPER SERIES
Working paper n. 13 | 2013

This project is funded by the European Union under the 7th Framework Programme (FP7-SSH/2007-2013) Grant Agreement n° 320270

A proposal for the treatment of "don't know" responses

Marica Manisera and Paola Zuccolotto*

2013, December

Summary. In this paper we propose a probabilistic framework for the treatment of "don't know" responses in surveys aimed at investigating human perceptions through expressed ratings. The rationale behind the proposal is that "don't know" is a valid response to all extent, as it informs about a specific state of mind of the respondent, and then it is not correct to treat it as a missing value, as usual. The actual insightfulness of the proposed model depends on the choices about the involved probability laws. The required assumptions regard the probability distribution of, firstly, the expressed ratings and, secondly, the state of mind of "don't know" respondents toward the ratings. About the former, in this paper we choose to work in the framework of CUB models, while for the latter we propose to use the Uniform distribution and we justify this choice with a number of motivations, based on formal and empirical rationales. We show that these two choices provide a solution both tractable and easy to interpret, where "don't know" responses can be taken into account with a simple adjustment of one parameter of the model.
Keywords: Rating data, "don't know" responses; CUB models

1 Introduction

Questionnaires and other rating scales are commonly used in a wide range of disciplines. In general, they are aimed at collecting respondents' judgments about a given issue, in order to measure a latent trait Y. For example, when the latent trait Y under evaluation is the customer satisfaction with a product, it is common to ask for customers' opinions about a set of features (items) of the product at hand, by requiring them to express a rating on a given response scale. However, the quality of the results from whatever analysis can be affected by nonresponse: for many reasons, subjects may decide not to respond to some items or they may unintentionally skip some questions. Nonresponse leads to the problem of missing data, which is related to less efficient estimates, difficulties in using standard statistical techniques conceived to deal with complete data, and bias due to systematic differences between respondents and nonrespondents (Schafer, 1997). In the literature, several procedures have been proposed to face this issue (Carpita and Manisera, 2011, e.g.); a very frequent option (the default in many statistical packages) is to ignore the missing data and/or omit subjects with missing data from the study (listwise and pairwise deletions); alternatively, weighting and model-based procedures can be implemented (e.g., the well-known EM algorithm and the data augmentation technique); otherwise, imputation procedures can be used to fill in missing values with one (single imputation) or many (multiple imputation) "plausible" values to create a completed data set that can be analysed with techniques requiring complete data (Rubin, 1987). Finally, missing data may be imputed with interval-valued or histogram-valued data and then treated with symbolic data analysis techniques (Zuccolotto, 2011, 2012). In this paper we deal with a very specific kind of nonresponse, occurring when the respondent, although having experience about the latent trait Y under evaluation, feels unable to formulate the requested rating, due to uncertainty, reluctance, reticence, or because he/she has never formulated a point of view, etc. In this situation, the most appropriate response should be "don't know" (*dk*).

In the literature, *dk* responses have been studied for a long time (for a review, see Krosnick 2002). Whether to include or not the *dk* option in a response scale is an open issue (after the seminal work of Converse 1964, see, among others, Beatty and Herrmann 1995; Lietz 2010; Poe et al. 1988 and the references therein). The main motivation supporting the inclusion of the *dk* option is that to force the respondent to answer all the questions, even when he/she is really unable to face the decision process leading to express the requested rating, could seriously affect the quality of data (Schuman and Presser, 1996), as respondents would provide meaningless answers (Bishop et al., 1986). On the other hand, some

*University of Brescia, c.da S. Chiara 50 - 25122 Brescia, Italy.
email: marica.manisera@eco.unibs.it, paola.zuccolotto@eco.unibs.it

researchers state that *dk* responses may be given prematurely, as several respondents may be able to provide adequate answers with some additional effort (Cannell et al., 1979).

Another open issue is the statistical treatment of such responses (Rubin et al., 1995; Schafer and Graham, 2002, e.g.). Although a *dk* response contains a specific information about the state of mind of the respondent, it is often treated as a missing value. Imputation procedures replace *dk* responses with estimates of the corresponding rating, usually by drawing information from the observed data, implicitly assuming the same probability distribution of the random variable R generating the observed ratings. As a matter of fact, most of the existing imputation procedures are not appropriate to handle *dk* responses (Feick, 2005, e.g.) and some alternatives have been proposed in the literature (Feick, 2005; Kroh, 2006). Following another approach, *dk* could be treated as one of the possible response categories. However, the addition of the *dk* category to the ordered ratings imposes a nominal scaling level to the random variable generating responses and this prevents the use of the statistical methods usually employed to model ratings, as they are mostly conceived for ordinal data (with very few exceptions, see Meulman et al. 2004, e.g.).

In this paper we propose to consider *dk* responses as valid responses to all extent and define a probabilistic framework able to exploit the information contained in the *dk* responses while preserving the ordinal nature of the random variable R generating the ratings. This is done by assuming two different distributions for R, conditionally on whether the respondent is able or not to formulate a rating for the analysed latent trait Y. One important feature of our proposal is that the aim is not to replace each single *dk* with a substantive response, i.e. we do not make conjectures about the possible response of each given subject if he/she were forced to answer. This is a main difference with other methods, which replace each *dk* answer with a (sometimes questionable) plausible value and then carry out all the following analyses as if that value were the real one. On the contrary, our method treats the attitudes of the respondents who "don't know" only at an aggregate level and this leads to an adjustment of the parameters' estimates, but no further reasoning (e.g. evaluation of association between variables, clusterization, etc.) is made by imputing the single *dk* responses. This may be a significant advantage when a well-founded hypothesis for imputation is not available. From a probabilistic point of view, we will show that our idea can be effectively formulated by modelling the expressed ratings within the framework of the CUB models (D'Elia and Piccolo, 2005; Piccolo, 2003).

2 "Don't know" responses treatment with CUB models

In this Section we show how the probabilistic framework introduced in Section 1 can be studied making use of the CUB models. The marginal distribution of the ratings is given by

$$\begin{aligned} Pr(R=r;\theta) &= f\left[\pi_0 b_r(\xi_0) + (1-\pi_0)Pr(U=r)\right] + \\ &+ (1-f)Pr(U=r) \end{aligned}$$

which, after simple algebra, can be written as

$$Pr(R=r;\theta,f) = \pi_{adj} b_r(\xi_0) + (1-\pi_{adj})Pr(U=r), \tag{1}$$

where

$$\pi_{adj} = f\pi_0. \tag{2}$$

In other words, we have $R \sim CUB(\theta)$, with $\theta = (\xi_0, \pi_{adj})'$. The estimation procedure works as follows:

1. f is estimated by the relative frequency of expressed ratings,
2. ML estimates of $\theta_0 = (\xi_0, \pi_0)'$ can be computed via EM algorithm using the observed sample of $n_{lw} = n - n_{dk}$ observations obtained with listwise deletion of *dk* responses, i.e. by estimating a CUB model with $n = n_{lw}$, $\theta = \theta_0$ and $p_r(\theta_0) = Pr(R = r_i; \theta_0 | A = 0)$.

Finally, the estimate of π_{adj} is given by

$$\hat{\pi}_{adj} = \hat{f}\hat{\pi}_0 \tag{3}$$

and θ is estimated by $\hat{\theta} = (\hat{\xi}_0, \hat{\pi}_{adj})$. Equation (3) shows that considering *dk* responses as missing values and simply estimating the CUB model with listwise deletion of the *dk* responses can lead to underestimate the uncertainty component of the analysed population. Instead, a very simple adjustment is possible, obtained by simply multiplying the parameter estimated with listwise deletion by the relative frequency of expressed ratings, $(n - n_{dk})/n$.

3 Real data analysis

This Section presents a case study dealing with data coming from the wave 78.1 of the Eurobarometer survey, called Standard Eurobarometer 78, carried out by the consortium TNS Opinion & Social on request of the European Commission (European Commission, 2013). The Standard Eurobarometer 78 is a sample survey covering the national population of citizens of the 27 European Union Member States. In this work, we did not consider the six candidate Countries (Croatia, Turkey, the Former Yugoslav Republic of Macedonia, Iceland, Montenegro and Serbia) and the Turkish Cypriot Community, where the survey was also conducted. The number of interviewees ranges from 500 (in Malta, MT) to 1561 (in Germany, DE), with a mean of 986 over the 27 considered Countries. More details, included the abbreviations used for the Countries' names, can be found on the technical notes and the annexes to the report on Standard Eurobarometer 78 (European Commission, 2013). Our analysis focuses on the responses given to the questions QA20.7, QA20.8, QA20.11, QA20.12 with possible responses given by "totally disagree", "tend to disagree", "tend to agree", "totally agree" and *dk*. We can reasonably expect that each of the interviewees has at least heard about these very general issues on globalisation and European Union, so the *dk* responses come from their inability to express the requested rating. The percentage of *dk* responses is rather high on these questions, with an average (over Countries) equal to 14.67, 11.18, 9.52, 21.22 for the four questions, respectively. The proposal for modelling *dk* responses is here considered within the framework of CUB models, as suggested in Section 2. A CUB model is fitted to the frequency distribution of each variable under study for each of the Countries, firstly deleting all the *dk* responses and then including them and adjusting the estimates of the uncertainty parameter.

Focusing on the first statement (QA20.7), Figure 1 (left) shows the estimates of the feeling and uncertainty parameters when listwise deletion of *dk* responses is used. Most Countries show a null uncertainty, although with a varying feeling. Only few Countries are associated with a degree of uncertainty higher than 0.2: Cyprus (CY) with a low level of feeling ($1-\hat{\xi}_0 = 0.16$, $1-\hat{\pi}_0 = 0.44$) and others with higher levels of feeling, namely Ireland (IE; $1-\hat{\xi}_0 = 0.67$, $1-\hat{\pi}_0 = 0.25$), Spain (ES; $1-\hat{\xi}_0 = 0.54$, $1-\hat{\pi}_0 = 0.35$), Slovenia (SI; $1-\hat{\xi}_0 = 0.50$, $1-\hat{\pi}_0 = 0.29$). The middle panel in Figure 1 interestingly shows the more realistic positioning of the 27 Countries when the *dk* responses are modelled as proposed in this paper and the uncertainty is adjusted for the presence of *dk*s. The adjustment, that is the increase of uncertainty due to the presence of *dk* responses, is shown in the arrows in the right panel of Figure 1. As expected, the Countries with the highest percentage of *dk* responses also have the highest increase in the uncertainty estimate; the highest differences between $\hat{\pi}_{adj}$ and $\hat{\pi}_0$ are observed for Romania (RO, $dk = 34\%$), Malta (MT, $dk = 27\%$), Bulgaria (BG, $dk = 26\%$), Lithuania (LT, $dk = 26\%$), Spain (ES, $dk = 25\%$).

Note that without the adjustment in the uncertainty estimates, the CUB models represented by points in Figure 1 (left) cannot be fairly compared, because the different percentages of *dk* responses are not taken into account in the uncertainty estimation. For example, Belgium (BE) and Portugal (PT), showing the same behaviour before the adjustment, with a feeling of $1-\hat{\xi}_0 = 0.47$ and an almost null uncertainty in both cases (Figure 1, left), have instead a quite different position after taking the *dk* responses into account ($dk = 4\%$ and $dk = 21\%$, respectively), with adjusted uncertainty parameters given by $1-\hat{\pi}_{adj} = 0.04$ for Belgium and $1-\hat{\pi}_{adj} = 0.21$ for Portugal (Figure 1, middle and right).

The same analysis is carried out for questions QA20.8, QA20.11, QA20.1. The arrows representing the adjustment in the estimates is shown (Figure 2). The right panel (QA20.12) shows the most relevant shifts of the original positions, since the percentages of *dk* responses to this question are quite high, ranging from 6 (in Denmark, DK) to 46 (in Malta, MT).

In this case, we show that the adjustment is also important for the comparison of different statements evaluated in the same Country. For example, before adjustment, Spain (ES) exhibits a very different uncertainty in questions QA20.11 and QA20.12 ($1-\hat{\pi}_0 = 0.57$ and $1-\hat{\pi}_0 = 0.38$, respectively) but the different percentages of *dk* responses ($dk = 13\%$ and $dk = 31\%$) lead to very similar adjusted uncertainty parameters, given by $1-\hat{\pi}_{adj} = 0.62$ and $1-\hat{\pi}_{adj} = 0.57$ (Figure 2, middle and right).

4 Conclusions

The questionnaires of surveys aimed at investigating human perceptions often contain questions where the "don't know" option (*dk*) is available to respondents who do not feel able to formulate a response, in spite of having experienced the trait under study. However, as a matter of fact, the statistical models subsequently used to elaborate the collected data are usually unable to properly take account of this

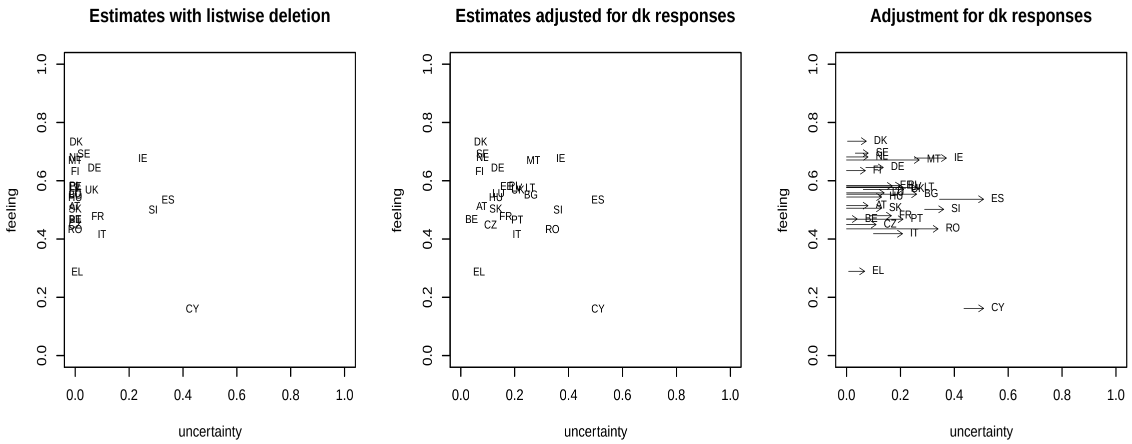

Figure 1: Question QA20.7. *Left*: Uncertainty $(1-\hat{\pi}_0)$ *(x-axis)* and feeling $(1-\hat{\xi}_0)$ *(y-axis)* estimates for each Country with listwise deletion of *dk* responses; *Middle*: Uncertainty $(1-\hat{\pi}_{adj})$ *(x-axis)* and feeling $(1-\hat{\xi}_0)$ *(y-axis)* estimates for each Country adjusted for *dk* responses; *Right*: Adjustment of the uncertainty estimates for the *dk* responses; for each Country, an arrow is drawn from $(1-\hat{\pi}_0, 1-\hat{\xi}_0)$ to $(1-\hat{\pi}_{adj}, 1-\hat{\xi}_0)$.

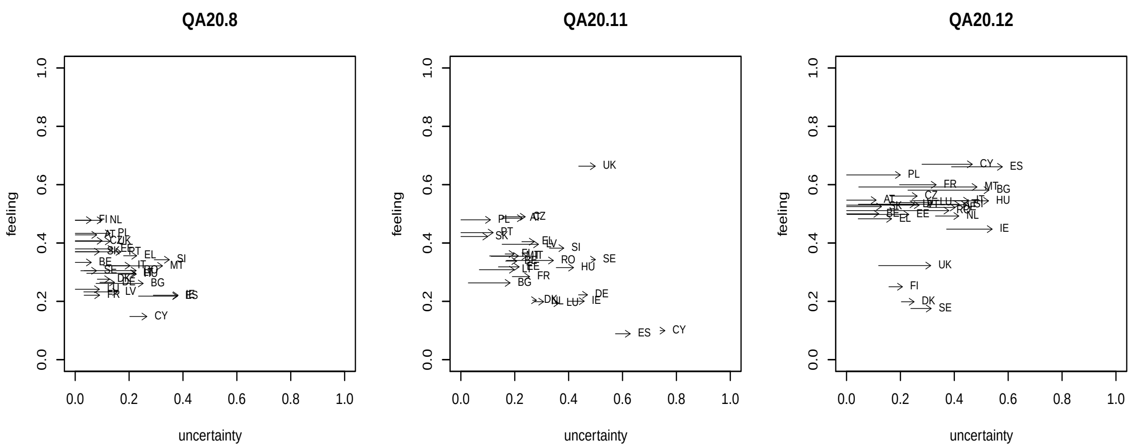

Figure 2: Questions QA20.8 (*Left*), QA20.11 (*Middle*), QA20.12 (*Right*): Adjustment of the uncertainty estimates for the *dk* responses; for each Country, an arrow is drawn from $(1-\hat{\pi}_0, 1-\hat{\xi}_0)$ to $(1-\hat{\pi}_{adj}, 1-\hat{\xi}_0)$.

kind of response, which often ends to be treated as a missing value. In this paper we start postulating that *dk* is a valid response to all extent, as it contains an important information about the uncertainty of the subject in formulating the requested response. For this reason, the information supplied by *dk* responses can be effectively exploited within a statistical model containing, beyond the usual parameter aimed at measuring the attractiveness toward the item (feeling), also a parameter accounting for the indecision in the response (uncertainty). In this framework, after conditioning on the state of mind of the respondent (able/unable to formulate an evaluation), we derive the marginal distribution of responses, accounting for the uncertainty of respondents who choose the *dk* option. As underlined above, the proposed probabilistic framework makes sense only if the resulting marginal distribution is able to properly summarise the parameters describing feeling and uncertainty in the conditional distributions, so the distributional assumption is a crucial choice. In this paper we decline our proposal in the context of CUB models, for two main reasons. Firstly, the marginal distribution obtained by applying our proposed probabilistic framework very efficiently summarises the parameters of the conditional distributions, so it allows us very interesting insights about the presence of a group of respondents unable to express an evaluation. Secondly, the procedure to correct the parameters obtained with listwise deletion of *dk* responses results to be extremely simple, as it consists only of multiplying the parameter to be corrected by the relative frequency of the expressed evaluations. Although deriving from a manifestly unpretentious computation, the proposed correction leads to a very reasonable treatment of *dk* responses. In fact, taking *dk* responses into account ends up in an increased estimate of the uncertainty in the population under study, and the increment is directly related to the estimated proportion of subjects unable to express an evaluation in the population. Some real data case studies clearly show that correcting for *dk* responses can be important in empirical analysis, especially when the aim of the study is the comparison of feeling and uncertainty over a multiplicity of items, or over multiple groups of people expressing evaluations on a single item of interest.

Acknowledgments

The authors thank the anonymous reviewers for their valuable comments, which greatly improved the paper; Prof. Domenico Piccolo for many interesting discussions and Prof. Gerhard Tutz for his stimulating plenary talk at the 2013 ASMOD Conference.

This research was partially funded by STAR project (University of Naples Federico II - CUP: E68C13000020003) and partially by a grant from the European Union Seventh Framework Programme *'Cooperation - Socio-Economic Sciences and Humanities'* (FP7-SSH/2007-2013); *'SYstemic Risk TOmography: Signals, Measurements, Transmission Channels, and Policy Interventions' - SYRTO* - Project ID: 320270.

References

Beatty, P., Herrmann, D., 1995. A framework for evaluating "don't know" responses in surveys, in: Proc of the section on Survey Research Methods. Am Stat Assoc, pp. 1005–1010.

Bishop, G.F., Tuchfarber, A.J., Oldendick, R.W., 1986. Opinions on fictitious issues: The pressure to answer survey questions. Public Opin Quart 50, 240–250.

Cannell, C.F., Oksenberg, L., Converse, J.M., 1979. Experiments in Interviewing Techniques. Inst. Social Research, The Univ. of Michigan.

Carpita, M., Manisera, M., 2011. On the imputation of missing data in surveys with likert-type scales. J Class 28, 93–112.

Converse, P., 1964. The nature of belief systems in mass publics, in: Apter, D. (Ed.), Idealogy and Discontent. Free Press, pp. 206–261.

D'Elia, A., Piccolo, D., 2005. A mixture model for preference data analysis. Comput Stat Data An 49, 917–934.

European Commission, 2013. Eurobarometer 78.1 (2012). Technical Report. TNS Opinion & Social, Brussels.

Feick, L.F., 2005. Latent class analysis of survey questions that include don't know responses. Public Opin Quart 53, 525–547.

Kroh, M., 2006. Taking don't knows as valid responses: A complete random imputation of missing data. Qual Quant 40, 225–244.

Krosnick, J.A., 2002. The causes of no-opinion responses to attitude measures in surveys: They are rarely what they appear to be, in: Groves, R., Dillman, D., Eltinge, J., Little, J. (Eds.), Survey nonresponse. Wiley, pp. 1005–1010.

Lietz, P., 2010. Research into questionnaire design - a summary of the literature. Int J Market Res 52, 249–272.

Meulman, J.J., Van der Kooij, A.J., Heiser, W.J., 2004. Principal components analysis with nonlinear optimal scaling transformations for ordinal and nominal data, in: Kaplan, D. (Ed.), Handbook of Quantitative Methods in the Social Sciences. Sage, Newbury Park, CA, pp. 49–70.

Piccolo, D., 2003. On the moments of a mixture of uniform and shifted binomial random variables. Quad Stat 5, 85–104.

Poe, G.S., Seeman, I., McLaughlin, J., Mehl, E., Dietz, M., 1988. Don't know boxes in factual questions in a mail questionnaire. Public Opin Quart 52, 212–222.

Rubin, D.B., 1987. Multiple Imputation for Nonresponse in Surveys. Wiley, NY.

Rubin, D.B., Stern, H.S., Vehovarc, V., 1995. Handling "don't know" survey responses: The case of the slovenian plebiscite. J Am Stat Assoc 90, 822–828.

Schafer, J.L., 1997. Analysis Of Incomplete Multivariate Data. Chapman & Hall, London.

Schafer, J.L., Graham, J.W., 2002. Missing data: our view of the state of the art. Psychol Methods 7, 147–177.

Schuman, H., Presser, S., 1996. Questions and Answers in Attitude Surveys Experiments on Question Form, Wording, and Context. Sage, Beverly Hills, CA.

Zuccolotto, P., 2011. Symbolic missing data imputation in principal component analysis. Stat Anal Data Min 4, 171–183.

Zuccolotto, P., 2012. Principal component analysis with interval imputed missing values. AStA Adv Stat Anal 96, 1–23.